Factor Models: A Review

James J. Heckman
The University of Chicago

Econ 350, Winter 2023

Factor Models: A Review of General Models

$$E(\theta) = 0; \qquad E(\varepsilon_i) = 0; \;\; i = 1, \dots, 5$$

$$\theta \perp\!\!\!\perp (\varepsilon_i, \dots, \varepsilon_5)$$

$$R_1 = \alpha_1 \theta + \varepsilon_1, \quad R_2 = \alpha_2 \theta + \varepsilon_2, \quad R_3 = \alpha_3 \theta + \varepsilon_3,$$
$$R_4 = \alpha_4 \theta + \varepsilon_4, \quad R_5 = \alpha_5 \theta + \varepsilon_5, \quad \varepsilon_i \perp\!\!\!\perp \varepsilon_j, \; i \neq j$$

$$Cov(R_1, R_2) = \alpha_1 \alpha_2 \sigma_\theta^2$$
$$Cov(R_1, R_3) = \alpha_1 \alpha_3 \sigma_\theta^2$$
$$Cov(R_2, R_3) = \alpha_2 \alpha_3 \sigma_\theta^2$$

- Normalize $\alpha_1 = 1$

$$\frac{Cov(R_2, R_3)}{Cov(R_1, R_2)} = \alpha_3$$

THE UNIVERSITY OF CHICAGO

- $\therefore$ We know σ_θ^2 from $Cov\,(R_1, R_2)$.
- From $Cov\,(R_1, R_3)$ we know

$$\alpha_3, \alpha_4, \alpha_5.$$

- Can get the variances of the ε_i from variances of the R_i

$$Var(R_i) = \alpha_i^2 \sigma_\theta^2 + \sigma_{\varepsilon_i}^2.$$

- If $T = 2$, all we can identify is $\alpha_1 \alpha_2 \sigma_\theta^2$.
- If $\alpha_1 = 1$, $\sigma_\theta^2 = 1$, we identify α_2.
- Otherwise model is fundamentally underidentified.

2 Factors: (Some Examples)

$$\theta_1 \perp\!\!\!\perp \theta_2$$

$$\varepsilon_i \perp\!\!\!\perp \varepsilon_j \ \ \forall i \neq j$$

$$\begin{aligned}
R_1 &= \alpha_{11}\theta_1 + (0)\theta_2 + \varepsilon_1 \\
R_2 &= \alpha_{21}\theta_1 + (0)\theta_2 + \varepsilon_2 \\
R_3 &= \alpha_{31}\theta_1 + \alpha_{32}\theta_2 + \varepsilon_3 \\
R_4 &= \alpha_{41}\theta_1 + \alpha_{42}\theta_2 + \varepsilon_4 \\
R_5 &= \alpha_{51}\theta_1 + \alpha_{52}\theta_2 + \varepsilon_5
\end{aligned}$$

Let $\alpha_{11} = 1$, $\alpha_{32} = 1$. (Set scale)

$$Cov(R_1, R_2) = \alpha_{21}\sigma^2_{\theta_1}$$
$$Cov(R_1, R_3) = \alpha_{31}\sigma^2_{\theta_1}$$
$$Cov(R_2, R_3) = \alpha_{21}\alpha_{31}\sigma^2_{\theta_1}$$

- Form ratio of $\dfrac{Cov(R_2, R_3)}{Cov(R_1, R_2)} = \alpha_{31}, \quad \therefore$ we identify $\alpha_{31}, \alpha_{21}, \sigma^2_{\theta_1}$, as before.

$$Cov(R_1, R_4) = \alpha_{41}\sigma^2_{\theta_1}, \qquad \therefore \text{ since we know } \sigma^2_{\theta_1} \therefore \text{ we get } \alpha_{41}.$$
$$\vdots$$
$$Cov(R_1, R_k) = \alpha_{k1}\sigma^2_{\theta_1}$$

- $\therefore$ we identify α_{k1} for all k and $\sigma^2_{\theta_1}$.

