STAT 7164 Multivariate Analysis

Homework 5

Chapter 12: PCA

Chapter 13: Factor Analysis

+8.5

Qiangian Yao
12531797

Homework 5 Completion

Question 1 PCA

y_1	y_2	y_3	y_4	y_5	y_6

$$R = \begin{bmatrix} 1.000 & & & & & \\ -0.5442 & 1.000 & & & & \\ 0.0215 & 0.8226 & 1.000 & & & \\ 0.3108 & -0.6638 & -0.5851 & 1.000 & & \\ -0.0100 & -0.5091 & -0.6195 & 0.9410 & 1.000 & \\ -0.0412 & 0.2089 & 0.1900 & 0.5499 & 0.6103 & 1.000 \end{bmatrix}$$

$$C_1 = \begin{pmatrix} 0.1134 \\ -0.4593 \\ -0.4324 \\ 0.5344 \\ 0.5025 \\ 0.1843 \end{pmatrix} \quad C_2 = \begin{pmatrix} 0.3446 \\ -0.4352 \\ -0.2193 \\ -0.1795 \\ -0.3108 \\ -0.6964 \end{pmatrix} \quad C_3 = \begin{pmatrix} 0.8099 \\ -0.0446 \\ 0.4925 \\ 0.1317 \\ -0.1302 \\ 0.2553 \end{pmatrix} \quad C_4 = \begin{pmatrix} 0.0215 \\ -0.2928 \\ -0.3966 \\ -0.4097 \\ -0.4168 \\ 0.6492 \end{pmatrix} \quad C_5 = \begin{pmatrix} -0.4400 \\ -0.6418 \\ 0.5193 \\ 0.2432 \\ -0.2587 \\ 0.0286 \end{pmatrix} \quad C_6 = \begin{pmatrix} 0.0324 \\ -0.3163 \\ 0.2601 \\ -0.6618 \\ 0.6211 \\ 0.0002 \end{pmatrix}$$

← eigenvectors of R

θ_1	θ_2	θ_3	θ_4	θ_5	θ_6
3.22553	1.65603	1.07631	0.04112	0.00092	0.00003

← eigenvalues of R

---

a.) Use an orthogonal matrix A to transform $\underline{y}$ to $\underline{z}$: $\underline{z} = A\underline{y}$

$$\underline{z}_i = A\underline{y}_i = \begin{bmatrix} a_{11}y_{i1} + a_{12}y_{i2} + \cdots + a_{1p}y_{ip} \\ a_{21}y_{i1} + a_{22}y_{i2} + \cdots + a_{2p}y_{ip} \\ \vdots \\ a_{p1}y_{i1} + a_{p2}y_{i2} + \cdots + a_{pp}y_{ip} \end{bmatrix} = \begin{bmatrix} z_{i1} \\ z_{i2} \\ \vdots \\ z_{ip} \end{bmatrix}$$

$\underline{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{ip} \end{bmatrix}$

$E^{-1}H$

$$A = C' = \begin{pmatrix} \underline{a}_1' \\ \underline{a}_2' \\ \vdots \\ \underline{a}_p' \end{pmatrix}$$

a_j is the jth normalized eigenvector of S.

Principle components (to find the linear combination of the variables in an observation vector that have maximum variance.

$$\underline{z}_1 = \underline{a}_1'\underline{y} = a_{11}y_1 + a_{12}y_2 + \cdots + a_{1p}y_p \rightarrow s_{z_1}^2 = \lambda_1$$
$$\underline{z}_2 = \underline{a}_2'\underline{y} = a_{21}y_1 + a_{22}y_2 + \cdots + a_{2p}y_p \rightarrow s_{z_2}^2 = \lambda_2$$
$$\vdots$$
$$z_p = \underline{a}_p'\underline{y} = a_{p1}y_1 + a_{p2}y_2 + \cdots + a_{pp}y_p \rightarrow s_{z_p}^2 = \lambda_p$$

Eigenvalues are the variances of the principle components, the proportion of variance explained by the first k principle component is:

$$\text{proportion of variance} = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_k}{\sum_{i=1}^{p} \lambda_i}$$

$$\sum_{i=1}^{6} \theta_i = 3.22553 + 1.65603 + 1.07631 + 0.04112 + 0.00092 + 0.00003 = 6$$

$$\frac{\theta_1}{\sum \theta_i} = \frac{3.22553}{6} = 0.53716$$

$$\frac{\theta_1 + \theta_2}{\sum \theta_i} = \frac{3.22553 + 1.65603}{6} = 0.8136$$

Thus, two principle components are necessary to explain 80% of the total sample Variance of the standardized variables.

b.) $Z_1 = \underline{c}_1' \underline{y}$

$= 0.11134\, y_1 - 0.4593 y_2 - 0.4324 y_3 + 0.5344 y_4 + 0.5025 y_5 + 0.1842 y_6$

Variances

$S^2_{z_1} = \theta_1 = 3.22553$

$Z_2 = \underline{c}_2' \underline{y}$

$= 0.3446 y_1 - 0.4352 y_2 - 0.2193 y_3 - 0.1195 y_4 - 0.3108 y_5 - 0.6964 y_6$

$S^2_{z_2} = \theta_2 = 1.65603$

c.) $Z_3 = \underline{c}_3' \underline{y}$

$= 0.8099\, y_1 - 0.0446 y_2 + 0.4925 y_3 + 0.1211 y_4 - 0.1302 y_5 + 0.2552 y_6$

$S^2_{z_3} = \theta_3 = 1.07631$

$Z_4 = \underline{c}_4' \underline{y}$

$= 0.0215 y_1 - 0.2928 y_2 - 0.3966 y_3 - 0.4091 y_4 - 0.4168 y_5 + 0.6492 y_6$

$S^2_{z_4} = \theta_4 = 0.04112$

$Z_5 = \underline{c}_5' \underline{y}$

$= -0.4400 y_1 - 0.6418 y_2 + 0.5113 y_3 + 0.2432 y_4 - 0.2581 y_5 + 0.0286 y_6$

$S^2_{z_5} = \theta_5 = 0.00092$

$Z_6 = \underline{c}_6' \underline{y}$

$= 0.0321 y_1 - 0.3163 y_2 + 0.2601 y_2 - 0.6618 y_4 + 0.6211 y_5 + 0.0002 y_6$

$S^2_{z_6} = \theta_6 = 0.00003$

amount of the total sample variance explained by each principle component

d.)

$$\frac{\theta_1}{\sum \theta_i} = \frac{3.22553}{6} = 0.53716$$

$$\frac{\theta_2}{\sum \theta_i} = \frac{1.65603}{6} = 0.27160$$

$$\frac{\theta_3}{\sum \theta_i} = \frac{1.07631}{6} = 0.11194$$

$$\frac{\theta_4}{\sum \theta_i} = \frac{0.04112}{6} = 6.853 \times 10^{-3}$$

$$\frac{\theta_5}{\sum \theta_i} = \frac{0.00092}{6} = 1.533 \times 10^{-4}$$

$$\frac{\theta_6}{\sum \theta_i} = \frac{0.00003}{6} = 5.0 \times 10^{-6}$$

Question 2. Factor Analysis

a.) Completion

Orthogonal two-factor Model

$\underline{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix}$

i)

$$\begin{aligned}
y_1 &= \lambda_{11} f_1 + \lambda_{12} f_2 + \epsilon_1 \\
y_2 &= \lambda_{21} f_1 + \lambda_{22} f_2 + \epsilon_2 \\
y_3 &= \lambda_{31} f_1 + \lambda_{32} f_2 + \epsilon_3 \\
y_4 &= \lambda_{41} f_1 + \lambda_{42} f_2 + \epsilon_4 \\
y_5 &= \lambda_{51} f_1 + \lambda_{52} f_2 + \epsilon_5 \\
y_6 &= \lambda_{61} f_1 + \lambda_{62} f_2 + \epsilon_6
\end{aligned}$$

(i = 1, 2, 3, 4, 5, 6; j = 1, 2)

Assumptions of the factor model:

$y_i = \lambda_{i1} f_1 + \lambda_{i2} f_2 + \epsilon_i$

$E(f_j) = 0$ for $j = 1, 2$

$Var(f_j) = 1$ for $j = 1, 2$

$Cov(f_j, f_k) = 0$ for $j \neq k$

$E(\epsilon_i) = 0$ for $i = 1, 2, 3, 4, 5, 6$

$Var(\epsilon_i) = \psi_i$ for $i = 1, 2, 3, 4, 5, 6$

$Cov(\epsilon_i, \epsilon_k) = 0$ for $i \neq k$

$Cov(\epsilon_i, f_j) = 0 \quad \forall i, j$

ii) $Var(y_i) = Var(\lambda_{i1} f_1 + \lambda_{i2} f_2 + \epsilon_i)$ ← $Var(ay) = a^2 Var(y)$

$= Var(\lambda_{i1} f_1) + Var(\lambda_{i2} f_2) + Var(\epsilon_i)$

$= \lambda_{i1}^2 Var(f_1) + \lambda_{i2}^2 Var(f_2) + Var(\epsilon_i)$

$= \lambda_{i1}^2 + \lambda_{i2}^2 + \psi_i$

$\therefore Var(y_i) = \lambda_{i1}^2 + \lambda_{i2}^2 + \psi_i$

iii) $Cov(y_i) = E[(y_i - \mu_i)(y_i - \mu_i)]$

$= E[y_i \cdot y_i]$

$= E[(\lambda_{i1} f_1 + \lambda_{i2} f_2 + \lambda_{i3} f_3)(\lambda_{i1} f_1 + \lambda_{i2} f_2 + \lambda_{i3} f_3)]$

