More on Estimation of Factor Scores

In this document I show that there is a simple relationship between the standardized factor scoring coefficients and the factor loadings and eigenvalues for the important special case of Principal Factoring without uniqueness (PRIORS = ONE in SAS). For convenience and to make my discussion self-contained, I collect the specifications of the Factor Analysis Model:

The Factor Analysis Model

$$\left\{\begin{array}{ccccc} X_1 & = & \lambda_{11}\xi_1 + \lambda_{12}\xi_2 + \cdots + \lambda_{1m}\xi_m & + & \varepsilon_1 \\ X_2 & = & \lambda_{21}\xi_1 + \lambda_{22}\xi_2 + \cdots + \lambda_{2m}\xi_m & + & \varepsilon_2 \\ X_3 & = & \lambda_{31}\xi_1 + \lambda_{32}\xi_2 + \cdots + \lambda_{3m}\xi_m & + & \varepsilon_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ X_p & = & \lambda_{p1}\xi_1 + \lambda_{p2}\xi_2 + \cdots + \lambda_{pm}\xi_m & + & \varepsilon_p \end{array}\right\} \text{[Eq 1]}$$

The Factor Analysis Model (FAM) specifies that each of p manifest variables $X_1, X_2, ..., X_p$ is a linear combination of $m < p$ common factors $\xi_1, \xi_2, ..., \xi_m$ and a single unique factor ε. Under the standard assumptions of FAM, the common factors are all uncorrelated with each other and standardized. The unique factors are all uncorrelated with each other and with all common factors, but are only centered (mean 0), rather than completely standardized.

The manifest variables $X_1, X_2, ..., X_p$ are all measured or observed for each of n individuals. So there are n observations on each equation. Therefore, the data consist of pn values $\{X_{1k}, X_{2k}, ..., X_{pk}, k = 1, ..., n\}$. For the n individuals, the FAM can be written:

$$\left\{\begin{array}{cccccc} X_{1k} & = & \lambda_{11}\xi_{1k} + \lambda_{12}\xi_{2k} + \cdots + \lambda_{1m}\xi_{mk} & + & \varepsilon_{1k} & ,k = 1,...,n \\ X_{2k} & = & \lambda_{21}\xi_{1k} + \lambda_{22}\xi_{2k} + \cdots + \lambda_{2m}\xi_{mk} & + & \varepsilon_{2k} & ,k = 1,...,n \\ X_{3k} & = & \lambda_{31}\xi_{1k} + \lambda_{32}\xi_{2k} + \cdots + \lambda_{3m}\xi_{mk} & + & \varepsilon_{3k} & ,k = 1,...,n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ X_{pk} & = & \lambda_{p1}\xi_{1k} + \lambda_{p2}\xi_{2k} + \cdots + \lambda_{pm}\xi_{mk} & + & \varepsilon_{pk} & ,k = 1,...,n \end{array}\right\} \text{[Eq 2]}$$

In Principal Factor Analysis (SAS option METHOD=PRINCIPAL) with uniqueness suppressed (SAS option PRIORS=ONE), the FAM becomes

$$\left\{\begin{array}{cccc} X_{1k} & = & \lambda_{11}\xi_{1k} + \lambda_{12}\xi_{2k} + \cdots + \lambda_{1m}\xi_{mk} & ,k = 1,...,n \\ X_{2k} & = & \lambda_{21}\xi_{1k} + \lambda_{22}\xi_{2k} + \cdots + \lambda_{2m}\xi_{mk} & ,k = 1,...,n \\ X_{3k} & = & \lambda_{31}\xi_{1k} + \lambda_{32}\xi_{2k} + \cdots + \lambda_{3m}\xi_{mk} & ,k = 1,...,n \\ \vdots & \vdots & \vdots & \vdots \\ X_{pk} & = & \lambda_{p1}\xi_{1k} + \lambda_{p2}\xi_{2k} + \cdots + \lambda_{pm}\xi_{mk} & ,k = 1,...,n \end{array}\right\} \text{[Eq 3]}$$

