Barra Model

Reading Notes of Factor Investing #Ch7

Sylvia Xu

January 16, 2025

1

[1] Chuan Shi, 2020. "Factor Investing: Methodology and Practice".

Contents

Relationship of Stock and Factor Covariance Matrix

The following equation gives the relationship of stock covariance matrix and factor covariance matrix:

$$\Sigma = \beta \Sigma_\lambda \beta^\top + \Sigma_\varepsilon \tag{1}$$

where:

- $\beta = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_N \end{bmatrix}^\top \in \mathbb{R}^{N \times K}$ is the factor loading matrix;
- $\Sigma \in \mathbb{R}^{N \times N}, \Sigma_\lambda \in \mathbb{R}^{K \times K}, \Sigma_\varepsilon \in \mathbb{R}^{N \times N}$ are covariance matrices of stocks, factors and idiosyncratic return, respectively. $(\Sigma_\varepsilon = \mathrm{diag}(\sigma^2_{\varepsilon,1}, \sigma^2_{\varepsilon,2}, \cdots, \sigma^2_{\varepsilon,N}))$

Barra Multi-factor Model

CNE5 Model as an example

CNE5 Model

Introduction, Interpretation and Solution

CNE5 Model

In the **CNE5 model**, there is **one country factor**, P **industry factors** (denoted as $I_1, \ldots, I_P$), and Q **style factors** (denoted as $S_1, \ldots, S_Q$). Let $K = 1 + P + Q$, representing a total of K factors. At time t, the multi-factor model is given as:

$$\begin{aligned}
\begin{bmatrix} R^e_{1t} \\ R^e_{2t} \\ \vdots \\ R^e_{Nt} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \lambda_{Ct} + \begin{bmatrix} \beta^{I_1}_{1t-1} \\ \beta^{I_1}_{2t-1} \\ \vdots \\ \beta^{I_1}_{Nt-1} \end{bmatrix} \lambda_{I_1 t} + \cdots + \begin{bmatrix} \beta^{I_P}_{1t-1} \\ \beta^{I_P}_{2t-1} \\ \vdots \\ \beta^{I_P}_{Nt-1} \end{bmatrix} \lambda_{I_P t} \\
+ \begin{bmatrix} \beta^{S_1}_{1t-1} \\ \beta^{S_1}_{2t-1} \\ \vdots \\ \beta^{S_1}_{Nt-1} \end{bmatrix} \lambda_{S_1 t} + \cdots + \begin{bmatrix} \beta^{S_Q}_{1t-1} \\ \beta^{S_Q}_{2t-1} \\ \vdots \\ \beta^{S_Q}_{Nt-1} \end{bmatrix} \lambda_{S_Q t} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Nt} \end{bmatrix}
\end{aligned} \tag{2}$$

CNE5 Model (continued)

- R^e_{it} is the excess return of stock i at time t.
- $\beta^{I_p}_{it-1}$ is the factor loading of stock i on industry I_p at time $t-1$, and $\beta^{I_p}_{it-1} \in \{0, 1\}$.
- $\beta^{S_q}_{it-1}$ is the factor loading of stock i on style factor S_q at time $t-1$.
- λ_{C_t} is the return of country factor at time t; λ_{I_pt} is the return of industry factor I_p at time t; λ_{S_qt} is the return of style factor S_q at time t.
- u_{it} is the idiosyncratic return of stock i as time t.

CNE5 - Constraints

Note that there is collinearity between the factor loading of the country factor and the P industry factors. The factor loading vector of the country factor can be expressed as a linear combination of the factor exposure vectors of the P industry factors, but this would result in the solution to model (2) being non-unique.

Therefore, the factor returns of the industry factors are restricted as follows:

$$s_{I_1}\lambda_{I_1 t} + s_{I_2}\lambda_{I_2 t} + \cdots + s_{I_P}\lambda_{I_P t} = 0 \tag{3}$$

Where s_{I_p} is the sum of the market capitalization weights of all stocks belonging to industry I_p.