$= E[\lambda_{i1}^2 f_1 f_1 + \lambda_{i1}\lambda_{i2} f_1 f_2 + \lambda_{i1}\lambda_{i3} f_1 f_3 +$
$\lambda_{i2}\lambda_{i1} f_2 f_1 + \lambda_{i2}^2 f_2 f_2 + \lambda_{i2}\lambda_{i3} f_2 f_3 +$
$\lambda_{i3}\lambda_{i1} f_3 f_1 + \lambda_{i3}\lambda_{i2} f_3 f_2 + \lambda_{i3}^2 f_3 f_3]$

$= \lambda_{i1}^2 Cov(f_1, f_1) + \lambda_{i1}\lambda_{i2} Cov(f_1, f_2) + \lambda_{i1}\lambda_{i3} Cov(f_1, f_3) +$
$\lambda_{i2}\lambda_{i1} Cov(f_2, f_1) + \lambda_{i2}^2 Cov(f_2, f_2) + \lambda_{i2}\lambda_{i3} Cov(f_2, f_3) +$
$\lambda_{i3}\lambda_{i1} Cov(f_3, f_1) + \lambda_{i3}\lambda_{i2} Cov(f_3, f_2) + \lambda_{i3}^2 Cov(f_3, f_3)$

← $Cov(f_i, f_j) = 0$ for $i \neq j$; $Cov(f_j) = 1$

$= \lambda_{i1}^2 + \lambda_{i2}^2 + \lambda_{i3}^2$ demenstrated by 3 factor model, two factor model same way.

$\therefore Cov(y_1) = \lambda_{11}^2 + \lambda_{12}^2$

(IV) $Cov(y_i, y_j)$

$= E[(y_i - \mu_i)(y_j - \mu_j)]$

$= E[y_i \cdot y_j]$

$= E[(\lambda_{i1} f_1 + \lambda_{i2} f_2)(\lambda_{j1} f_1 + \lambda_{j2} f_2)]$

$= E[\lambda_{i1}\lambda_{j1} f_1 f_1 + \lambda_{i1}\lambda_{j2} f_1 f_2 + \lambda_{i2} f_2 \lambda_{j1} f_1 + \lambda_{i2}\lambda_{j2} f_2 f_2]$

$= \lambda_{i1}\lambda_{j1} Cov(f_1, f_1) + \lambda_{i1}\lambda_{j2} Cov(f_1, f_2) + \lambda_{i2}\lambda_{j1} Cov(f_2, f_1) + \lambda_{i2}\lambda_{j2} Cov(f_2, f_2)$

$= \lambda_{i1}\lambda_{j1} + \lambda_{i2}\lambda_{j2}$

$\therefore Cov(y_1, y_6) = \lambda_{11}\lambda_{61} + \lambda_{12}\lambda_{62}$

b. principle component method of estimation of three-factor Model.

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{pmatrix} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \\ \lambda_{41} & \lambda_{42} & \lambda_{43} \\ \lambda_{51} & \lambda_{52} & \lambda_{53} \\ \lambda_{61} & \lambda_{62} & \lambda_{63} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \epsilon_4 \\ \epsilon_5 \\ \epsilon_6 \end{pmatrix}$$

$\underline{y}$ $\quad \Lambda \quad$ $\underline{f}$ $\quad \underline{\epsilon}$

$\underline{y} = \Lambda \underline{f} + \underline{\epsilon}$ Assumptions: $E(\underline{f}) = \underline{0}$

$Cov(\underline{f}) = \underline{I}$

$E(\underline{\epsilon}) = \underline{0}$

$Cov(\underline{\epsilon}) = \underline{\Psi} = diag(\Psi_1, \Psi_2, \dots \Psi_6)$

$Cov(\underline{f}, \underline{\epsilon}) = 0$

$$\begin{aligned} R &= Cov(\underline{y}) \\ &= Cov(\Lambda \underline{f} + \underline{\epsilon}) \\ &= Cov(\Lambda \underline{f}) + Cov(\underline{\epsilon}) \\ &= \Lambda Cov(f) \Lambda' + \Psi \\ &= \Lambda I \Lambda' + \Psi \\ &= \Lambda \Lambda' + \Psi \end{aligned}$$

We attempt to find an estimator $\hat{\Lambda}$.

In principle component approach, Ψ is neglected (at first)

$$\begin{aligned} R &= \hat{\Lambda} \hat{\Lambda}' \\ &= C D C' \\ &= C D^{1/2} D^{1/2} C' \\ &= (C D^{1/2})(C' D^{1/2})' \end{aligned} \qquad C = (\underline{c}_1\ \underline{c}_2\ \underline{c}_3\ \underline{c}_4\ \underline{c}_5\ \underline{c}_6) \quad D = diag(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$$

We define $C_1 = (\underline{c}_1, \underline{c}_2, \underline{c}_3)$ $D = diag(\theta_1, \theta_2, \theta_3)$

$$\begin{aligned} \hat{\Lambda} &= C_1 D_1^{1/2} \\ &= (\sqrt{\theta_1}\, \underline{c}_1 \quad \sqrt{\theta_2}\, \underline{c}_2 \quad \sqrt{\theta_3}\, \underline{c}_3) \end{aligned}$$

Finally complete approximation of S by defining: $\hat{\Psi}_i = s_{ii} - \sum_{j=1}^{m} \hat{\lambda}_{ij}^2 = s_{ii} - \hat{h}_i^2$

$$\begin{array}{c} \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{array}\begin{array}{c} \begin{array}{cccccc} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 \end{array} \\ \begin{bmatrix} \underline{s_{11}} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16} \\ s_{21} & \underline{s_{22}} & s_{23} & s_{24} & s_{25} & s_{26} \\ s_{31} & s_{32} & \underline{s_{33}} & s_{34} & s_{35} & s_{36} \\ s_{41} & s_{42} & s_{43} & \underline{s_{44}} & s_{45} & s_{46} \\ s_{51} & s_{52} & s_{53} & s_{54} & \underline{s_{55}} & s_{56} \\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & \underline{s_{66}} \end{bmatrix} \end{array}$$

$$\Sigma = \Lambda\Lambda' + \Psi$$

$$s_{ii} = \hat{h}_i^2 + \psi_i = \hat{\lambda}_{i1}^2 + \hat{\lambda}_{i2}^2 + \hat{\lambda}_{i3}^2 + \psi_i$$

s_{ii}: the variance of ith variable

$\hat{h}_i^2$: factors (communality)

ψ_i: specific variance

$\hat{\lambda}_{i1}^2 + \hat{\lambda}_{i2}^2 + \hat{\lambda}_{i3}^2$: A row of $\hat{\Lambda}$

Total variance: $tr(S) = s_{11} + s_{22} + s_{33} + s_{44} + s_{55} + s_{66}$

Variance due to jth factor: $\sum_{i=1}^{p} \lambda_{ij}^2 = \lambda_{1j}^2 + \lambda_{2j}^2 + \lambda_{3j}^2 + \lambda_{4j}^2 + \lambda_{5j}^2 + \lambda_{6j}^2 = \underline{\theta_j}$

A column of $\hat{\Lambda}$

proportion of total sample variance due to jth factor:

$\frac{\theta_j}{tr(S)}$ or $\frac{\theta_j}{p}$ by correlation matrix R

i) $\frac{\theta_1 + \theta_2 + \theta_3}{6} = \frac{3.22553 + 1.65603 + 1.07634}{6} = 0.993$ ✓

ii) $\hat{\Lambda} = C_1 P_1^{1/2} = (\sqrt{\theta_1}C_1 \quad \sqrt{\theta_2}C_2 \quad \sqrt{\theta_3}C_3)$

$$= \left[\sqrt{3.22553} \cdot \begin{pmatrix} 0.1734 \\ -0.4593 \\ -0.4324 \\ 0.5344 \\ 0.5025 \\ 0.1843 \end{pmatrix} \quad \sqrt{1.65603} \begin{pmatrix} 0.3446 \\ -0.4352 \\ -0.2193 \\ -0.1795 \\ -0.3108 \\ -0.6964 \end{pmatrix} \quad \sqrt{1.07634} \begin{pmatrix} 0.8099 \\ -0.0446 \\ 0.4925 \\ 0.1317 \\ -0.1302 \\ 0.2553 \end{pmatrix} \right]$$

$$= \begin{array}{c} \begin{array}{ccc} \text{factor 1} & \text{factor 2} & \text{factor 3} \end{array} \\ \begin{pmatrix} 0.3114 & 0.4435 & \boxed{0.8403} \\ \boxed{-0.8249} & -0.5601 & -0.0463 \\ \boxed{-0.17166} & -0.3594 & 0.5110 \\ \boxed{0.9598} & -0.2310 & 0.1366 \\ \boxed{0.9025} & -0.4000 & -0.1351 \\ 0.3310 & \boxed{-0.8962} & 0.2649 \end{pmatrix} \end{array}$$ ✓

iii) $\sum_{j=1}^{3} \lambda_{ij}^2 = \hat{h}_i^2 \rightarrow$ a row of $\hat{\Lambda}$

