Solutions to Problem Set 7: Factor Modeling

Jura Ivanković

MIT Financial Mathematics course website: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/`

Problem sets: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/assignments/`

Problem set 7: `https://ocw.mit.edu/courses/mathematics/18-s096-topics-in-mathematics-with-applications-in-finance-fall-2013/assignments/MIT18_S096F13_pset7.pdf`

1. (a)

$$|\Sigma - \lambda I| = \left(\sigma_1^2 - \lambda\right)\left(\sigma_2^2 - \lambda\right) - \rho^2\sigma_1^2\sigma_2^2 = \lambda^2 - \left(\sigma_1^2 + \sigma_2^2\right)\lambda + \left(1 - \rho^2\right)\sigma_1^2\sigma_2^2$$

$$\begin{aligned}
\therefore \lambda_1, \lambda_2 &= \frac{\sigma_1^2 + \sigma_2^2 \pm \sqrt{\left(\sigma_1^2 + \sigma_2^2\right)^2 - 4\left(1 - \rho^2\right)\sigma_1^2\sigma_2^2}}{2} \\
&= \frac{\sigma_1^2 + \sigma_2^2 \pm \sqrt{\sigma_1^4 + 2\sigma_1^2\sigma_2^2 + \sigma_2^4 - 4\sigma_1^2\sigma_2^2 + 4\rho^2\sigma_1^2\sigma_2^2}}{2} \\
&= \frac{\sigma_1^2 + \sigma_2^2 \pm \sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}}{2}
\end{aligned}$$

(b) Case $\rho = 0$:

$$\Sigma\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\sigma_1^2\\0\end{pmatrix}, \quad \Sigma\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\\sigma_2^2\end{pmatrix}$$

Case $\rho \neq 0$: let $\gamma_i = \begin{pmatrix}\gamma_{i,1} & \gamma_{i,2}\end{pmatrix}'$ for $1 \leq i \leq 2$.

$$\begin{aligned}
\sigma_1^2\gamma_{i,1} + \rho\sigma_1\sigma_2\gamma_{i,2} &= \lambda_i\gamma_{i,1} \\
\gamma_{i,2} &= \frac{\left(\lambda_i - \sigma_1^2\right)\gamma_{i,1}}{\rho\sigma_1\sigma_2} \\
\gamma_{i,1}^2 + \frac{\left(\lambda_i - \sigma_1^2\right)^2\gamma_{i,1}^2}{\rho^2\sigma_1^2\sigma_2^2} &= 1
\end{aligned}$$

$$\begin{aligned}
\gamma_{i,1}^2 &= \frac{1}{1 + \frac{\left(\lambda_i - \sigma_1^2\right)^2}{\rho^2\sigma_1^2\sigma_2^2}} \\
&= \frac{\rho^2\sigma_1^2\sigma_2^2}{\rho^2\sigma_1^2\sigma_2^2 + \left(\lambda_i - \sigma_1^2\right)^2} \\
&= \frac{\rho^2\sigma_1^2\sigma_2^2}{\rho^2\sigma_1^2\sigma_2^2 + \left(\frac{\sigma_2^2 - \sigma_1^2 \pm \sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}}{2}\right)^2}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{4\rho^2\sigma_1^2\sigma_2^2}{4\rho^2\sigma_1^2\sigma_2^2 + \left(\sigma_2^2 - \sigma_1^2 \pm \sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}\right)^2} \\
&= \frac{4\rho^2\sigma_1^2\sigma_2^2}{4\rho^2\sigma_1^2\sigma_2^2 + \left(\sigma_2^2 - \sigma_1^2\right)^2 \pm 2\left(\sigma_2^2 - \sigma_1^2\right)\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} + \left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \\
&= \frac{4\rho^2\sigma_1^2\sigma_2^2}{2\left(\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2 \pm \left(\sigma_2^2 - \sigma_1^2\right)\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}\right)} \\
&= \frac{2\rho^2\sigma_1^2\sigma_2^2}{\left(\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \pm \left(\sigma_2^2 - \sigma_1^2\right)\right)\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}}
\end{aligned}
$$

Let

$$
\gamma_{i,1} = \frac{\sqrt{2}\rho\sigma_1\sigma_2}{\sqrt{\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \pm \left(\sigma_2^2 - \sigma_1^2\right)}\left(\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}}
$$