Standardization of Raw Factor Loadings

To standardize the raw factor loadings, the model performs two steps:

1. **Mean Adjustment**: Assumption that the market portfolio is neutral to any style factor gives that the weighted exposure is zero. Thus, for any factor, the mean adjustment is calculated by subtracting the market-capitalization-weighted factor exposure from the raw factor loading.
2. **Normalization**: After mean adjustment, the factor loadings are divided by their standard deviation.

After mean adjustment, the factor loadings satisfy:

$$\sum_{i=1}^{N} s_i \beta_{it-1}^{q} = 0, \quad q = 1, \dots, Q,$$

where s_i represents the market-cap weight of stock i.

Country Factor

The main difference between the early Barra models for the Chinese market and the CNE5 model is the inclusion of the country factor. The country factor portfolio essentially represents the market portfolio weighted by market capitalization:

Country Factor (continued)

Proof.

Let $R^e_{M_t}$ denote the excess return of the market portfolio at time t:

$$
\begin{aligned}
R^e_{M_t} =& \lambda_{C_t} + \sum_{p=1}^{P} \beta^{p}_{M_t-1} \lambda_{I_p t} + \sum_{q=1}^{Q} \beta^{q}_{M_t-1} \lambda_{S_q t} + u_{Mt} \\
=& \lambda_{C_t} + \sum_{p=1}^{P} \left(\sum_{i=1}^{N} s_i \beta^{p}_{it-1} \right) \lambda_{I_p t} + \sum_{q=1}^{Q} \left(\sum_{i=1}^{N} s_i \beta^{q}_{it-1} \right) \lambda_{S_q t} + \sum_{i=1}^{N} s_i u_{it} \\
=& \lambda_{C_t} + \sum_{p=1}^{P} s_{I_p} \lambda_{I_p t} + 0 \times \lambda_{S_q t} + \sum_{i=1}^{N} s_i u_{it} \\
=& \lambda_{C_t} + 0 + 0 + \sum_{i=1}^{N} s_i u_{it} \approx \lambda_{C_t}
\end{aligned}
$$

Interpretation of Country Factor

In the above derivation:

- The country factor's return rate closely approximates the excess return of the market portfolio. This implies that the country factor portfolio is essentially equivalent to the market portfolio.
- Mathematically, in the CNE5 model, all stocks are assumed to have an exposure of 1 to the country factor. Thus, during cross-sectional regression, the country factor effectively serves as the intercept term.

Solve CNE5 (I)

Apply **Weighted least square (WLS)** to estimate factor returns.
For simplicity, the time subscripts t and $t-1$ are omitted. Vectors and matrices are used for representation.
Model (2) can be expressed as

$$R^e = \beta\lambda + u \tag{4}$$

where β is the factor loading matrix:

$$\boldsymbol{\beta} = \begin{bmatrix} 1 & \beta_1^{I_1} & \cdots & \beta_1^{I_P} & \beta_1^{S_1} & \cdots & \beta_1^{S_Q} \\ 1 & \beta_2^{I_1} & \cdots & \beta_2^{I_P} & \beta_2^{S_1} & \cdots & \beta_2^{S_Q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & \beta_N^{I_1} & \cdots & \beta_N^{I_P} & \beta_N^{S_1} & \cdots & \beta_N^{S_Q} \end{bmatrix} \in \mathbb{R}^{N\times K} \tag{5}$$

Solve CNE5 (II)

Let $\Omega \in \mathbb{R}^{K\times N}$ be the pure factor portfolio weight matrix for all factors.