$\hat{h}_1^2 = 0.3114^2 + 0.4435^2 + 0.8403^2 = 0.9998$

$\hat{h}_2^2 = 0.8249^2 + 0.5601^2 + 0.0463^2 = 0.9963$

$\hat{h}_3^2 = 0.17166^2 + 0.3594^2 + 0.5110^2 = 0.9934$

$\hat{h}_4^2 = 0.9598^2 + 0.2310^2 + 0.1366^2 = 0.9932$

$\hat{h}_5^2 = 0.9025^2 + 0.4000^2 + 0.1351^2 = 0.9928$

$\hat{h}_6^2 = 0.3310^2 + 0.8962^2 + 0.2649^2 = 0.9829$

$\Sigma = 5.9584$

Variance explained by factors ✓

iv.) $S_{ii} = \hat{h}_i^2 + \psi_i$

$r_{ii} = \hat{h}_i^2 + \psi_i$

$\psi_1 = 1 - \hat{h}_1^2 = 1 - 0.9998 = 0.0002$

$\psi_2 = 1 - \hat{h}_2^2 = 1 - 0.9963 = 0.0037$

$\psi_3 = 1 - \hat{h}_3^2 = 1 - 0.9934 = 0.0066$ ✓

$\psi_4 = 1 - \hat{h}_4^2 = 1 - 0.9932 = 0.0068$

$\psi_5 = 1 - \hat{h}_5^2 = 1 - 0.9928 = 0.0072$

$\psi_6 = 1 - \hat{h}_6^2 = 1 - 0.9829 = 0.0171$

v.) $y_i = a_{i1}f_1 + a_{i2}f_2 + a_{i3}f_3 + \epsilon_i$

$y_3 = -0.4766 f_1 - 0.3594 f_2 + 0.5110 f_3 + \epsilon_3$ (0.0066)

-0.5

$\text{var}(\epsilon_3) = 0.0066$

vi.) Error Matrix: $E = R - (\hat{\Lambda}\hat{\Lambda}' + \hat{\Psi})$

$$\begin{pmatrix} 0.3114 & 0.4435 & 0.8463 \\ -0.8249 & -0.5601 & -0.0463 \\ -0.4766 & -0.3594 & 0.5110 \\ 0.9598 & -0.2310 & 0.1366 \\ 0.9025 & -0.4000 & -0.1351 \\ 0.3310 & -0.8962 & 0.2649 \end{pmatrix}_{6\times3} \quad \begin{pmatrix} 0.3114 & -0.8249 & -0.4766 & 0.9598 & 0.9025 & 0.3310 \\ 0.4435 & -0.5601 & -0.3594 & -0.2310 & -0.4000 & -0.8962 \\ 0.8463 & -0.0463 & 0.5110 & 0.1366 & -0.1351 & 0.2649 \end{pmatrix}_{3\times6}$$

only calculated at index

row: 2

column: 5

$(\hat{\Lambda}\hat{\Lambda}')_{25} = -0.8249 \times 0.9025 + 0.5601 \times 0.4000 + 0.0463 \times 0.1351 = -0.5142$

$\hat{\Psi}_{25} = 0$

$R_{25} = -0.5091$ ✓

$E_{25} = R_{25} - (\hat{\Lambda}\hat{\Lambda}'_{25} + \hat{\Psi}_{25})$

$= -0.5091 - (-0.5142 + 0)$

$= -0.5091 + 0.5142$

$\doteq 0.0051$

c.) $\hat{\Lambda}^* = \hat{\Lambda} \cdot T$

$$\begin{pmatrix} 0.3114 & 0.4435 & 0.8403 \\ -0.8249 & -0.5601 & -0.0463 \\ -0.1166 & -0.3594 & 0.5110 \\ 0.9598 & -0.2310 & 0.1366 \\ 0.9025 & -0.4000 & -0.1351 \\ 0.3310 & -0.8962 & 0.2649 \end{pmatrix}_{6\times3} \quad \begin{pmatrix} -0.1738 & 0.5163 & 0.2630 \\ 0.4454 & 0.7902 & -0.4210 \\ 0.4504 & 0.2086 & 0.8681 \end{pmatrix}_{3\times3}$$

$-0.1738 \times 0.3114 + 0.4454 \times 0.4435 + 0.4504 \times 0.8403 = 0.3350$

$\times -0.8249 + \quad -0.5601 \quad -0.0463 = 0.3680$

$-0.1166 \quad -0.3594 \quad 0.5110 = 0.6110$

$0.9598 \quad -0.2310 \quad 0.1366 = -0.1841$

$\vdots$

The result is calculated by numpy in python: numpy.dot($\hat{\Lambda}$, T)

$$\hat{\Lambda}^* = \begin{bmatrix} 0.3350 & 0.7052 & 0.5142 \\ 0.3680 & -0.9216 & -0.0186 \\ 0.6110 & -0.6250 & 0.3600 \\ -0.1841 & 0.3991 & 0.4601 \\ -0.9374 & 0.1758 & 0.2966 \\ -0.5360 & -0.4622 & 0.6184 \end{bmatrix}$$

ok

d.) $\sum_{j=1}^{3} \lambda_{ij}^2 = \hat{h}_i^2 \rightarrow$ a row of $\hat{\Lambda}^*$

$\hat{h}_1^2 = 0.9392$

$\hat{h}_2^2 = 0.9962$

$\hat{h}_3^2 = 0.9105$

$\hat{h}_4^2 = 0.9858$

$\hat{h}_5^2 = 0.9976$

$\hat{h}_6^2 = 0.9612$

$\Sigma = 5.8505$

ok

e.) $\hat{\Lambda}^* = \hat{\Lambda} \cdot Q$

$$\begin{pmatrix} 0.3114 & 0.4435 & 0.8403 \\ -0.8249 & -0.5601 & -0.0463 \\ -0.7166 & -0.3594 & 0.5110 \\ 0.9598 & -0.2310 & 0.1366 \\ 0.9025 & -0.4000 & -0.1351 \\ 0.3310 & -0.8962 & 0.2649 \end{pmatrix}_{6\times3} \quad \begin{pmatrix} -0.1210 & 0.4399 & 0.1443 \\ 0.5401 & 0.9081 & -0.3526 \\ 0.6249 & 0.3123 & 0.9682 \end{pmatrix}_{3\times3}$$

CODE: numpy.dot($\hat{\Lambda}$, Q)

$$\hat{\Lambda}^* = \begin{pmatrix} 0.5404 & 0.8022 & 0.1021 \\ 0.2630 & -0.8860 & 0.0336 \\ 0.6849 & -0.5084 & 0.5094 \\ -0.1316 & 0.2551 & 0.3522 \\ -0.9514 & -0.0084 & 0.1405 \\ -0.5511 & -0.5855 & 0.6202 \end{pmatrix}$$ ✓

f.) $\sum_{j=1}^{3} \lambda_{ij}^2 = \hat{h}_i^2 \rightarrow$ a row of $\hat{\Lambda}^*$:

$\hat{h}_1^2 = 1.4285$

$\hat{h}_2^2 = 0.8553$

$\hat{h}_3^2 = 0.9870$ ✓

$\hat{h}_4^2 = 0.1244$

$\hat{h}_5^2 = 0.9250$

$\hat{h}_6^2 = 1.0385$

$\Sigma = 5.9587$

g.) Covariance of factor vector $\underline{f}$

$\underline{f}^* = Q'\underline{f}$

$Cov(\underline{f}^*) = Cov(Q'\underline{f})$

$= Q' Cov(\underline{f}) Q$

$= Q'Q \neq I$

$$= \begin{pmatrix} -0.1210 & 0.5401 & 0.6249 \\ 0.4399 & 0.9081 & 0.3123 \\ 0.1443 & -0.3526 & 0.9682 \end{pmatrix} \begin{pmatrix} -0.1210 & 0.4399 & 0.1443 \\ 0.5401 & 0.9081 & -0.3526 \\ 0.6249 & 0.3123 & 0.9682 \end{pmatrix}$$

$$= \begin{pmatrix} 1.2021 & 0.3690 & 0.3103 \\ 0.3690 & 1.1151 & 0.0457 \\ 0.3103 & 0.0457 & 1.0826 \end{pmatrix}$$ ✓

h.)

$\hat{\Lambda}$ =

θ_1	θ_2	θ_3	$\hat{h}_i^2$
0.3114	0.4435	**0.8403**	0.9998
-0.8249	-0.5601	-0.0463	0.9963
-0.1166	-0.3594	0.5110	0.9934
0.9598	-0.2310	0.1366	0.9932
0.9025	-0.4000	-0.1351	0.9928
0.3310	**-0.8962**	0.2649	0.9829
			$\Sigma = 5.9584$

$\hat{\Lambda}^* = \hat{\Lambda}T$

			$\hat{h}_i^2$
0.3350	**0.1052**	0.5142	0.9392
0.3680	**-0.9216**	-0.0186	0.9962
0.6710	-0.6250	0.3600	0.9105
-0.1841	0.3991	0.4601	0.9858
-0.3714	0.1758	0.2966	0.9976
-0.5360	-0.4622	**0.6184**	0.9612
			$\Sigma = 5.8505$

$\hat{\Lambda}^* = \hat{\Lambda}Q$

			$\hat{h}_i^2$
0.5404	**0.8022**	0.1021	1.4285
0.2630	**-0.8860**	0.0336	0.8552
0.6849	-0.5289	0.5094	0.9810
-0.1316	0.2551	0.3522	0.1244
-0.9514	-0.0084	0.1405	0.9250
-0.5517	-0.5855	**0.6202**	1.0385
			$\Sigma = 5.9587$

There's No obvious improvement in interpretation of the factors After Loading rotated.

