Factor Models

Feng Li
feng.li@cufe.edu.cn

School of Statistics and Mathematics
Central University of Finance and Economics

Revised on May 28, 2015

Today we are going to learn...

1. Introduction to Factor Models
2. Macroeconometric factor models
3. Fundamental factor models
4. Statistical factor analysis

Background I

- Most financial portfolios consist of multiple assets, and their returns depend concurrently and dynamically on many economic and financial variables. Therefore, it is important to use proper multivariate statistical analyses to study the behavior and properties of portfolio returns.
- Analysis of multiple asset returns often requires high-dimensional statistical models that are complicated and hard to apply.
- To simplify the task of modeling multiple returns, one can use dimension reduction methods to search for the underlying structure of the assets.
- In practice, observed return series often exhibit similar characteristics leading to the belief that they might be driven by some common sources, often referred to as **common factors**.

Background II

- Three types of factor models are available for studying asset returns.
  1. **macroeconomic factor models**. They use macroeconomic variables such as growth rate of GDP, interest rates, inflation rate, and unemployment rate to describe the common behavior of asset returns. Here the factors are observable and the model can be estimated via linear regression methods.
  2. **fundamental factor models**. They use firm or asset specific attributes such as firm size, book and market values, and industrial classification to construct common factors.
  3. **statistical factor models** They treat the common factors as unobservable or latent variables to be estimated from the returns series.

Factor Models I

- Suppose that there are k assets and T time periods. Let r it be the return of asset i in the time period t. A general form for the **factor model** is

$$r_{it} = \alpha_i + \beta_{i1} f_{1t} + \ldots + \beta_{im} f_{mt} + \epsilon_{it}$$

or in matrix form

$$\begin{aligned} \mathbf{r}_t &= \boldsymbol{\alpha} + \boldsymbol{\beta}_i \mathbf{f}_t + \boldsymbol{\epsilon}_t \\ &= \boldsymbol{\xi} \mathbf{g}_t + \boldsymbol{\epsilon}_t \end{aligned}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_k)$ is a constant representing the intercept vector, $\mathbf{f}_t = (f_{1t}, \ldots, f_{mt})$ are m common factors, $\boldsymbol{\beta} = \beta_{ij}$ is the **factor loading** matrix for asset i on the j th factor, and ϵ_{it} is the **specific factor** of asset i with $Cov(\boldsymbol{\epsilon}_t) = \mathbf{D} = diag(\sigma_1^2, \ldots, \sigma_k^2)$.

Factor Models II

- For asset returns, the factor $\mathbf{f}_t = (f_{1t}, ..., f_{mt})$ is assumed to be an m-dimensional stationary process such that

$$
\begin{aligned}
E(\mathbf{f}_t) &= \boldsymbol{\mu}_f, \\
Cov(\mathbf{f}_t) &= \boldsymbol{\Sigma}_f
\end{aligned}
$$

- and the asset specific factor ϵ_{it} is a white noise series and uncorrelated with the common factors f_{jt} and other specific factors.
- The common factors are uncorrelated with the specific factors, and the specific factors are uncorrelated among each other. The common factors, however, need not be uncorrelated with each other in some factor models.
- In some applications, the number of assets k may be larger than the number of time periods T.
- The model presented above is in a cross-sectional regression form if the factors f_{jt} are observed.
- It is also common to assume that the factors, hence $\mathbf{r}_t$, are serially uncorrelated in factor analysis.

Factor Models III

- The covariance matrix of the return $\mathbf{r}_t$ is then

$$\mathrm{Cov}(\mathbf{r}_t) = \beta\boldsymbol{\Sigma}_f\boldsymbol{\beta}' + \mathbf{D}$$

- We can treat the factor model as a time series, we have

$$\mathbf{R}_i = \boldsymbol{\alpha}_i\mathbf{1}_T + \mathbf{F}\boldsymbol{\beta}_i' + \mathbf{E}_i$$

for the ith asset, where $\mathbf{R}_i = (r_{i1}, ..., r_{iT})'$, $\mathbf{1}_t$ is a T-dimensional vector of ones, $\mathbf{F}$ is a $T \times m$ matrix whose tth row is f_t', and $E_i = (\epsilon_{i1}, ..., \epsilon_{iT})'$. The covariance matrix of E_i is $\mathrm{Cov}E_i = \sigma_i^2\mathbf{I}$, a $T \times T$ diagonal matrix.

- Taking the transpose of the matrix equation and stacking all data together, we have

$$\mathbf{R} = \mathbf{G}\boldsymbol{\xi}' + \mathbf{E}$$

where $\mathbf{R}$ is a $T \times k$ matrix of returns whose tth row is r_t', G is a $T \times (m+1)$ matrix whose tth row is g_t', and E is a $T \times k$ matrix of specific factors whose tth row is ϵ_t'

Factor Models IV

- If the common factors f_t are observed, then the above model is a special form of the **multivariate linear regression** (MLR) model.