- The k-th row represents the weights of all N stocks in the pure factor portfolio for factor k.
- Since WLS (Weighted Least Squares) is used, it is necessary to specify a regression weight matrix when solving for factor returns and constructing pure factor portfolios.
- The Barra model assumes that the variance of the idiosyncratic returns of individual stocks is inversely proportional to the square root of their market capitalization, and the following regression weight matrix W is chosen:

$$W = \begin{bmatrix} \frac{\sqrt{s_1}}{\sum_{i=1}^N \sqrt{s_i}} & 0 & \cdots & 0 \\ 0 & \frac{\sqrt{s_2}}{\sum_{i=1}^N \sqrt{s_i}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\sqrt{s_N}}{\sum_{i=1}^N \sqrt{s_i}} \end{bmatrix} \tag{6}$$

Solve CNE5 (III)

The constraint matrix is (with only one constraint):

$$\begin{bmatrix} \lambda_C \\ \lambda_{I_1} \\ \vdots \\ \lambda_{I_P} \\ \lambda_{S_1} \\ \vdots \\ \lambda_{S_Q} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & -\frac{s_{I_1}}{s_{I_P}} & -\frac{s_{I_2}}{s_{I_P}} & \cdots & -\frac{s_{I_{P-1}}}{s_{I_P}} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \lambda_C \\ \lambda_{I_1} \\ \vdots \\ \lambda_{I_{P-1}} \\ \lambda_{S_1} \\ \vdots \\ \lambda_{S_Q} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{7}$$

The matrix on the right-hand side of the equality is the constraint matrix C, it is a $K \times (K-1)$ matrix.

Solve CNE5 (IV)

The stock weight matrix Ω of the pure factor portfolio can be obtained using the constrained least squares method.

$$\Omega = C(C^\top \beta^\top W \beta C)^{-1} C^\top \beta^\top W \tag{8}$$

With Ω, factor returns can be calculated easily

$$\lambda_{kt} = \sum_{i=1}^{N} w_{ki} R_{it}^e, k = 1, \cdots, K \tag{9}$$

[2]

[2] The detailed solution steps will be provided in the appendix.

Pure Factor Portfolios

Pure Factor Portfolios

$\Omega\beta$ is a K-dimensional matrix, where the k-th row represents the exposure of factor k's pure factor portfolio to all K factors.

$$\Omega\beta = \begin{bmatrix} \mathbf{D}_{(1+P)\times(1+P)} & \mathbf{0}_{(1+P)\times Q} \\ \mathbf{0}_{Q\times(1+P)} & \mathbf{I}_{Q\times Q} \end{bmatrix} \tag{10}$$

- The upper-right zero matrix $0_{(1+P)\times Q}$ indicates that the pure factor portfolios of the country factor and the P industry factors have zero exposure to all Q style factors.
- The lower-left zero matrix $0_{Q\times(1+P)}$ indicates that the pure factor portfolios of the Q style factors have zero exposure to the country and industry factors.
- The lower-right identity matrix $I_{Q\times Q}$ indicates that each style factor's pure factor portfolio has a unit exposure to itself and zero exposure to other style factors.
- The upper-left submatrix $D_{(1+P)\times(1+P)}$ represents the exposure of the pure factor portfolios of the country and industry factors to these factors themselves.

Facts about Pure Factor Portfolios

- The pure factor portfolio for the country factor approximates the market portfolio. It is a purely long-only portfolio, with an exposure of 1 to the country factor, positive exposures to all industries, and zero exposure to all style factors.
- The pure factor portfolio for an industry factor is market-neutral. For industry I_p, this portfolio is 100% long in the industry and 100% short in the country factor portfolio, reflecting the excess return of the industry relative to the market portfolio. It has zero exposure to all style factors.
- The pure factor portfolio for a style factor is also market-neutral. For a style factor q, its pure factor portfolio has an exposure of 1 solely to that factor, with no exposure to other industry or style factors. The portfolio achieves the excess return of the risk factor through its exposure to the specific style factor.

Covariance Matrix Estimation

Bias Statistic

Bias Statistic

Bias statistics is a commonly used metric for evaluating the accuracy of risk models (Menchero et al., 2011). It is used to measure **the error between predicted risk values and actual risk values**.