The goal of rotation is to obtain a simple structure of loadings in which each variable loads highly on only one factor, with small loadings on all other factors. ✓

value of 0.5 or 0.6 is used as threshhold to area of significance.

In $\Lambda^* = \hat{\Lambda}T$: the third row and sixth row have multiple large loadings

In $\Lambda^* = \hat{\Lambda}Q$: the first, third, sixth row have multiple large loadings, which mean that the variance of the corresponding y_i can not be explained mostly by 1 factor.

In $\hat{\Lambda}$: The pattern is Loading matrix without rotation is much better.

y_1 is explained by factor 3

y_2, y_3, y_4, y_5 are explained by factor 1

y_6 is explained by factor 2.

Question 3. Completion

page14 of output

a. PRINIT + PRIORS = SMC

$$\hat{\Lambda} = \begin{bmatrix} 0.31245 & -0.46333 & 0.82686 \\ -0.82192 & 0.56043 & -0.03093 \\ -0.77895 & 0.34537 & 0.52132 \\ 0.95869 & 0.24015 & 0.14845 \\ 0.90119 & 0.41668 & -0.11846 \\ 0.31755 & 0.85672 & 0.27341 \end{bmatrix} \quad \begin{matrix} 0.9960033 \\ 1.0004985 \\ 0.9978061 \\ 0.9987979 \\ 0.9998781 \\ 0.9095502 \end{matrix}$$

$\Sigma = 5.902534$

Factor Pattern			
	Factor1	Factor2	Factor3
redshift	0.31245	-0.46333	0.82686
lineflux	-0.82792	0.56043	-0.03093
luminosity	-0.77895	0.34537	0.52132
ab1450	0.95869	0.24015	0.14845
absmag	0.90119	0.41668	-0.11876
rfewidth	0.31755	0.85672	0.27341

$\hat{\Lambda}$

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
3.2219128	1.6133031	1.0673182

Final Communality Estimates: Total = 5.902534					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
0.9960033	1.0004985	0.9978061	0.9987979	0.9998781	0.9095502

Varimax + PRINIT

page21 of output

$$\hat{\Lambda} = \begin{bmatrix} -0.06152 & 0.00512 & 0.99821 \\ 0.86998 & -0.05430 & -0.49083 \\ 0.98986 & -0.07156 & 0.08855 \\ -0.56026 & 0.78066 & 0.27287 \\ -0.56180 & 0.82585 & -0.04885 \\ 0.26407 & 0.91573 & -0.05981 \end{bmatrix} \quad \begin{matrix} 1.0002424 \\ 1.0007257 \\ 0.9927907 \\ 0.9977780 \\ 1.0000291 \\ 0.9118729 \end{matrix}$$

$\Sigma = 5.903439$

Rotated Factor Pattern			
	Factor1	Factor2	Factor3
redshift	-0.06152	0.00512	0.99821
lineflux	0.86998	-0.05430	-0.49083
luminosity	0.98986	-0.07156	0.08855
ab1450	-0.56026	0.78066	0.27287
absmag	-0.56180	0.82585	-0.04885
rfewidth	0.26407	0.91573	-0.05981

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
2.4397192	2.1381129	1.3256071

HK + PRINIT

page29 of output

$$\hat{\Lambda} = \begin{bmatrix} 0.05864 & -0.00508 & 1.01495 \\ 0.86191 & 0.01224 & -0.35891 \\ 1.05857 & 0.01102 & 0.25605 \\ -0.39171 & 0.11402 & 0.18874 \\ -0.42297 & 0.11809 & -0.14183 \\ 0.48356 & 1.00528 & -0.01494 \end{bmatrix} \quad \begin{matrix} 1.0002424 \\ 1.0001254 \\ 0.9927907 \\ 0.9977780 \\ 1.0000291 \\ 0.9118729 \end{matrix}$$

$\Sigma = 5.903439$

Final Communality Estimates: Total = 5.903439					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1.0002427	1.0007257	0.9927907	0.9977780	1.0000291	0.9118729

Rotated Factor Pattern (Standardized Regression Coefficients)			
	Factor1	Factor2	Factor3
redshift	0.05864	-0.00508	1.01495
lineflux	0.86191	0.07227	-0.35897
luminosity	1.05857	0.07102	0.25605
ab1450	-0.39171	0.74021	0.18874
absmag	-0.42297	0.78809	-0.14183
rfewidth	0.48356	1.00528	-0.01494

$\hat{\Lambda}$ in Question 2. part b.

b).

$$\begin{pmatrix} 0.3114 & 0.4435 & 0.8402 \\ -0.8249 & -0.5601 & -0.0462 \\ -0.1166 & -0.3594 & 0.5110 \\ 0.9598 & -0.2310 & 0.1366 \\ 0.9625 & -0.4000 & -0.1351 \\ 0.3310 & -0.8960 & 0.2649 \end{pmatrix} \quad \begin{matrix} 0.9998 \\ 0.9963 \\ 0.9934 \\ 0.9932 \\ 0.9928 \\ 0.9829 \end{matrix}$$

$\Sigma = 5.9584$

I really don't see any obvious advantages of estimated Loadings in a. than the one calculated in Question 2.b.

Because All the four Loading matrix have similar pattern and each row have one large loading value. Thus, I would say that these Loading matrix have similar performance.

dataset

Obs	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1	2.81	-13.48	45.29	19.50	-26.27	117
2	3.07	-13.73	45.13	19.65	-26.26	82
3	3.45	-13.87	45.11	18.93	-27.17	33
4	3.19	-13.27	45.63	18.59	-27.39	92
5	3.07	-13.56	45.30	19.59	-26.32	114
6	4.15	-13.95	45.20	19.42	-26.97	50
7	3.26	-13.83	45.08	19.18	-26.83	43
8	2.81	-13.50	45.27	20.41	-25.36	259
9	3.83	-13.66	45.41	18.93	-27.34	58
10	3.32	-13.71	45.23	20.00	-26.04	126
11	2.81	-13.50	45.27	18.45	-27.32	42
12	4.40	-13.96	45.25	20.55	-25.94	146
13	3.45	-13.91	45.07	20.45	-25.65	124
14	3.70	-13.85	45.19	19.70	-26.51	75
15	3.07	-13.67	45.19	19.54	-26.37	85
16	4.34	-13.93	45.27	20.17	-26.29	109
17	3.00	-13.75	45.08	19.30	-26.58	55
18	3.88	-14.17	44.92	20.68	-25.61	91
19	3.07	-13.92	44.94	20.51	-25.41	116
20	4.08	-14.28	44.86	20.70	-25.67	75
21	3.62	-13.82	45.20	19.45	-26.73	63
22	3.07	-14.08	44.78	19.90	-26.02	46
23	2.94	-13.82	44.99	19.49	-26.35	55
24	3.20	-14.15	44.75	20.89	-25.09	99
25	3.24	-13.74	45.17	19.17	-26.83	53

prin as example13.3.1

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

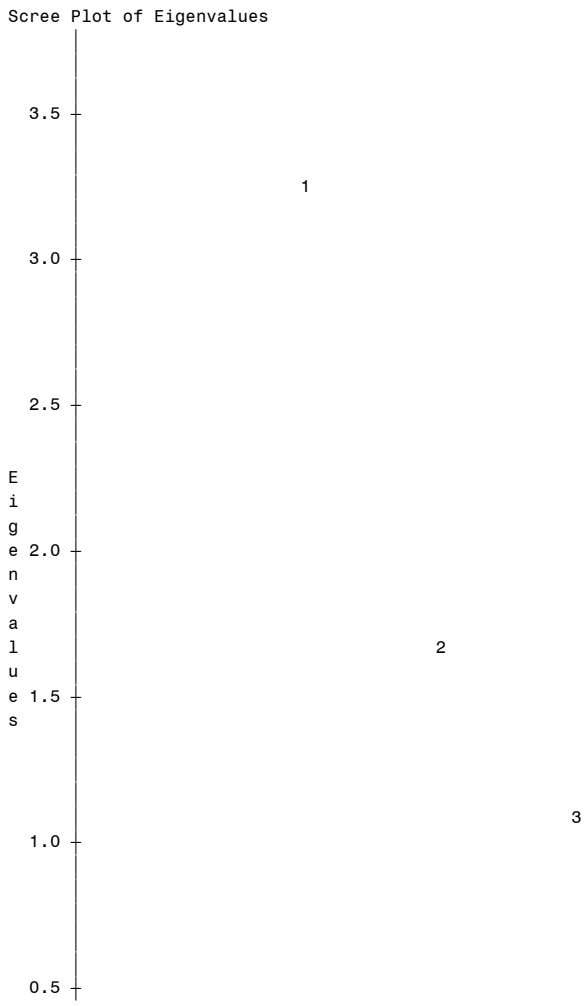
prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