Then,

$$
\begin{aligned}
\gamma_{i,2} &= \frac{\left(\lambda_i - \sigma_1^2\right)\sqrt{2}\rho\sigma_1\sigma_2}{\rho\sigma_1\sigma_2\sqrt{\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \pm \left(\sigma_2^2 - \sigma_1^2\right)}\left(\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \\
&= \frac{\left(\sigma_2^2 - \sigma_1^2 \pm \sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}\right)\sqrt{2}}{2\sqrt{\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \pm \left(\sigma_2^2 - \sigma_1^2\right)}\left(\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \\
&= \frac{\sigma_2^2 - \sigma_1^2 \pm \sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2}}{\sqrt{2}\sqrt{\sqrt{\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2} \pm \left(\sigma_2^2 - \sigma_1^2\right)}\left(\left(\sigma_2^2 - \sigma_1^2\right)^2 + 4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}}
\end{aligned}
$$

$$
\therefore \gamma_1 = \begin{cases}
\begin{pmatrix} 1 \\ 0 \end{pmatrix} & \rho = 0, \sigma_1^2 \geq \sigma_2^2 \\
\begin{pmatrix} 0 \\ 1 \end{pmatrix} & \rho = 0, \sigma_1^2 < \sigma_2^2 \\
\begin{pmatrix} \frac{2\rho\sigma_1\sigma_2}{\sqrt{2}\sqrt{\sqrt{\left(\sigma_2^2-\sigma_1^2\right)^2+4\rho^2\sigma_1^2\sigma_2^2}+\left(\sigma_2^2-\sigma_1^2\right)}\left(\left(\sigma_2^2-\sigma_1^2\right)^2+4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \\ \frac{\sigma_2^2-\sigma_1^2+\sqrt{\left(\sigma_2^2-\sigma_1^2\right)^2+4\rho^2\sigma_1^2\sigma_2^2}}{\sqrt{2}\sqrt{\sqrt{\left(\sigma_2^2-\sigma_1^2\right)^2+4\rho^2\sigma_1^2\sigma_2^2}+\left(\sigma_2^2-\sigma_1^2\right)}\left(\left(\sigma_2^2-\sigma_1^2\right)^2+4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \end{pmatrix} & \rho \neq 0
\end{cases}
$$

$$\gamma_2 = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \rho = 0, \sigma_1^2 \geq \sigma_2^2 \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \rho = 0, \sigma_1^2 < \sigma_2^2 \\ \begin{pmatrix} \frac{2\rho\sigma_1\sigma_2}{\sqrt{2}\sqrt{\sqrt{(\sigma_2^2-\sigma_1^2)^2+4\rho^2\sigma_1^2\sigma_2^2}-(\sigma_2^2-\sigma_1^2)}\left((\sigma_2^2-\sigma_1^2)^2+4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \\ \frac{\sigma_2^2-\sigma_1^2-\sqrt{(\sigma_2^2-\sigma_1^2)^2+4\rho^2\sigma_1^2\sigma_2^2}}{\sqrt{2}\sqrt{\sqrt{(\sigma_2^2-\sigma_1^2)^2+4\rho^2\sigma_1^2\sigma_2^2}-(\sigma_2^2-\sigma_1^2)}\left((\sigma_2^2-\sigma_1^2)^2+4\rho^2\sigma_1^2\sigma_2^2\right)^{\frac{1}{4}}} \end{pmatrix} & \rho \neq 0 \end{cases}$$

(c) Since Σ is a real symmetric matrix, Σ is orthonormally diagonalizable.

$$\therefore \Sigma = \begin{pmatrix} \gamma_1 & \gamma_2 & \cdots & \gamma_m \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \end{pmatrix} \begin{pmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_m' \end{pmatrix} = \begin{pmatrix} \lambda_1\gamma_1 & \lambda_2\gamma_2 & \cdots & \lambda_m\gamma_m \end{pmatrix} \begin{pmatrix} \gamma_1' \\ \gamma_2' \\ \vdots \\ \gamma_m' \end{pmatrix} = \sum_{i=1}^{m} \lambda_i\gamma_i\gamma_i'$$

(d)

$$E[p_i] = E[\gamma_i'(x-\alpha)] = \gamma_i' E[x-\alpha] = \gamma_i'\mathbf{0} = 0$$
$$Var(p_i) = E\left[p_i^2\right] = E[\gamma_i'(x-\alpha)(x-\alpha)'\gamma_i] = \gamma_i'\Sigma\gamma_i = \lambda_1\gamma_i'\gamma_1\gamma_1'\gamma_i + \lambda_2\gamma_i'\gamma_2\gamma_2'\gamma_i = \lambda_i$$
$$Cov(p_1,p_2) = E[p_1p_2] = E[\gamma_1'(x-\alpha)(x-\alpha)'\gamma_2] = \gamma_1'\Sigma\gamma_2 = \lambda_1\gamma_1'\gamma_1\gamma_1'\gamma_2 + \lambda_2\gamma_1'\gamma_2\gamma_2'\gamma_2 = 0$$

(e) For the given solution γ_1, γ_2, the following are also solutions: $\gamma_1, -\gamma_2$; $-\gamma_1, \gamma_2$; $-\gamma_1, -\gamma_2$. Additionally, for case $\rho = 0, \sigma_1^2 = \sigma_2^2$, γ_1 and γ_2 are any orthogonal unit vectors.

