Factor models

Camilo A. García Trillos

Market risk and portfolio theory

Factor Models

A factor model for a random variable Y is an expression of Y in terms of other random variables $F_1, \dots, F_n$ called *the factors* such that

$$Y = h(F_1, \dots, F_n) + \epsilon$$

where ϵ is a random *model error*.

Factor Models are usually ”statistically” driven: data most prominent than economics and mathematical reduction.

Statistical analysis can suggest possible relations that with economic grounds but also consequences of co-dependence.

Why factor models?

There are many market instruments...

Exchange	Number of listed entries
NYSE	2400
LSE	2292
Euronext Paris	1078
Shanghai SE	1041

Plus thousands of standardised futures and options on equity, interest rates and commodities available in different exchanges (ex. CME).

On top of all of these standard instruments, financial institutions produce millions of customised over the counter operations.

Factor models help ease tasks like risk reduction or identifying opportunities to improve performance. They can also be used for pricing purposes.

Types of factors

Observable factors: Are the ones for which we can obtain reliable values from sources known to the market: an index value, some economic indices (GDP, inflation,...).

Unobservable factors: are abstract and are indirectly deduced from the data.

Several models use one of the two or even a mix of these types of variables.

- Example of models based on observable: Fama-French models.
- Examples of models based on unobservable: Black and Scholes, pricing kernel algorithms.

Linear models

The simplest and more common type of factor model occurs when h is affine.

We say that a random variable $Y \in \mathbb{R}$ has a linear model in terms of the factors $F_1, \ldots, F_\kappa \in \mathbb{R}$, if there exist $\alpha, \beta_1, \ldots, \beta_\kappa \in \mathbb{R}$ such that

$$Y = \alpha + \beta_1 F_1 + \ldots + \beta_\kappa F_\kappa + \epsilon$$

where ϵ is a random variable in $\mathbb{R}$ with zero expectation and such that $\mathbb{E}[\epsilon] = 0$.

The parameters $\alpha, \beta_1, \ldots, \beta_\kappa$ are chosen to have ϵ as small as possible (more precisely to minimise $\mathrm{var}(\epsilon)$.

Linear models

If we call $\bar{\Sigma}$ the variance covariance matrix of $\boldsymbol{F} = (F_1, \ldots, F_K)^\top$, and we assume it is invertible. We then have

Linear models(cont.)

We have that $\alpha = \mathbb{E}[Y] - \sum_{i=1}^{\kappa} \beta_i \mathbb{E}[F_i]$. and

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix} = \bar{\bar{\bar{\Sigma}}}^{-1} \begin{bmatrix} \mathrm{cov}(Y, F_1) \\ \vdots \\ \mathrm{cov}(Y, F_\kappa) \end{bmatrix}$$

If we are allowed to modify the factor variables, we can also “avoid” α:

$$Y = \sum_{i=1}^{\kappa} \beta_i (F_i + \alpha_i) + \epsilon$$

where $\sum_{i=1}^{\kappa} \alpha_i = \alpha$. Note that the expression for each β_i is as before.

In python: it can be computed by ”hand” or using sklearn.linear_model.LinearRegression.

Linear model (cont.)

In a discrete probability space (Ex: empirical distribution)

$$\boldsymbol{Y} \to \mathcal{Y} \in \mathbb{R}^{n \times 1}, \qquad \boldsymbol{F} \to \mathcal{M}_{\boldsymbol{F}} \in \mathbb{R}^{n \times k}.$$

We then have

Beta Pricing

Take $Y := R - R^0$ (excess return), and take expectation. We get

$$\mathbb{E}[R] - R^0 = \alpha + \hat{\beta} \cdot \mathbb{E}[\boldsymbol{F}]$$

We say that there is a *multifactor* beta pricing model if there exists $\boldsymbol{\lambda} \in \mathbb{R}^k$ such that

$$\mathbb{E}[R] - R^0 = \boldsymbol{\lambda}^\top \bar{\bar{\bar{\Sigma}}}_F^{-1} cov(\boldsymbol{F}, R).$$

We call λ the *factor risk premium* and $\bar{\bar{\bar{\Sigma}}}_F^{-1} \mathrm{cov}(\boldsymbol{F}, R)$ the vector of betas for some return R.

