# Quantitative Investment

S. G. Kou
Questrom School of Business
Boston University

This version Spring 2020

# Chapter 4

# Black-Litterman Model for Asset Allocation

## 4.1 The Basic Setting

Black and Litterman propose a Bayesian model that tries to solve the problems we have with the mean variance optimizer that (1) portfolios are often highly leveraged, and (2) portfolio allocations change drastically with small changes in the forecasts.

The ingredients of the Black-Litterman (BL) model are: a model for market returns, a prior on expected return (e.g. the CAPM), and views of how certain asset classes are going to behave. The result is a set of return forecasts that significantly reduce the pitfalls of the mean variance optimizer, and give a way to blend model and views in a consistent way.

More to the point, the BL model does not give a "best" model based on the data. Rather, it gives a model that is make your subjective views consistent with the market data.

Let $\mathbf{X}_{N\times 1}$ be the observed excess return of $N$ risky asset with the mean $\mathbf{R}_{N\times 1}$ and covariance matrix $\mathbf{C}_{N\times N}$. Note that the risk-free rate has already been taken out from $\mathbf{R}_{N\times 1}$.

The BL model has three components.

(1) The BL model assumes the normality and that $\mathbf{C}_{N\times N}$ is known. The assumption is based on the belief that it is much harder to estimate the mean $\mathbf{R}_{N\times 1}$ than covariance matrix $\mathbf{C}_{N\times N}$, and the fact that assuming normality the impact of the mean $\mathbf{R}_{N\times 1}$ to the optimal asset allocation is much more than that of the covariance matrix $\mathbf{C}_{N\times N}$. In terms of mathematics, this assumption implies that

$$\mathbf{X}_{N\times 1}|\mathbf{R}_{N\times 1} \sim N(\mathbf{R}_{N\times 1},\ \mathbf{C}_{N\times N}).$$

Of course, one can use various ways to estimate $\mathbf{C}_{N\times N}$.

Recall that the $d$-dim normal density with mean $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Sigma}$ is given by

$$\begin{aligned} f(x_1, x_2, ..., x_d) &= \frac{1}{\sqrt{(2\pi)^d |\boldsymbol{\Sigma}|}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})\right\} \\ &\propto \exp\left\{-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x} + \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right\}, \end{aligned}$$

where $|\boldsymbol{\Sigma}|$ is the determinant of $\boldsymbol{\Sigma}$.

(2) The BL model assumes that $\mathbf{R}_{N\times 1}$ has a prior distribution

$$\mathbf{R}_{N\times 1} \sim N(\boldsymbol{\pi}_{N\times 1},\ \boldsymbol{\Sigma}_{N\times N}).$$

The two parameters $\boldsymbol{\pi}_{1\times N}$ and $\boldsymbol{\Sigma}_{N\times N}$ are given.

(3) The subjective views are given in a linear form

$$\mathbf{P}_{k\times N} \cdot \mathbf{R}_{N\times 1} = \mathbf{q}_{k\times 1} + \boldsymbol{\varepsilon}_{k\times 1},$$

where $\mathbf{P}_{k\times N}$ and $\mathbf{q}_{k\times 1}$ are given, and $\boldsymbol{\varepsilon}_{k\times 1}$ is a noise vector with

$$\boldsymbol{\varepsilon}_{k\times 1} \sim N(\mathbf{0}_{k\times 1}, \boldsymbol{\Omega}_{k\times k}),$$

and the covariance matrix $\boldsymbol{\Omega}_{k\times k}$ is assumed to be known.

## 4.2 How to Specify the Parameters in the BL Model

To illustrate the model, we shall consider a simple case that there are three global assets, $N = 3$, US stocks, Euro stocks, and Japanese stocks, with the following table

| Asset | Market Cap |
|---|---|
| US Stocks | 35 |
| Euro Stocks | 10 |
| Japanese Stocks | 5 |

Also assume that we have estimated the variance-covariance matrix $\mathbf{C}_{3\times 3}$ as

| | US | Euro | Japan |
|---|---|---|---|
| US | 7.344% | 2.015% | 3.309% |
| Euro | 2.015% | 4.410% | 1.202% |
| Japan | 3.309% | 1.202% | 3.497% |