If there is a significant error between these two values, the risk prediction is considered biased. Hence, this metric is referred to as bias statistics.

Assume there are T periods, then the **bias statistic** of stock i is defined as

$$B_i = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(b_{it} - \bar{b}_i)^2} \tag{11}$$

$\bar{b}_i$ is the mean of b_{it}, and b_{it} is defined as:

$$b_{it} = \frac{R_{it}}{\sigma_{it}} \tag{12}$$

where R_{it} is the return of stock i at time t, σ_{it} is the estimate at the beginning of time t. Hence, b_{it} standardizes R_{it} with σ_{it}.

Bias Statistic (continued)

If the ex-ante risk prediction σ_{it} is accurate, the standard deviation of the standardized return b_{it} will be 1.
Assuming that the asset return R_{it} follows a normal distribution and the ex-ante risk prediction is accurate, b_i should also follow a normal distribution when T is sufficiently large. Its 95% confidence interval is given by:

$$B_i \in [1 - \sqrt{2/T}, 1 + \sqrt{2/T}]$$

If B_i is not in this range, we believe that ex-ante risk estimation is inaccurate.

Eigenfactor Adjustment

For systematic risk

Motivation

- Multi-factor models play a crucial role in estimating the relationship between asset returns and factor exposures.
- Accurate estimation of the risk covariance matrix is essential for portfolio optimization.
- A common question: Is the risk estimate $\sigma_p^2 = \omega^\top \Sigma \omega$ unbiased?

$$\sigma_p^2 = \omega^\top \beta \Sigma_\lambda \beta^\top \omega + \omega^\top \Sigma_\varepsilon \omega \approx \omega^\top \beta \Sigma_\lambda \beta^\top \omega$$

- Define $\omega_\lambda \equiv (w^\top \beta)^\top$, then $\sigma_p^2 \approx \omega_\lambda^\top \Sigma_\lambda \omega_\lambda$. It can be interpreted as the portfolio to invest in pure factor portfolios with weight ω_λ.
- Studies show that pre-risk estimates for certain factor combinations are not accurate, which affects portfolio configurations.

Decomposing the Factor Covariance Matrix

The factor covariance matrix $\hat{\Sigma}_\lambda$ can be decomposed as:

$$\hat{\Sigma}_\lambda = UDU^\top$$

- U: Eigenvector matrix of $\hat{\Sigma}_\lambda$ (orthogonal, $U \top U = I$), the i-th column of U is the i-th eigenvector of $\hat{\Sigma}_\lambda$.
- D: Diagonal matrix of eigenvalues, representing the variance of each eigenfactor portfolio.

To interpret:

$$U^\top \hat{\Sigma}_\lambda U = D$$

- Each row of U defines a unique eigenfactor portfolio.
- Eigenfactor portfolios are uncorrelated, with variances along the diagonal of D.

Bias in Factor Estimates

- Studies show that bias exists in factor risk estimates:
  - Smaller variances tend to have larger biases (estimated bias $\gg 1.0$).
  - Larger variances tend to have smaller biases.
- Accurate adjustment of these biases is crucial for effective risk management.
- Menchero et al.(2011): **Eigenfactor Adjustment Method**.

Key Steps in Adjustment

The core of the factor adjustment method is the bootstrap approach. This method uses the sample factor covariance matrix, $\hat{\Sigma}_\lambda$, as the known parameter for the data generating process. By simulating the parameters, the method estimates an adjusted factor covariance matrix. The steps are as follows:

1. Perform eigen decomposition on $\hat{\Sigma}_\lambda$:

$$\hat{\Sigma}_\lambda = UDU^\top$$

2. Let $D(k)$ denote the kth diagonal element of D. For the kth factor portfolio, assume returns satisfy $N(0, D(k))$.
   For the mth bootstrap iteration, simulate T-period returns for the kth factor portfolio, obtaining the factor return matrix $\hat{b}_m \in \mathbb{R}^{K\times T}$.
   Compute the factor return matrix:

$$\tilde{\lambda}_m = Ub_m$$

Key Steps in Adjustment (continued)