Hence, λ_l is the additional expected-return of an asset for each unit increase in its *beta*.

The single factor case

A particular case of interest is the one of a single factor: we say there is a *single-factor* beta pricing model with factor f for some random variable f if there exists a constant λ such that

$$\mathbb{E}[R] - R^0 = \lambda \frac{\mathrm{cov}(f, R)}{\mathrm{var}[f]}. \tag{1}$$

Example

There exists as single-factor linear model for the excess of return, with factor M, an SDF for the market, since we showed that for all market instruments,

$$\mathbb{E}[R^i] - R^0 = -R^0 \mathrm{cov}(M, R^i).$$

This is another example of unobservable factor.

Single-factor models with returns as factors

A very particular case occurs if we choose as factor the return of a certain asset.

Assume that $f = R_*$ is the return of a certain asset. Then, the single-factor pricing beta model reads

$$\mathbb{E}[R] - R^0 = \lambda \frac{\mathrm{cov}(R_*, R)}{\mathrm{var}[R_*]}.$$

Plugging in $R = R_*$ we obtain

$$\mathbb{E}[R_*] - R^0 = \lambda \frac{\mathrm{cov}(R_*, R_*)}{\mathrm{var}[R_*]} = \lambda.$$

Thus, the factor risk premium is the ordinary risk premium of R_*. In this case,

$$\mathbb{E}[R] - R^0 = (\mathbb{E}[R_*] - R^0) \frac{\mathrm{cov}(R_*, R)}{\mathrm{var}[R_*]}.$$

Single-factor models with returns as factors (cont.)

Theorem

There is a beta pricing model with a return R_* as the single factor if and only if the return is on the mean-variance frontier (with risk-free asset) and does not equal the risk-free rate.

In particular, as a consequence, we verify again that if all the assumptions of CAPM are correct, the market portfolio coincides with the tangency portfolio (if...).

Proof of the theorem

Beta and Alpha

$$\mathbb{E}[R_t] - R_t^0 = \alpha + \beta(\mathbb{E}[R_t^M] - R_t^0) + \epsilon_t$$

- β : Correlation with market return
- α : Idiosyncratic excess return

Adjustments for illiquidity

When investing in illiquid securities is necessary to adjust for prices that do not change frequently

$$\mathbb{E}[R_t] - R_t^0 = \alpha^{adj} + \beta_0(\mathbb{E}[R_t^M] - R_t^0) + \beta_1(\mathbb{E}[R_{t-1}^M] - R_{t-1}^0)$$
$$+ \ldots + \beta_L(\mathbb{E}[R_{t-L}^M] - R_{t-L}^0) + \epsilon_t$$

Fama-French models

- 3-factor:

$$\mathbb{E}[R] - R^0 = \alpha + \beta_1(\mathbb{E}[R_M] - R^0) + \beta_2 SMB + \beta_3 HML + \epsilon$$

  where *SMB* is difference of returns of two portfolios of 'small' and 'big' stocks; and HML difference in return of two portfolios with high and low Book-to-Market ratio.

- 5-factor:

$$\begin{aligned}\mathbb{E}[R] - R^0 =& \alpha + \beta_1(\mathbb{E}[R_M] - R^0) + \beta_2 SMB + \beta_3 HML \\ &+ \beta_4 RMW + \beta_5 CMA + \epsilon\end{aligned}$$

  where *RMW* is difference of returns of two portfolios with 'strong' and 'weak' profitability; and CMA difference in return of two portfolio with low and high inner investment.

To summarise

We have introduced the notion of factor models to simplify analysis and estimation.

Some of the models we have made before produce single-factor models (SDF, CAPM).

However, the framework extends, essentially relying on data (Ex: Fama-French factors)

A major limitation: we depend on past information, and we need to assume some sort of stationarity (more on that later).