Macroeconometric factor models I

- For macroeconomic factor models, the factors are observed and we can apply the least-squares method to the MLR model to perform estimation.
  1. The estimate is
  $$\hat{\boldsymbol{\xi}}' = (\mathbf{G}'\mathbf{G})^{-1}\mathbf{G}'\mathbf{R}$$
  2. The residuals are
  $$\hat{\mathbf{E}} = \mathbf{R} - \mathbf{G}\hat{\boldsymbol{\xi}}'$$
  3. Based on the model assumption, the covariance matrix of ϵ_t is estimated by
  $$\hat{\mathbf{D}} = diag\{\hat{\sigma}_1^2, ..., \hat{\sigma}_k^2\}$$
  where $\hat{\sigma}_i^2$ is the (i, i)-th element of $\hat{\mathbf{E}}'\hat{\mathbf{E}}/(T - m - 1)$

Macroeconometric factor models II

4. Furthermore, the R^2 of the ith asset is

$$\mathbf{R}_i^2 = 1 - \frac{[\hat{\mathbf{E}}'\hat{\mathbf{E}}]_{i,i}}{[\mathbf{R}'\mathbf{R}]_{i,i}}$$

- Note that the aforementioned least-squares estimation does not impose the constraint that the specific factors ϵ_{it} are uncorrelated with each other. Consequently, the estimates obtained are not efficient in general.
- The best known macroeconomic factor model in finance is the **market model**. This is a **single-factor model** and can be written as

$$r_{it} = \alpha_i + \beta_i r_{mt} + \epsilon_{it}$$

  - **R example** we consider monthly returns of 13 stocks and use the return of the S&P 500 index as the market return.

Macroeconometric factor models III

- **Multifactor model** for stock returns. The factors used consist of unexpected changes or surprises of macroeconomic variables. Here unexpected changes denote the residuals of the macroeconomic variables after removing their dynamic dependence. A simple way to obtain unexpected changes is to fit a VAR model of Chapter 8 to the macroeconomic variables
  - **R example** Consumer price index (CPI) for all urban consumers, Civilian employment numbers 16 years and over.

Fundamental factor models I

- **Fundamental factor models** use observable asset specific fundamentals such as industrial classification, market capitalization, book value, and style classification (growth or value) to construct common factors that explain the excess returns.
- There are two approaches to fundamental factor models available in the literature.
  - The first approach is proposed by Bar Rosenberg, founder of BARRA Inc., and is referred to as the **BARRA approach**. In contrast to the macroeconomic factor models, this approach treats the observed asset specific fundamentals as the factor betas, β_i , and estimates the factors f_t at each time index t via regression methods. The betas are time invariant, but the realizations f_t evolve over time.
    - Assume that the excess returns and, hence, the factor realizations are mean corrected. At each time index t, the factor model reduces to

$$\tilde{\mathbf{r}}_t = \boldsymbol{\beta} \mathbf{f}_t + \boldsymbol{\epsilon}_t$$

where $\tilde{\mathbf{r}}_t$ denotes the (sample) mean-corrected excess returns and, for simplicity in notation, we continue to use f_t as factor realizations.

Fundamental factor models II

- However, the regression is not homogeneous because the covariance matrix of ϵ_t is $D = diag\{\sigma_1^2, ..., \sigma_k^2\}$.
- Consequently, the factor realization at time index t can be estimated by the weighted least-squares (WLS) method using the standard errors of the specific factors as the weights.

$$\hat{f}_t = (\beta' D^{-1} \beta)^{-1} (\beta' D^{-1} \tilde{\mathbf{r}}_t)$$

- In practice, the covariance matrix D is unknown so that we use a two-step procedure to perform the estimation. In step one, the ordinary least-squares (OLS) method is used at each time index t to obtain a preliminary estimate of f_t. Second, we plug in the estimate, $\hat{D}$ to obtain a refined estimate of the factor realization.

- The second approach is the **Fama-French approach** proposed by Fama and French (1992). In this approach, the factor realization f_{jt} for a given specific fundamental is obtained by constructing some hedge portfolio based on the observed fundamental.
  - For a given asset fundamental (e.g., ratio of book-to-market value), Fama and French (1992) determined factor realizations using a two-step procedure.

Fundamental factor models III

- First, they sorted the assets based on the values of the observed fundamental. Then they formed a hedge portfolio, which is long in the top quintile (1/3) of the sorted assets and short in the bottom quintile of the sorted assets. The observed return on this hedge portfolio at time t is the observed factor realization for the given asset fundamental.
- The procedure is repeated for each asset fundamental under consideration.
- Finally, given the observed factor realizations $\{f_t | t = 1, ..., T\}$, the betas for each asset are estimated using a time series regression method.
- These authors identify three observed fundamentals that explain high percentages of variability in excess returns. The three fundamentals used by Fama and French are (a) the overall market return (market excess return), (b) the performance of small stocks relative to large stocks (SMB, small minus big), and (c) the performance of value stocks relative to growth stocks (HML, high minus low). The size sorted by market equity and the ratio of book equity to market equity is used to define value and growth stocks with value stocks having high book equity to market equity ratio.

Statistical factor analysis I

Suggested Reading

- Tsay (2010) **Chapter 9**
- Tsay (2014) **Chapter 6**