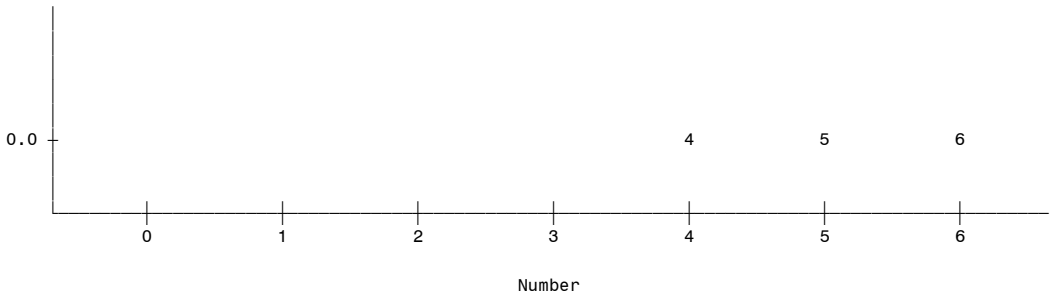
Eigenvalues of the Correlation Matrix: Total = 6 Average = 1				
	Eigenvalue	**Difference**	**Proportion**	**Cumulative**
1	3.22553194	1.56950063	0.5376	0.5376
2	1.65603131	0.57966002	0.2760	0.8136
3	1.07637129	1.03525276	0.1794	0.9930
4	0.04111853	0.04020171	0.0069	0.9998
5	0.00091682	0.00088671	0.0002	1.0000
6	0.00003011		0.0000	1.0000

3 factors will be retained by the NFACTOR criterion.



prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components



Eigenvectors			
	1	**2**	**3**
redshift	0.17344	-0.34462	0.80994
lineflux	-0.45925	0.43519	-0.04463
luminosity	-0.43238	0.27933	0.49246
ab1450	0.53436	0.17948	0.13166
absmag	0.50254	0.31076	-0.13022
rfewidth	0.18427	0.69645	0.25533

Factor Pattern			
	Factor1	**Factor2**	**Factor3**
redshift	0.31149	-0.44348	0.84030
lineflux	-0.82481	0.56003	-0.04630
luminosity	-0.77655	0.35946	0.51092
ab1450	0.95970	0.23096	0.13660
absmag	0.90256	0.39990	-0.13510
rfewidth	0.33095	0.89624	0.26490

Variance Explained by Each Factor		
Factor1	**Factor2**	**Factor3**
3.2255319	1.6560313	1.0763713

prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components

Final Communality Estimates: Total = 5.957935					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99980348	0.99609392	0.99328650	0.99303228	0.99278308	0.98293528

Residual Correlations With Uniqueness on the Diagonal						
	redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
redshift	0.00020	-0.00000	-0.00056	-0.00046	-0.00026	0.00056
lineflux	-0.00000	0.00391	0.00447	0.00480	0.00516	-0.00777
luminosity	-0.00056	0.00447	0.00671	0.00679	0.00668	-0.01049
ab1450	-0.00046	0.00480	0.00679	0.00697	0.00695	-0.01084
absmag	-0.00026	0.00516	0.00668	0.00695	0.00722	-0.01105
rfewidth	0.00056	-0.00777	-0.01049	-0.01084	-0.01105	0.01706

Root Mean Square Off-Diagonal Residuals: Overall = 0.00642747					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.00042529	0.00509995	0.00664879	0.00685904	0.00695288	0.00906026

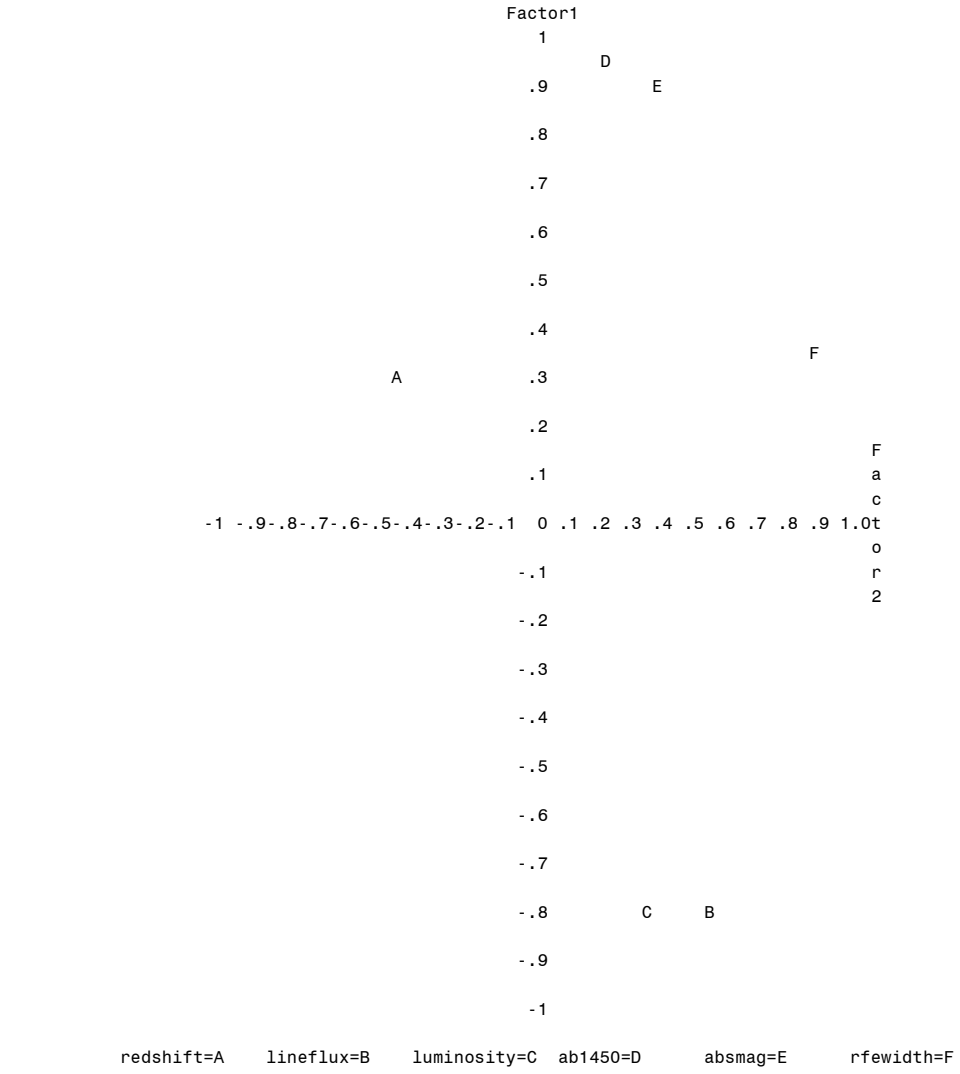
Partial Correlations Controlling Factors						
	redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
redshift	1.00000	-0.00003	-0.48641	-0.39361	-0.22101	0.30442
lineflux	-0.00003	1.00000	0.87243	0.91920	0.97273	-0.95206
luminosity	-0.48641	0.87243	1.00000	0.99275	0.95950	-0.98012
ab1450	-0.39361	0.91920	0.99275	1.00000	0.98020	-0.99455
absmag	-0.22101	0.97273	0.95950	0.98020	1.00000	-0.99551
rfewidth	0.30442	-0.95206	-0.98012	-0.99455	-0.99551	1.00000

Root Mean Square Off-Diagonal Partials: Overall = 0.80826925					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.32650866	0.83170802	0.87915198	0.88716568	0.87949034	0.88767929

prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components

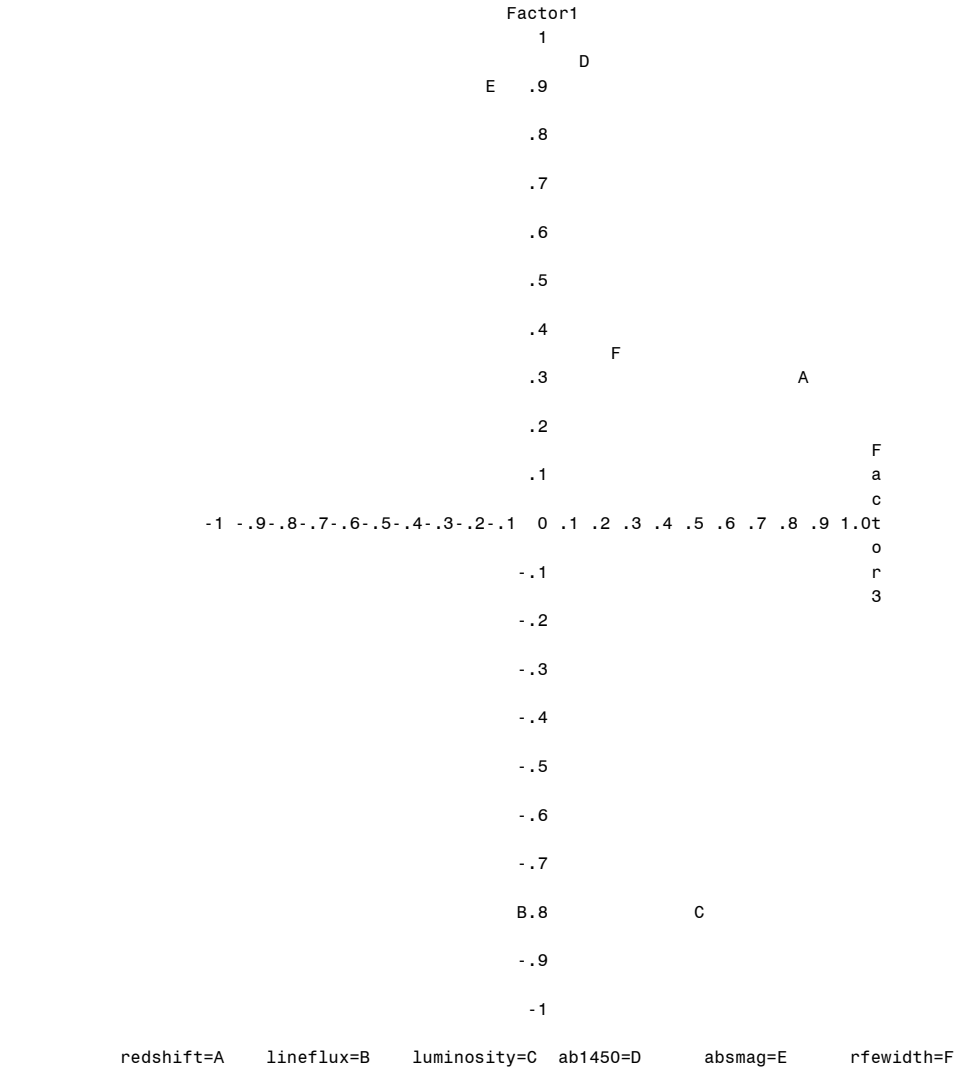
Plot of Factor Pattern for Factor1 and Factor2



prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components

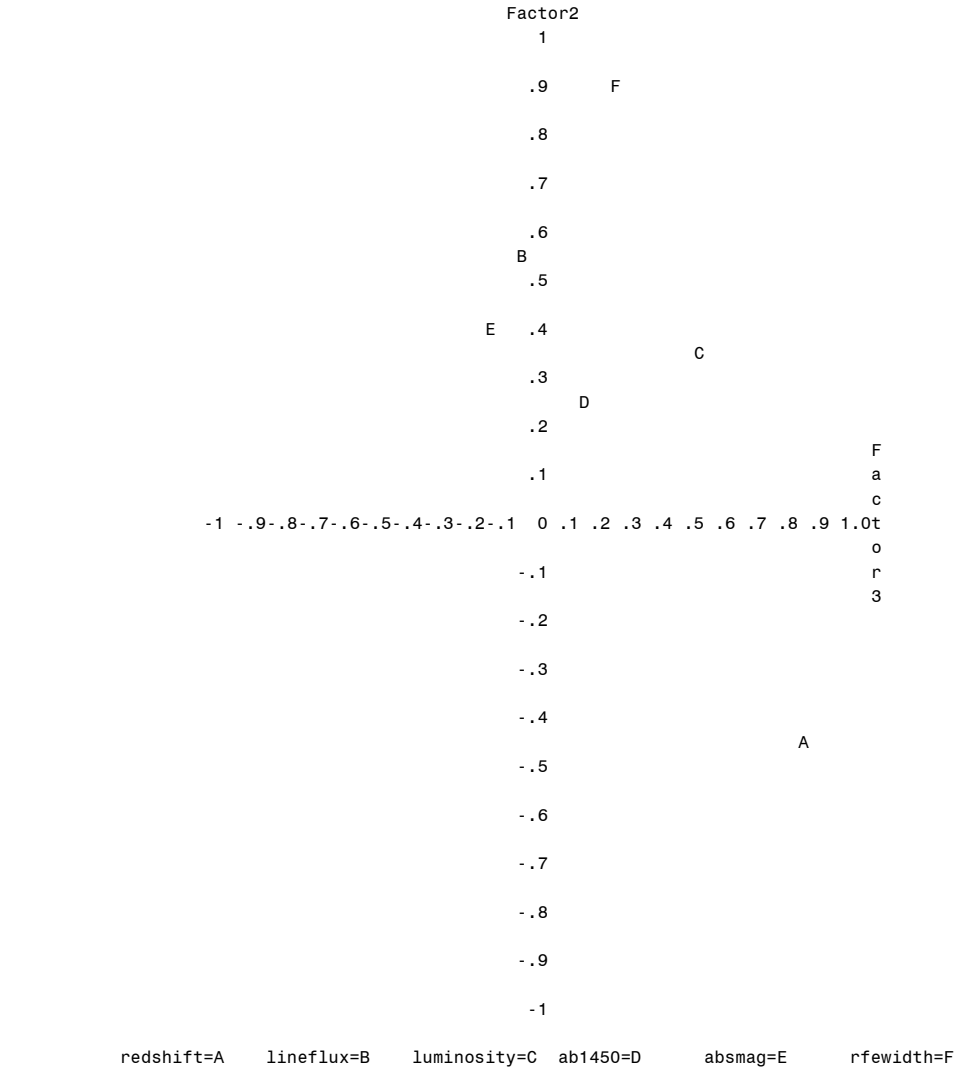
Plot of Factor Pattern for Factor1 and Factor3



prin as example13.3.1

The FACTOR Procedure
Initial Factor Method: Principal Components

Plot of Factor Pattern for Factor2 and Factor3



prin + priors=smc

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prin + priors=smc

The FACTOR Procedure
Initial Factor Method: Principal Factors

Prior Communality Estimates: SMC					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99595959	0.99973499	0.99960684	0.99993158	0.99992389	0.91190712

Eigenvalues of the Reduced Correlation Matrix: Total = 5.90706401 Average = 0.98451067				
	Eigenvalue	**Difference**	**Proportion**	**Cumulative**
1	3.22236811	1.60836678	0.5455	0.5455
2	1.61400132	0.54629561	0.2732	0.8187
3	1.06770572	1.06382769	0.1808	0.9995
4	0.00387802	0.00396454	0.0007	1.0002
5	-.00008652	0.00071611	-0.0000	1.0001
6	-.00080264		-0.0001	1.0000

3 factors will be retained by the NFACTOR criterion.

Eigenvectors			
	1	**2**	**3**
redshift	0.17401	-0.36428	0.80051
lineflux	-0.46111	0.44107	-0.03049
luminosity	-0.43403	0.27233	0.50455
ab1450	0.53416	0.18902	0.14362
absmag	0.50204	0.32777	-0.11521
rfewidth	0.17700	0.67486	0.26416

Factor Pattern			
	Factor1	**Factor2**	**Factor3**
redshift	0.31237	-0.46279	0.82717
lineflux	-0.82775	0.56035	-0.03151
luminosity	-0.77913	0.34597	0.52135
ab1450	0.95888	0.24014	0.14841
absmag	0.90122	0.41641	-0.11905
rfewidth	0.31772	0.85737	0.27296

prin + priors=smc

The FACTOR Procedure
Initial Factor Method: Principal Factors

Variance Explained by Each Factor		
Factor1	**Factor2**	**Factor3**
3.2223681	1.6140013	1.0677057

Final Communality Estimates: Total = 5.904075					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.9959538	1.0001448	0.9985480	0.9991350	0.9997579	0.9105357

prinit priors=smc

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prinit priors=smc

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Prior Communality Estimates: SMC					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
0.99595959	0.99973499	0.99960684	0.99993158	0.99992389	0.91190712

Preliminary Eigenvalues: Total = 5.90706401 Average = 0.98451067				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22236811	1.60836678	0.5455	0.5455
2	1.61400132	0.54629561	0.2732	0.8187
3	1.06770572	1.06382769	0.1808	0.9995
4	0.00387802	0.00396454	0.0007	1.0002
5	-.00008652	0.00071611	-0.0000	1.0001
6	-.00080264		-0.0001	1.0000

3 factors will be retained by the NFACTOR criterion.

Warning: Too many factors for a unique solution.

Iteration	Change	Communalities					
1	0.0014	0.99595	1.00000	0.99855	0.99913	0.99976	0.91054
2	0.0010	0.99600	1.00000	0.99781	0.99880	0.99988	0.90955

Convergence criterion satisfied.

Eigenvalues of the Reduced Correlation Matrix: Total = 5.90393039 Average = 0.9839884				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22191281	1.60860967	0.5457	0.5457
2	1.61330315	0.54598499	0.2733	0.8190
3	1.06731816	1.06444752	0.1808	0.9998
4	0.00287064	0.00341657	0.0005	1.0002
5	-.00054593	0.00038250	-0.0001	1.0002
6	-.00092843		-0.0002	1.0000

prinit priors=smc

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Eigenvectors			
	1	2	3
redshift	0.17407	-0.36478	0.80036
lineflux	-0.46125	0.44123	-0.02994
luminosity	-0.43396	0.27191	0.50461
ab1450	0.53410	0.18907	0.14369
absmag	0.50207	0.32806	-0.11495
rfewidth	0.17691	0.67449	0.26465

Factor Pattern			
	Factor1	Factor2	Factor3
redshift	0.31245	-0.46333	0.82686
lineflux	-0.82792	0.56043	-0.03093
luminosity	-0.77895	0.34537	0.52132
ab1450	0.95869	0.24015	0.14845
absmag	0.90119	0.41668	-0.11876
rfewidth	0.31755	0.85672	0.27341

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
3.2219128	1.6133031	1.0673182

Final Communality Estimates: Total = 5.902534					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
0.9960033	1.0004985	0.9978061	0.9987979	0.9998781	0.9095502

prin + varimax rotation

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prin + varimax rotation

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 6 Average = 1				
	Eigenvalue	**Difference**	**Proportion**	**Cumulative**
1	3.22553194	1.56950063	0.5376	0.5376
2	1.65603131	0.57966002	0.2760	0.8136
3	1.07637129	1.03525276	0.1794	0.9930
4	0.04111853	0.04020171	0.0069	0.9998
5	0.00091682	0.00088671	0.0002	1.0000
6	0.00003011		0.0000	1.0000

3 factors will be retained by the NFACTOR criterion.