2. (a) Let θ be the angle between ϕ_1 and γ_1 and ψ the angle between ϕ_1 and γ_2. Since γ_1 and γ_2 are orthogonal, $\cos^2\psi = \sin^2\theta$. Then,

$$\begin{aligned} Var\left(\phi_1^T X\right) &= E\left[\left(\phi_1^T(X-E[X])\right)^2\right] \\ &= E\left[\phi_1^T(X-E[X])(X-E[X])^T\phi_1\right] \\ &= \phi_1^T\Sigma\phi_1 \\ &= \lambda_1\left(\phi_1^T\gamma_1\right)^2 + \lambda_2\left(\phi_1^T\gamma_2\right)^2 \\ &= \lambda_1\cos^2\theta + \lambda_2\sin^2\theta \\ &= (\lambda_1-\lambda_2)\cos^2\theta + \lambda_2 \end{aligned}$$

$$\underset{\theta\in\mathbb{R}}{\arg\max}\, Var\left(\phi_1^T X\right) = \{k\pi \mid k \in \mathbb{Z}\}$$

$$\therefore \quad \phi_1 = \pm\gamma_1, \quad Var\left(\phi_1^T X\right) = \lambda_1$$

(b) Let θ be the angle between ϕ_2 and γ_1. Then, $Var\left(\phi_2^T X\right) = (\lambda_1-\lambda_2)\cos^2\theta + \lambda_2$. Then,

$$\underset{\theta\in\mathbb{R}}{\arg\min}\, Var\left(\phi_2^T X\right) = \left\{\frac{(2k+1)\pi}{2} \,\middle|\, k \in \mathbb{Z}\right\}$$

$$\therefore \quad \phi_2 = \pm\gamma_2, \quad Var\left(\phi_2^T X\right) = \lambda_2$$

(c) Principal components analysis T of a random m-vector X is a translation to $X - E[X]$ followed by a rotation such that for any $1 \leq i \leq m$, the variance of the ith entry of $T(X)$ is maximised subject to the value of the jth entry of $T(X)$ for all $1 \leq j < i$.

3. (a) The difference in sign is not meaningful, as an eigendecomposition of a matrix does not change if any eigenvector is multiplied by -1.

   (b) Let $X = \begin{pmatrix} X_1 & X_2 & \cdots & X_m \end{pmatrix}^T$ be the vector of yield changes, $\phi_1 = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,m} \end{pmatrix}^T$ a unit vector, $\sigma_i = \sigma(X_i)$ and $\rho_{i,j} = Corr(X_i, X_j)$ for $1 \leq i, j \leq m$.

$$\phi_1^T Cov(X)\phi_1 = \sum_{i=1}^{m} \phi_{1,i}\sigma_i \sum_{j=1}^{m} \phi_{1,j}\sigma_j\rho_{i,j}, \quad \phi_1^T Corr(X)\phi_1 = \sum_{i=1}^{m} \phi_{1,i} \sum_{j=1}^{m} \phi_{1,j}\rho_{i,j}$$

   Clearly, the ϕ_1 that maximises $\phi_1^T Cov(X)\phi_1$ will give lesser weight to less variable yield changes and greater weight to more variable yield changes than the ϕ_1 that maximises $\phi_1^T Corr(X)\phi_1$. This explains why the first principal component of the correlation matrix has a smaller range of magnitudes of loadings than that of the covariance matrix.

   (c) In the correlation matrix case, the first principal component is the negative of the average yield change across tenors weighted by correlation, the second principal component is the spread between short-term and long-term tenors with a negative score indicating that long-term tenors have greater yield changes than short-term tenors and the third principal component is the curvature of the yield curve, with a low score indicating high curvature.

   In the covariance matrix case, the first principal component is the positive average yield change weighted by covariance, the second principal component is spread with a positive score indicating that long-term tenors have larger yield changes than short-term tenors and the third principal component is similar to the correlation matrix case.

   (d) Principal components analysis of the correlation matrix might be preferred to that of the covariance matrix because in the latter case the first principal component will give greater weight to tenors with more highly variable yield changes and insufficient weight to tenors with less variable yield changes.

4. (a) The p-value of the F-statistic is greater than 0.05, so the estimation results are inconclusive.

   (b) The F-statistic is significant at the 0.01% level. The only variable that is significant at the 5% level is Comp.2.l1, which is significant at the 0.0001% level. The estimate of Comp.2.l1 is 0.138152 with standard error 0.028048. Therefore, Comp.2 can be modelled as an $AR(1)$ model with mean 0 and coefficient 0.138152.

   (c) The p-value of the F-statistic is greater than 0.05, so the estimation results are inconclusive.

   (d) Since Comp.2 can be modelled as an $AR(1)$ model with coefficient of magnitude less than 1, there is evidence of mean-reversion for Comp.2.

   There is no evidence for either mean-reversion or momentum for Comp.1 or Comp.3.