3. Calculate the "sample" covariance matrix $\tilde{\Sigma}_{\lambda_m}$ of $\tilde{\lambda}_m$:

$$\tilde{\Sigma}_{\lambda_m} = \mathrm{cov}(\tilde{\lambda}_m, \tilde{\lambda}_m)$$

4. Perform eigen decomposition on $\tilde{\Sigma}_{\lambda_m}$ to derive $\tilde{U}_m$ and $\tilde{D}_m$.
5. Use the "true" covariance matrix $\hat{\Sigma}_\lambda$ and $\tilde{D}_m$ to calculate the true covariance matrix:

$$D_m = \tilde{U}_m^\top \hat{\Sigma}_\lambda \tilde{U}_m$$

6. Repeat steps (2)-(5) for M iterations. For the kth diagonal element, calculate the average bias:

$$\nu(k) = \frac{1}{M} \sum_{m=1}^{M} \sqrt{\frac{D_m(k)}{\tilde{D}_m(k)}}$$

Key Steps in Adjustment (continued)

7. The normality and stationarity assumptions of return in simulations are not satisfied in reality. To mitigate the bias, apply a shrinkage adjustment based on:

$$v_s(k) = a\,(v(k) - 1) + 1$$

where a is an empirical constant, typically in $[1, 2]$.

8. Finally, use $v_s(k)$ to adjust D of $\hat{\Sigma}_\lambda$, and use U to recover the covariance matrix:

$$D_0 = v^2 D$$

$$\hat{\Sigma}_{\lambda_0} = U D_0 U^\top$$

Here, v^2 is a K order diagonal matrix, with the kth diagonal entry $v_s(k)^2$; Σ_{λ_0} is the final adjusted factor covariance matrix.

Bayesian Shrinkage

For idiosyncratic risk

Empirical Observations on Idiosyncratic Volatility

- **Key Observation:**
  - **Idiosyncratic volatility (standard deviation of idiosyncratic returns) has poor persistence out of sample.**
  - This indicates that the sample covariance matrix is not a reliable estimate.
- **Example - Analysis by Deciles:**
  - Stocks are divided into deciles based on sample idiosyncratic volatility.
  - The average bias statistics for each decile reveal:
    - **Low idiosyncratic volatility deciles:** Bias statistic significantly greater than 1.0 $\Longrightarrow$ underestimates the ex-post idiosyncratic volatility out of sample.
    - **High idiosyncratic volatility deciles:** Bias statistic significantly less than 1.0 $\Longrightarrow$ overestimates the ex-post idiosyncratic volatility out of sample.

Bayesian Shrinkage for Adjustment

- Use Bayesian shrinkage to adjust sample estimates with priors.
- **Process:**
  - Combine prior beliefs about idiosyncratic volatility with sample data.
  - Linearly combine prior and sample data to estimate the posterior covariance matrix.
- The process pulls sample volatility estimates closer to the prior, a concept termed as **shrinkage**.

Constructing Prior for Idiosyncratic Volatility

- Divide all stocks into 10 groups based on market capitalization.
- For any given group g, calculate the weight for stock i in the group:

$$\omega_i = \frac{s_i}{\sum_{j \in M_g} s_j}$$

- Here, M_g: set of all stocks in group g, s_i: market capitalization of stock i, and $\hat{\sigma}_i^i$: sample idiosyncratic volatility of stock i.
- Calculate the weighted average idiosyncratic volatility for all stocks in M_g:

$$\bar{\sigma}_g^i = \sum_{i \in M_g} \omega_i \hat{\sigma}_i^i$$

It is used as the prior idiosyncratic volatility for all stocks in M_g.