Factor Pattern			
	Factor1	**Factor2**	**Factor3**
redshift	0.31149	-0.44348	0.84030
lineflux	-0.82481	0.56003	-0.04630
luminosity	-0.77655	0.35946	0.51092
ab1450	0.95970	0.23096	0.13660
absmag	0.90256	0.39990	-0.13510
rfewidth	0.33095	0.89624	0.26490

Variance Explained by Each Factor		
Factor1	**Factor2**	**Factor3**
3.2255319	1.6560313	1.0763713

Final Communality Estimates: Total = 5.957935					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99980348	0.99609392	0.99328650	0.99303228	0.99278308	0.98293528

prin + varimax rotation

The FACTOR Procedure
Rotation Method: Varimax

Orthogonal Transformation Matrix			
	1	**2**	**3**
1	-0.77380	0.57626	0.26299
2	0.44540	0.79019	-0.42097
3	0.45040	0.20861	0.86812

Rotated Factor Pattern			
	Factor1	**Factor2**	**Factor3**
redshift	-0.06009	0.00436	0.99809
lineflux	0.86682	-0.04243	-0.49287
luminosity	0.99111	-0.05687	0.08800
ab1450	-0.57822	0.76404	0.27374
absmag	-0.58112	0.80793	-0.04827
rfewidth	0.26241	0.95417	-0.06029

Variance Explained by Each Factor		
Factor1	**Factor2**	**Factor3**
2.4781963	2.1520040	1.3277343

Final Communality Estimates: Total = 5.957935					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99980348	0.99609392	0.99328650	0.99303228	0.99278308	0.98293528

prinit + varimax rotation

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prinit + varimax rotation

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Prior Communality Estimates: ONE

Preliminary Eigenvalues: Total = 6 Average = 1				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22553194	1.56950063	0.5376	0.5376
2	1.65603131	0.57966002	0.2760	0.8136
3	1.07637129	1.03525276	0.1794	0.9930
4	0.04111853	0.04020171	0.0069	0.9998
5	0.00091682	0.00088671	0.0002	1.0000
6	0.00003011		0.0000	1.0000

3 factors will be retained by the NFACTOR criterion.

Warning: Too many factors for a unique solution.

Iteration	Change	Communalities					
1	0.0171	0.99980	0.99609	0.99329	0.99303	0.99278	0.98294
2	0.0126	1.00000	0.99522	0.98996	0.99119	0.99074	0.97029
3	0.0100	1.00000	0.99557	0.98836	0.99128	0.99080	0.96031
4	0.0082	1.00000	0.99639	0.98772	0.99202	0.99167	0.95213
5	0.0069	1.00000	0.99731	0.98760	0.99290	0.99278	0.94527
6	0.0058	1.00000	0.99818	0.98778	0.99372	0.99389	0.93946
7	0.0049	1.00000	0.99896	0.98812	0.99445	0.99493	0.93452
8	0.0042	1.00000	0.99962	0.98856	0.99507	0.99585	0.93031
9	0.0036	1.00000	1.00000	0.98904	0.99559	0.99665	0.92671
10	0.0031	1.00000	1.00000	0.98954	0.99603	0.99735	0.92365
11	0.0026	1.00000	1.00000	0.99007	0.99642	0.99795	0.92105
12	0.0022	1.00000	1.00000	0.99060	0.99675	0.99846	0.91885
13	0.0019	1.00000	1.00000	0.99110	0.99703	0.99889	0.91697
14	0.0016	1.00000	1.00000	0.99158	0.99726	0.99925	0.91538
15	0.0014	1.00000	1.00000	0.99202	0.99746	0.99956	0.91402
16	0.0012	1.00000	1.00000	0.99243	0.99763	0.99981	0.91286
17	0.0010	1.00000	1.00000	0.99279	0.99778	1.00000	0.91187

prinit + varimax rotation

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Convergence criterion satisfied.

Eigenvalues of the Reduced Correlation Matrix: Total = 5.90273507 Average = 0.98378918				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22054393	1.60613811	0.5456	0.5456
2	1.61440582	0.54591640	0.2735	0.8191
3	1.06848942	1.06624710	0.1810	1.0001
4	0.00224232	0.00308888	0.0004	1.0005
5	-.00084655	0.00125331	-0.0001	1.0004
6	-.00209986		-0.0004	1.0000

Factor Pattern			
	Factor1	Factor2	Factor3
redshift	0.31321	-0.46653	0.82734
lineflux	-0.82797	0.56075	-0.02757
luminosity	-0.77728	0.34307	0.52051
ab1450	0.95859	0.23898	0.14758
absmag	0.90145	0.41637	-0.11853
rfewidth	0.31821	0.85682	0.27655

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
3.2205439	1.6144058	1.0684894

Final Communality Estimates: Total = 5.903439					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1.0002427	1.0007257	0.9927907	0.9977780	1.0000291	0.9118729

prinit + varimax rotation

The FACTOR Procedure
Rotation Method: Varimax

Orthogonal Transformation Matrix			
	1	2	3
1	-0.76647	0.58573	0.26353
2	0.44263	0.77902	-0.44408
3	0.46541	0.22373	0.85635

Rotated Factor Pattern			
	Factor1	Factor2	Factor3
redshift	-0.06152	0.00512	0.99821
lineflux	0.86998	-0.05430	-0.49083
luminosity	0.98986	-0.07156	0.08855
ab1450	-0.56026	0.78066	0.27287
absmag	-0.56180	0.82585	-0.04885
rfewidth	0.26407	0.91573	-0.05981

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
2.4397192	2.1381129	1.3256071

Final Communality Estimates: Total = 5.903439					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1.0002427	1.0007257	0.9927907	0.9977780	1.0000291	0.9118729

prin + hk rotation

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations

	redshift	lineflux	luminosity	ab1450	absmag	rfewidth
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prin + hk rotation

The FACTOR Procedure
Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 6 Average = 1				
	Eigenvalue	**Difference**	**Proportion**	**Cumulative**
1	3.22553194	1.56950063	0.5376	0.5376
2	1.65603131	0.57966002	0.2760	0.8136
3	1.07637129	1.03525276	0.1794	0.9930
4	0.04111853	0.04020171	0.0069	0.9998
5	0.00091682	0.00088671	0.0002	1.0000
6	0.00003011		0.0000	1.0000

3 factors will be retained by the NFACTOR criterion.

Factor Pattern			
	Factor1	**Factor2**	**Factor3**
redshift	0.31149	-0.44348	0.84030
lineflux	-0.82481	0.56003	-0.04630
luminosity	-0.77655	0.35946	0.51092
ab1450	0.95970	0.23096	0.13660
absmag	0.90256	0.39990	-0.13510
rfewidth	0.33095	0.89624	0.26490

Variance Explained by Each Factor		
Factor1	**Factor2**	**Factor3**
3.2255319	1.6560313	1.0763713

Final Communality Estimates: Total = 5.957935					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99980348	0.99609392	0.99328650	0.99303228	0.99278308	0.98293528

prin + hk rotation

The FACTOR Procedure
Rotation Method: Harris-Kaiser (hkpower = 0)

Oblique Transformation Matrix			
	1	**2**	**3**
1	-0.72100	0.43994	0.14428
2	0.54071	0.90806	-0.35255
3	0.62487	0.31225	0.96821

Inter-Factor Correlations			
	Factor1	**Factor2**	**Factor3**
Factor1	1.00000	-0.31950	-0.27322
Factor2	-0.31950	1.00000	0.04942
Factor3	-0.27322	0.04942	1.00000

Rotated Factor Pattern (Standardized Regression Coefficients)			
	Factor1	**Factor2**	**Factor3**
redshift	0.06071	-0.00329	1.01488
lineflux	0.86856	0.13122	-0.36127
luminosity	1.07352	0.14431	0.25592
ab1450	-0.48170	0.67459	0.18929
absmag	-0.51893	0.71802	-0.14157
rfewidth	0.41152	1.04214	-0.01174

Reference Axis Correlations			
	Factor1	**Factor2**	**Factor3**
Factor1	1.00000	0.31849	0.27200
Factor2	0.31849	1.00000	0.04155
Factor3	0.27200	0.04155	1.00000

prin + hk rotation

The FACTOR Procedure
Rotation Method: Harris-Kaiser (hkpower = 0)

Reference Structure (Semipartial Correlations)			
	Factor1	**Factor2**	**Factor3**
redshift	0.05536	-0.00311	0.97542
lineflux	0.79201	0.12424	-0.34723
luminosity	0.97889	0.13663	0.24597
ab1450	-0.43924	0.63868	0.18194
absmag	-0.47319	0.67980	-0.13606
rfewidth	0.37525	0.98667	-0.01128

Variance Explained by Each Factor Eliminating Other Factors		
Factor1	**Factor2**	**Factor3**
2.1462283	1.8776603	1.1842507

Factor Structure (Correlations)			
	Factor1	**Factor2**	**Factor3**
redshift	-0.21553	0.02747	0.99813
lineflux	0.92535	-0.16414	-0.59210
luminosity	0.95749	-0.18603	-0.03026
ab1450	-0.74895	0.83785	0.35424
absmag	-0.70966	0.87682	0.03570
rfewidth	0.08176	0.91008	-0.07267