$$
\begin{aligned}
Cov\left(R_3, R_4\right) - \alpha_{31}\alpha_{41}\sigma^2_{\theta_1} &= \alpha_{42}\sigma^2_{\theta_2} \\
Cov\left(R_3, R_5\right) - \alpha_{31}\alpha_{51}\sigma^2_{\theta_1} &= \alpha_{52}\sigma^2_{\theta_2} \\
Cov\left(R_4, R_5\right) - \alpha_{41}\alpha_{51}\sigma^2_{\theta_1} &= \alpha_{52}\alpha_{42}\sigma^2_{\theta_2},
\end{aligned}
$$

- By same logic,

$$
\frac{Cov\left(R_4, R_5\right) - \alpha_{41}\alpha_{51}\sigma^2_{\theta_1}}{Cov\left(R_3, R_4\right) - \alpha_{31}\alpha_{41}\sigma^2_{\theta_1}} = \alpha_{52}
$$

- $\therefore$ get $\sigma^2_{\theta_2}$ of "2" loadings.

- If we have dedicated measurements on each factor do not need a normalization on the factors of R.
- Dedicated measurements set the scales and make factor models interpretable:

$$M_1 = \theta_1 + \varepsilon_{1M}$$
$$M_2 = \theta_2 + \varepsilon_{2M}$$

$$\text{Cov}\left(R_1, M\right) = \alpha_{11}\sigma^2_{\theta_1}$$
$$\text{Cov}\left(R_2, M\right) = \alpha_{21}\sigma^2_{\theta_1}$$
$$\text{Cov}\left(R_3, M\right) = \alpha_{31}\sigma^2_{\theta_1}$$

$$\text{Cov}\left(R_1, R_2\right) = \alpha_{11}\alpha_{12}\sigma^2_{\theta_1},$$
$$\text{Cov}\left(R_1, R_3\right) = \alpha_{11}\alpha_{13}\sigma^2_{\theta_1}, \qquad \text{so we can identify } \alpha_{12}\sigma^2_{\theta_1}$$

- $\therefore$ We can get $\alpha_{12}, \sigma^2_{\theta_1}$ and the other parameters.

General Case

$$\underset{T\times 1}{R} = \underset{T\times 1}{M} + \underset{T\times K}{\Lambda}\underset{K\times 1}{\theta} + \underset{T\times 1}{\varepsilon}$$

- θ are factors, ε uniquenesses, $\theta \perp\!\!\!\perp \varepsilon$

$$E(\varepsilon) = 0$$

$$Var(\varepsilon\varepsilon') = D = \begin{pmatrix} \sigma^2_{\varepsilon_1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\varepsilon_2} & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma^2_{\varepsilon_T} \end{pmatrix}$$

$$E(\theta) = 0$$

$$Var(R) = \Lambda\Sigma_\theta\Lambda' + D \qquad \Sigma_\theta = E(\theta\theta')$$

- The only source of information on Λ and Σ_θ is from the covariances.
- (Each variance is "contaminated" by a uniqueness.)
- Associated with each variance of R_i is a $\sigma^2_{\varepsilon_i}$.
- Each uniqueness variance contributes one new parameter.
- How many unique covariance terms do we have?
- $\dfrac{T(T-1)}{2}$.

- We have T uniquenesses; TK elements of Λ.
- $\frac{K(K-1)}{2}$ elements of Σ_θ.
- $\frac{K(K-1)}{2} + TK$ parameters (Σ_θ, Λ).
- Need this many covariances to identify model "*Ledermann Bound*":

$$\frac{T(T-1)}{2} \geq TK + \frac{K(K-1)}{2}$$

Lack of Identification Up to Rotation

- Observe that if we multiply Λ by an orthogonal matrix C, $(CC' = I)$, we obtain

$$Var\left(R\right) = \Lambda C \left[C'\Sigma_{\theta} C\right] C'\Lambda' + D$$

- C is a "rotation."
- Cannot separate ΛC from Λ.
- Model not identified against orthogonal transformations in the general case.