Combining Prior and Sample Data

- Combine prior and sample idiosyncratic volatility using Bayesian shrinkage:

$$\hat{\sigma}_i^{bs} = \eta_i \bar{\sigma}_g^i + (1 - \eta_i)\hat{\sigma}_i$$

- $\hat{\sigma}_i^{bs}$: posterior idiosyncratic volatility.
- η_i: shrinkage intensity, calculated as:

$$\eta_i = \frac{q|\hat{\sigma}_i - \bar{\sigma}_g^i|}{\sqrt{\frac{1}{N_g}\sum_{j \in M_g}(\hat{\sigma}_j - \bar{\sigma}_g^i)^2} + q|\hat{\sigma}_i - \bar{\sigma}_g^i|}$$

- N_g: number of stocks in group M_g.
- q: empirical shrinkage coefficient.

Interpretation of Shrinkage Intensity

- η_i depends on two components:
  - $|\hat{\sigma}_i - \bar{\sigma}_g^i|$: deviation between sample and prior.
  - $\sqrt{\frac{1}{N_g} \sum_{j \in M_g} (\hat{\sigma}_j - \bar{\sigma}_g^i)^2}$: standard deviation of deviations in M_g.
- Key takeaways:
  - If $|\hat{\sigma}_i - \bar{\sigma}_g^i|$ is large: rely more on the sample, decrease η_i.
  - If $|\hat{\sigma}_i - \bar{\sigma}_g^i|$ is small: rely more on the prior, increase η_i.
- q is chosen to minimize bias statistics for idiosyncratic volatility across all stocks.

Bayesian Shrinkage Workflow

1. Calculate group-level prior volatility $(\bar{\sigma}_g^i)$ based on market capitalization weights.
2. Combine prior and sample volatility using shrinkage formula:

$$\hat{\sigma}_i^{bs} = \eta_i \bar{\sigma}_g^i + (1 - \eta_i)\hat{\sigma}_i$$

3. Adjust shrinkage intensity η_i based on:
   - Deviation of sample from prior.
   - Variability within the group.
4. Select q to optimize overall statistical performance.

Appendix

CNE5 Solution Steps

Notation and Setup

- $R^e \in \mathbb{R}^N$: Excess returns of N stocks.
- $\beta \in \mathbb{R}^{N\times K}$: Factor loading matrix.
- $\lambda \in \mathbb{R}^K$: Factor returns.
- $W \in \mathbb{R}^{N\times N}$: Diagonal regression-weight matrix.
- $C \in \mathbb{R}^{K\times(K-1)}$: Constraint matrix for factors.
- $\Omega \in \mathbb{R}^{K\times N}$: Weight matrix of pure factor portfolios (each row k is the portfolio for factor k).

Goal: Construct Pure Factor Portfolios

Pure Factor Portfolio Requirements

For each factor k:

1. Exposure to factor k is 1.
2. Exposure to all other factors is 0.
3. Respects any additional constraints imposed by the model.

Exposure Definition

$$\beta^\top \omega_k = \begin{bmatrix} \text{exposure to factor 1} \\ \text{exposure to factor 2} \\ \vdots \\ \text{exposure to factor K} \end{bmatrix}$$

where ω_k is the $N \times 1$ weight vector for the k-th pure factor portfolio.

Weighted Least Squares (WLS) for Factor Portfolios

- We often want
$$\beta^\top \Omega^\top \approx I_{K\times K}$$
in a **weighted least-squares** sense.
- The weight matrix W accounts for the variance structure of the residuals.
- Without constraints, the WLS solution to
$$\min_{\Omega} \|\beta\,\Omega^\top - I\|_W^2$$
would yield
$$\Omega^\top = (\beta^\top\, W\beta)^{-1}\,\beta^\top\, W.$$

[3]

[3] Detailed solution steps of a WLS problem is provided in the next part.

Incorporating Constraints via C

Why Constraints?