Variance Explained by Each Factor Ignoring Other Factors		
Factor1	**Factor2**	**Factor3**
2.8907277	2.3613607	1.4797995

Final Communality Estimates: Total = 5.957935					
redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
0.99980348	0.99609392	0.99328650	0.99303228	0.99278308	0.98293528

prinit + hk rotation

The FACTOR Procedure

Input Data Type	Raw Data
Number of Records Read	25
Number of Records Used	25
N for Significance Tests	25

Correlations						
	redshift	**lineflux**	**luminosity**	**ab1450**	**absmag**	**rfewidth**
redshift	1.00000	-0.54419	0.02747	0.31083	-0.01000	-0.07123
lineflux	-0.54419	1.00000	0.82263	-0.66375	-0.50906	0.20892
luminosity	0.02747	0.82263	1.00000	-0.58566	-0.61948	0.19001
ab1450	0.31083	-0.66375	-0.58566	1.00000	0.94705	0.54995
absmag	-0.01000	-0.50906	-0.61948	0.94705	1.00000	0.61027
rfewidth	-0.07123	0.20892	0.19001	0.54995	0.61027	1.00000

prinit + hk rotation

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Prior Communality Estimates: ONE

Preliminary Eigenvalues: Total = 6 Average = 1				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22553194	1.56950063	0.5376	0.5376
2	1.65603131	0.57966002	0.2760	0.8136
3	1.07637129	1.03525276	0.1794	0.9930
4	0.04111853	0.04020171	0.0069	0.9998
5	0.00091682	0.00088671	0.0002	1.0000
6	0.00003011		0.0000	1.0000

3 factors will be retained by the NFACTOR criterion.

Warning: Too many factors for a unique solution.

Iteration	Change	Communalities					
1	0.0171	0.99980	0.99609	0.99329	0.99303	0.99278	0.98294
2	0.0126	1.00000	0.99522	0.98996	0.99119	0.99074	0.97029
3	0.0100	1.00000	0.99557	0.98836	0.99128	0.99080	0.96031
4	0.0082	1.00000	0.99639	0.98772	0.99202	0.99167	0.95213
5	0.0069	1.00000	0.99731	0.98760	0.99290	0.99278	0.94527
6	0.0058	1.00000	0.99818	0.98778	0.99372	0.99389	0.93946
7	0.0049	1.00000	0.99896	0.98812	0.99445	0.99493	0.93452
8	0.0042	1.00000	0.99962	0.98856	0.99507	0.99585	0.93031
9	0.0036	1.00000	1.00000	0.98904	0.99559	0.99665	0.92671
10	0.0031	1.00000	1.00000	0.98954	0.99603	0.99735	0.92365
11	0.0026	1.00000	1.00000	0.99007	0.99642	0.99795	0.92105
12	0.0022	1.00000	1.00000	0.99060	0.99675	0.99846	0.91885
13	0.0019	1.00000	1.00000	0.99110	0.99703	0.99889	0.91697
14	0.0016	1.00000	1.00000	0.99158	0.99726	0.99925	0.91538
15	0.0014	1.00000	1.00000	0.99202	0.99746	0.99956	0.91402
16	0.0012	1.00000	1.00000	0.99243	0.99763	0.99981	0.91286
17	0.0010	1.00000	1.00000	0.99279	0.99778	1.00000	0.91187

prinit + hk rotation

The FACTOR Procedure
Initial Factor Method: Iterated Principal Factor Analysis

Convergence criterion satisfied.

Eigenvalues of the Reduced Correlation Matrix: Total = 5.90273507 Average = 0.98378918				
	Eigenvalue	Difference	Proportion	Cumulative
1	3.22054393	1.60613811	0.5456	0.5456
2	1.61440582	0.54591640	0.2735	0.8191
3	1.06848942	1.06624710	0.1810	1.0001
4	0.00224232	0.00308888	0.0004	1.0005
5	-.00084655	0.00125331	-0.0001	1.0004
6	-.00209986		-0.0004	1.0000

Factor Pattern			
	Factor1	Factor2	Factor3
redshift	0.31321	-0.46653	0.82734
lineflux	-0.82797	0.56075	-0.02757
luminosity	-0.77728	0.34307	0.52051
ab1450	0.95859	0.23898	0.14758
absmag	0.90145	0.41637	-0.11853
rfewidth	0.31821	0.85682	0.27655

Variance Explained by Each Factor		
Factor1	Factor2	Factor3
3.2205439	1.6144058	1.0684894

Final Communality Estimates: Total = 5.903439					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1.0002427	1.0007257	0.9927907	0.9977780	1.0000291	0.9118729

prinit + hk rotation

The FACTOR Procedure
Rotation Method: Harris-Kaiser (hkpower = 0)

Oblique Transformation Matrix			
	1	2	3
1	-0.65858	0.50402	0.14420
2	0.59695	0.88803	-0.38014
3	0.65681	0.30381	0.95781

Inter-Factor Correlations			
	Factor1	Factor2	Factor3
Factor1	1.00000	-0.34410	-0.27545
Factor2	-0.34410	1.00000	0.07354
Factor3	-0.27545	0.07354	1.00000

Rotated Factor Pattern (Standardized Regression Coefficients)			
	Factor1	Factor2	Factor3
redshift	0.05864	-0.00508	1.01495
lineflux	0.86191	0.07227	-0.35897
luminosity	1.05857	0.07102	0.25605
ab1450	-0.39171	0.74021	0.18874
absmag	-0.42297	0.78809	-0.14183
rfewidth	0.48356	1.00528	-0.01494

Reference Axis Correlations			
	Factor1	Factor2	Factor3
Factor1	1.00000	0.33779	0.26714
Factor2	0.33779	1.00000	0.02354
Factor3	0.26714	0.02354	1.00000

prinit + hk rotation

The FACTOR Procedure
Rotation Method: Harris-Kaiser (hkpower = 0)

Reference Structure (Semipartial Correlations)			
	Factor1	Factor2	Factor3
redshift	0.05306	-0.00477	0.97541
lineflux	0.77986	0.06784	-0.34498
luminosity	0.95781	0.06667	0.24608
ab1450	-0.35442	0.69481	0.18139
absmag	-0.38271	0.73976	-0.13630
rfewidth	0.43753	0.94363	-0.01436

Variance Explained by Each Factor Eliminating Other Factors		
Factor1	Factor2	Factor3
1.9919052	1.9295219	1.1826829

Factor Structure (Correlations)			
	Factor1	Factor2	Factor3
redshift	-0.21919	0.04938	0.99842
lineflux	0.93592	-0.25071	-0.59107
luminosity	0.96360	-0.27440	-0.03032
ab1450	-0.69840	0.88887	0.35107
absmag	-0.65509	0.92321	0.03264
rfewidth	0.14176	0.83779	-0.07421

Variance Explained by Each Factor Ignoring Other Factors		
Factor1	Factor2	Factor3
2.7895092	2.4848865	1.4769460

prinit + hk rotation

The FACTOR Procedure
Rotation Method: Harris-Kaiser (hkpower = 0)

Final Communality Estimates: Total = 5.903439					
redshift	lineflux	luminosity	ab1450	absmag	rfewidth
1.0002427	1.0007257	0.9927907	0.9977780	1.0000291	0.9118729

```
ods rtf file='hw5_3 factors 20231207_V4.rtf';

data quasars;
infile 'quasar.csv' firstobs=2 dsd;
input redshift lineflux luminosity ab1450 absmag rfewidth;
RUN;

title color="purple" height=25pt bold italic underlin=1 "dataset";
proc print data=quasars;
run;


title color="purple" height=18pt bold italic underlin=1 "prin as example13.3.1";
proc factor data=quasars method=prin nfactors=3 corr scree ev residuals plot;
var redshift lineflux luminosity ab1450 absmag rfewidth;
run;



title color="purple" height=18pt bold italic underlin=1 "prin + priors=smc";
proc factor data=quasars method=prin nfactors=3 priors=smc corr ev;
var redshift lineflux luminosity ab1450 absmag rfewidth;
run;


title color="purple" height=18pt bold italic underlin=1 "prinit priors=smc";
proc factor data=quasars method=prinit nfactors=3 priors=smc heywood corr ev;
var redshift lineflux luminosity ab1450 absmag rfewidth;
run;


title color="purple" height=18pt bold italic underlin=1 "prin + varimax rotation";
proc factor data=quasars method=prin nfactors=3 rotate=varimax corr;
  var redshift lineflux luminosity ab1450 absmag rfewidth;
  run;

title color="purple" height=18pt bold italic underlin=1 "prinit + varimax rotation";
proc factor data=quasars method=prinit nfactors=3 rotate=varimax heywood corr;
  var redshift lineflux luminosity ab1450 absmag rfewidth;
  run;


title color="purple" height=18pt bold italic underlin=1 "prin + hk rotation";
proc factor data=quasars method=prin nfactors=3 rotate=hk corr;
var redshift lineflux luminosity ab1450 absmag rfewidth;
run;

title color="purple" height=18pt bold italic underlin=1 "prinit + hk rotation";
proc factor data=quasars method=prinit nfactors=3 rotate=hk heywood corr;
var redshift lineflux luminosity ab1450 absmag rfewidth;
run;

ods rtf close;
```