The factor scores for individual k are $(\xi_{1k}, \xi_{2k}, \cdots, \xi_{mk}), k = 1, ..., n$. These are estimated as linear combinations of the manifest variables:

$$\left.\begin{cases} \xi_{1k} & = & s_{11}X_{1k} + s_{12}X_{2k} + \cdots + s_{1p}X_{pk} & ,k=1,\ldots,n \\ \xi_{2k} & = & s_{21}X_{1k} + s_{22}X_{2k} + \cdots + s_{2p}X_{pk} & ,k=1,\ldots,n \\ \xi_{3k} & = & s_{31}X_{1k} + s_{32}X_{2k} + \cdots + s_{3p}X_{pk} & ,k=1,\ldots,n \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{mk} & = & s_{m1}X_{1k} + s_{m2}X_{2k} + \cdots + s_{mp}X_{pk} & ,k=1,\ldots,n \end{cases}\right\} \text{[Eq 4]}$$

The standardized factor scoring coefficients $\{s_{ij}\}$ provide the weights to apply to the (standardized) manifest variables in order to calculate factor scores for the individual. The standardized scoring coefficients are the same for every individual but different for each factor and manifest variable. Thus, there are *mp* scoring coefficients to be estimated.

In Principal Factoring, the same method of factor extraction is used as in Principal Component Analysis. That is, each extracted factor maximizes the variance of the factor over all possible linear combinations, subject to being orthogonal to previously extracted factors. Conceptually, you can think of the computational aspect of Principal Factoring as being a PCA. However, at the end of that process, all of the PCs would be standardized and some (*p* - *m*) PCs would be dropped. So the factors are the same as corresponding PCs, except that the factors are standardized. That is, $\xi_i^{FA} = \dfrac{\xi_i^{PCA} - E(\xi_i^{PCA})}{\sigma(\xi_i^{PCA})} = \dfrac{\xi_i^{PCA} - 0}{\sqrt{\lambda_i}} = \dfrac{\xi_i^{PCA}}{\sqrt{\lambda_i}}$ where λ_i is the i^{th} eigenvalue.

So in Eq 4 (which actually displays ξ_{ik}^{FA}), the factor scores of individual *k* on factor *i* are $\xi_{ik} = \xi_{ik}^{FA} = \dfrac{\xi_{ik}^{PCA}}{\sqrt{\lambda_i}}$:

$$\left.\begin{cases} \xi_{1k}^{FA} & = & \dfrac{\xi_{1k}^{PCA}}{\sqrt{\lambda_1}} = & s_{11}X_{1k} + s_{12}X_{2k} + \cdots + s_{1p}X_{pk} & ,k=1,\ldots,n \\ \xi_{2k}^{FA} & = & \dfrac{\xi_{2k}^{PCA}}{\sqrt{\lambda_2}} = & s_{21}X_{1k} + s_{22}X_{2k} + \cdots + s_{2p}X_{pk} & ,k=1,\ldots,n \\ \xi_{3k}^{FA} & = & \dfrac{\xi_{3k}^{PCA}}{\sqrt{\lambda_3}} = & s_{31}X_{1k} + s_{32}X_{2k} + \cdots + s_{3p}X_{pk} & ,k=1,\ldots,n \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \xi_{mk}^{FA} & = & \dfrac{\xi_{mk}^{PCA}}{\sqrt{\lambda_m}} = & s_{m1}X_{1k} + s_{m2}X_{2k} + \cdots + s_{mp}X_{pk} & ,k=1,\ldots,n \end{cases}\right\} \text{[Eq 5]}$$

After multiplying by $\sqrt{\lambda_i}$ in Eq 5, we get

$$\left.\begin{cases} \xi_{1k}^{PCA} & = & s_{11}\sqrt{\lambda_1}X_{1k} + s_{12}\sqrt{\lambda_1}X_{2k} + \cdots + s_{1p}\sqrt{\lambda_1}X_{pk} & ,k=1,\ldots,n \\ \xi_{2k}^{PCA} & = & s_{21}\sqrt{\lambda_2}X_{1k} + s_{22}\sqrt{\lambda_2}X_{2k} + \cdots + s_{2p}\sqrt{\lambda_2}X_{pk} & ,k=1,\ldots,n \\ \xi_{3k}^{PCA} & = & s_{31}\sqrt{\lambda_3}X_{1k} + s_{32}\sqrt{\lambda_3}X_{2k} + \cdots + s_{3p}\sqrt{\lambda_3}X_{pk} & ,k=1,\ldots,n \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{mk}^{PCA} & = & s_{m1}\sqrt{\lambda_m}X_{1k} + s_{m2}\sqrt{\lambda_m}X_{2k} + \cdots + s_{mp}\sqrt{\lambda_m}X_{pk} & ,k=1,\ldots,n \end{cases}\right\} \text{[Eq 6]}$$