Some common assumptions:

1. $\theta_i \perp\!\!\!\perp \theta_j, \forall\ i \neq j$

$$\Sigma_\theta = \begin{pmatrix} \sigma^2_{\theta_1} & 0 & \cdots & 0 \\ 0 & \sigma^2_{\theta_2} & 0 & \vdots \\ \vdots & 0 & \ddots & \vdots \\ 0 & \cdots & 0 & \sigma^2_{\theta_K} \end{pmatrix}$$

joined with

ii

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \alpha_{21} & 0 & 0 & 0 & \cdots & 0 \\ \alpha_{31} & 1 & 0 & 0 & \cdots & 0 \\ \alpha_{41} & \alpha_{42} & 0 & 0 & \cdots & 0 \\ \alpha_{51} & \alpha_{52} & 1 & 0 & \cdots & 0 \\ \alpha_{61} & \alpha_{62} & \alpha_{63} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & 1 & & \vdots \end{pmatrix}$$

- We know that we can identify of the Λ, Σ_θ parameters.

$$\underbrace{\frac{K(K-1)}{2} + TK}_{\#\text{ of free parameters}} \leq \underbrace{\frac{T(T-1)}{2}}_{\text{data}}$$

"Ledermann Bound"

- Can get more information by looking at higher order moments.
- (See, e.g., Bonhomme and Robin, 2009.)

- Normalize: $\alpha_{I^*} = 1$, $\alpha_1 = 1 \quad \therefore \sigma_\theta^2 \quad \therefore \alpha_1$.
- Can make alternative normalizations.

Recovering the Distributions Nonparametrically

Theorem 1

Suppose that we have two random variables T_1 and T_2 that satisfy:

$$T_1 = \theta + v_1$$
$$T_2 = \theta + v_2$$

with θ, v_1, v_2 mutually statistically independent, $E(\theta) < \infty$, $E(v_1) = E(v_2) = 0$, that the conditions for Fubini's theorem are satisfied for each random variable, and the random variables possess nonvanishing (a.e.) characteristic functions, then the densities $f(\theta), f(v_1)$, and $f(v_2)$ are identified.

Proof.

See Kotlarski (1967). □

- Suppose

$$
\begin{aligned}
I &= \mu_I(X,Z) + \alpha_I\theta + \varepsilon_I \\
Y_0 &= \mu_0(X) + \alpha_0\theta + \varepsilon_0 \\
Y_1 &= \mu_1(X) + \alpha_1\theta + \varepsilon_1 \\
M &= \mu_M(X) + \theta + \varepsilon_M.
\end{aligned}
$$

- System can be rewritten as

$$
\begin{aligned}
\frac{I - \mu_I(X,Z)}{\alpha_I} &= \theta + \frac{\varepsilon_I}{\alpha_I} \\
\frac{Y_0 - \mu_0(X)}{\alpha_0} &= \theta + \frac{\varepsilon_0}{\alpha_0} \\
\frac{Y_1 - \mu_1(X)}{\alpha_1} &= \theta + \frac{\varepsilon_1}{\alpha_1} \\
M - \mu_M(X) &= \theta + \varepsilon_M
\end{aligned}
$$

- Applying Kotlarski's theorem, identify the densities of $\theta, \frac{\varepsilon_I}{\alpha_I}, \frac{\varepsilon_0}{\alpha_0}, \frac{\varepsilon_1}{\alpha_1}, \varepsilon_M$.
- We know α_I, α_0 and α_1.
- Can identify the densities of $\theta, \varepsilon_I, \varepsilon_0, \varepsilon_1, \varepsilon_M$.
- Recover the joint distribution of (Y_1, Y_0).

$$F(Y_1, Y_0 \mid X) = \int F(Y_1, Y_0 \mid \theta, X)\, dF(\theta).$$

- $F(\theta)$ is known.

$$F(Y_1, Y_0 \mid \theta, X) = F(Y_1 \mid \theta, X)\, F(Y_0 \mid \theta, X).$$

- $F(Y_1 \mid \theta, X)$ and $F(Y_0 \mid \theta, X)$ identified

$$F(Y_1 \mid \theta, X, S = 1) = F(Y_1 \mid \theta, X)$$
$$F(Y_0 \mid \theta, X, S = 0) = F(Y_0 \mid \theta, X).$$

- Can identify the number of factors generating dependence among the Y_1, Y_0, C, S and M.