- Industry neutrality or other structure: impose $\beta^\top \omega_k$ must lie in a certain subspace.
- The matrix C encodes these linear constraints.
- For example, if we want certain sums of exposures to be zero (or a constant), C projects onto that constraint space.

Constrained WLS Solution

A classical result from linear algebra / regression with constraints is:

$$\Omega \;=\; C\left(C^\top \beta^\top W \beta\, C\right)^{-1} C^\top \beta^\top W.$$

Each row of Ω is the weight vector for a pure factor portfolio satisfying those constraints.

Putting It All Together

1. **Start** with the unconstrained WLS solution:

$$\Omega^\top_{\text{unconstrained}} = (\beta^\top W \beta)^{-1} \beta^\top W.$$

2. **Add linear constraints** using C, which modifies the solution:

$$\Omega = C\left(C^\top \beta^\top W \beta\, C\right)^{-1} C^\top \beta^\top W.$$

3. **Interpretation**: This ensures each row (factor portfolio) has the desired exposures while obeying the constraints.

$$\boxed{\Omega = C\left(C^\top \beta^\top W \beta\, C\right)^{-1} C^\top \beta^\top W}$$

WLS Solution Steps

Weighted Least Squares (WLS) Objective

Defining the Norm

$$\|\beta\,\Omega^\top - I\|_W^2 \;=\; (\beta\,\Omega^\top - I)^\top\, W(\beta\,\Omega^\top - I).$$

In matrix form (often convenient):

$$\|\beta\,\Omega^\top - I\|_W^2 = \mathrm{trace}\Big[(\beta\,\Omega^\top - I)^\top\, W(\beta\,\Omega^\top - I)\Big].$$

Our Goal

$$\min_{\Omega}\ \mathrm{trace}\Big[(\beta\,\Omega^\top - I)^\top\, W(\beta\,\Omega^\top - I)\Big].$$

Expanding the Objective (Conceptually)

$$\mathrm{trace}\Big[(\beta\,\Omega^\top - I)^\top\, W(\beta\,\Omega^\top - I)\Big] = \mathrm{trace}\Big[(\Omega\,\beta^\top - I)\, W(\beta\,\Omega^\top - I)\Big].$$

- Expands to terms involving $\Omega\,(\beta^\top\, W\beta)\,\Omega^\top$ (the quadratic part).
- Linear terms in Ω: $\mathrm{trace}(\Omega\,\beta^\top\, W)$ (and a similar one).
- Constant term: $\mathrm{trace}(W)$.

Taking Derivatives: Normal Equations

Derivative Approach (Sketch)

$$\frac{\partial}{\partial \Omega} \operatorname{trace}\left[(\beta\,\Omega^\top - I)^\top W(\beta\,\Omega^\top - I)\right] = 0.$$

- The key term is $\operatorname{trace}(\Omega\,A\,\Omega^\top B)$.
- Known result:

$$\frac{\partial}{\partial \Omega}\operatorname{trace}(\Omega\,A\,\Omega^\top B) = 2\,\Omega\,(A\,\Omega^\top B)_{\mathsf{sym}} \quad \text{(if } A, B \text{ are symmetric).}$$

- Linear terms: $\operatorname{trace}(M\,\Omega) \to$ derivative is $M^\top$.

Resulting Normal Equation

$$\Omega\,(\beta^\top W\beta) = W\beta.$$

Solving the Normal Equation

From

$$\Omega\,(\beta^\top W\beta) = W\beta,$$

we multiply both sides (on the right) by $(\beta^\top W\beta)^{-1}$:

$$\Omega = W\beta\,(\beta^\top W\beta)^{-1}.$$

Taking transpose:

$$\Omega^\top = (\beta^\top W\beta)^{-1}\,\beta^\top W.$$

Conclusion

$$\boxed{\Omega^\top = (\beta^\top W\beta)^{-1}\,\beta^\top W} \quad \iff \quad \Omega = W\beta\,(\beta^\top W\beta)^{-1}.$$