These linear equations in Eq 6 fit every individual perfectly, for all *k* – with zero residuals. Therefore, they are the multiple regressions for the principal components as functions of the *X*'s. But we also know that the principal component model in Eq 7 below fits the principal components perfectly as well.

$$\left.\begin{cases} \xi_1^{PCA} & = & \beta_{11}X_1 + \beta_{12}X_2 + \beta_{13}X_3 + \cdots + \beta_{1p}X_p \\ \xi_2^{PCA} & = & \beta_{21}X_1 + \beta_{22}X_2 + \beta_{23}X_3 + \cdots + \beta_{2p}X_p \\ \xi_3^{PCA} & = & \beta_{31}X_1 + \beta_{32}X_2 + \beta_{33}X_3 + \cdots + \beta_{3p}X_p \\ \vdots & \vdots & \vdots \\ \xi_p^{PCA} & = & \beta_{p1}X_1 + \beta_{p2}X_2 + \beta_{p3}X_3 + \cdots + \beta_{pp}X_p \end{cases}\right\} \text{[Eq 7]}$$

Therefore, the multiple regression coefficients in Eq 6 and Eq 7 must be the same: $s_{ij}\sqrt{\lambda_i} = \beta_{ij}$. Thus, the ratio of PCA coefficients to FA coefficients is $\frac{\beta_{ij}}{s_{ij}} = \sqrt{\lambda_i}$, which is the same value for each manifest variable in a given principal component ξ_i^{PCA} or factor ξ_i^{FA}. Thus, the ratio of principal component *i* score for individual *k* to factor *i* score for individual *k* is $\frac{\xi_{ik}^{PCA}}{\xi_{ik}^{FA}} = \frac{\beta_{i1}X_{1k} + \beta_{i2}X_{2k} + \cdots + \beta_{ip}X_{pk}}{s_{i1}X_{1k} + s_{i2}X_{2k} + \cdots + s_{ip}X_{pk}} = \frac{\sqrt{\lambda_i}s_{i1}X_{1k} + \sqrt{\lambda_i}s_{i2}X_{2k} + \cdots + \sqrt{\lambda_i}s_{ip}X_{pk}}{s_{i1}X_{1k} + s_{i2}X_{2k} + \cdots + s_{ip}X_{pk}} = \sqrt{\lambda_i}$.

Moreover, the form of the PCA coefficients is known (see PCA notes): $\beta_{ij} = \frac{corr(\xi_i^{PCA}, X_j)}{\sqrt{\lambda_i}}$.

Thus, $s_{ij}\sqrt{\lambda_i} = \beta_{ij} = \frac{corr(\xi_i^{PCA}, X_j)}{\sqrt{\lambda_i}}$. So $s_{ij} = \frac{corr(\xi_i^{PCA}, X_j)}{\lambda_i}$. But correlation is not affected by linear transformation of variables, so $corr(\xi_i^{PCA}, X_j) = corr(\sqrt{\lambda_i}\xi_i^{FA}, X_j) = corr(\xi_i^{FA}, X_j) = \lambda_{ij}$ in Eq 3. Thus $s_{ij} = \frac{\lambda_{ij}}{\lambda_i}$ is the form of the standardized factor scoring coefficients. This is a very simple form. It requires knowing only the eigenvalues (λ_i's) of the correlation matrix of the manifest variables and the estimated factor coefficients (λ_{ij}'s) from least squares estimation of these coefficients in Step 1 of "*Estimation of factor loadings and scores.pdf*". This special case (principal factoring and no uniqueness) is worth study because it does make clear that the estimation of factor scores is feasible, in spite of appearing to be a hopeless project of conjuring up values magically out of nothing. For the general case of factor score estimation, see "*Estimation of factor loadings and scores.pdf*".