### 4.2.1 How to choose $\boldsymbol{\pi}_{N\times 1}$

The BL model uses the expected return that are implied by the mean variance efficient market portfolio. More precisely, since in equilibrium we would expect that for a mean variance efficient portfolio

$$w = \xi \mathbf{C}^{-1}\boldsymbol{\pi},$$

where we can use (at least for a U.S. based investor) the implied $\xi$ from historical U.S. stock indices

$$\xi = \frac{\sigma_*^2}{\rho - r} = \frac{0.16^2}{0.06} = 0.42667,$$

we have an estimator for $\boldsymbol{\pi}$,

$$\boldsymbol{\pi} = \frac{1}{\xi}\mathbf{C}w.$$

Note that the risk-free rate has already been taken out from $\boldsymbol{\pi}_{1\times N}$. In our case, we can choose

$$w = (\frac{35}{35+10+5}, \frac{10}{35+10+5}, \frac{5}{35+10+5})^\top,$$

from which we can compute $\boldsymbol{\pi}$.

### 4.2.2 How to choose $\boldsymbol{\Sigma}$

The matrix $\boldsymbol{\Sigma}_{N\times N}$ reflects the confidence of the fund manager in the equilibrium market portfolio. A strong reliance will be described by a "small" $\boldsymbol{\Sigma}_{N\times N}$. Black and Litterman suggest to set

$$\boldsymbol{\Sigma}_{N\times N} = \tau \mathbf{C}_{N\times N},$$

where $\tau$ is a subjective constant that is needed to input. If $\tau$ is small, then the investor will have a great confidence in the market equilibrium; if $\tau$ is large, then the investor will have a small confidence in the market equilibrium.

#### How to Choose $\mathbf{P}_{k\times N}$, $\mathbf{q}_{k\times 1}$, and $\boldsymbol{\Omega}_{k\times k}$

The subjective view $\mathbf{P}_{k\times N}\cdot\mathbf{R}_{N\times 1} = \mathbf{q}_{k\times 1} + \boldsymbol{\varepsilon}_{k\times 1}$ takes into account the fact that fund managers generally have opinions about how certain markets will behave in the future in absolute and relative terms. These opinions will generate portfolios different from the equilibrium one.

The matrix $\mathbf{P}_{k\times N}$ has $k$ rows (the number of views) and $N$ columns (the number of assets). In general, $k < N$, i.e. it is not required to express views on all asset classes. The matrix $\mathbf{P}$ allows to consider either relative views, either absolute views.

For example, if a fund manager says that "I'm pretty sure that US stocks will grow by the 2.5% in the next year subject to 1% standard error," which is an absolute view, and also says that "I'm confident that European market will grow more than the US market by the 2% subject to 1.5% standard error," which is a relative view, then we have

$$\mathbf{P} = \begin{bmatrix} +1 & 0 & 0 \\ -1 & +1 & 0 \end{bmatrix}, \quad \mathbf{q} = \begin{bmatrix} +2.5\% \\ +2\% \end{bmatrix}.$$

**Assumption:** $\mathbf{P}$ has a full rank, i.e. all views are not redundant.

The covariance matrix $\mathbf{\Omega}_{k\times k}$ expresses the confidence of the fund manager in his opinions expressed in terms of $\mathbf{P}$ and $\mathbf{q}$. It is commonly assumed that $\mathbf{\Omega}$ is a diagonal matrix, which means that the views expressed have independent errors. In our example assuming independence we have

$$\mathbf{\Omega} = \begin{bmatrix} (1\%)^2 & 0 \\ 0 & (1.5\%)^2 \end{bmatrix}.$$

## 4.3 Solution of the BL Model

We have the following result:

**Theorem 1:** The posterior expectation is given by

$$E\left[\mathbf{R}|\mathbf{q}\right] = \boldsymbol{\pi} + \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{q} - \mathbf{P}\boldsymbol{\pi})$$

and the posterior covariance is given by

$$Var\left[\mathbf{R}|\mathbf{q}\right] = \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{P}\mathbf{\Sigma}).$$

In summary, since $\mathbf{X}_{N\times1}|\mathbf{R}_{N\times1} \sim N(\mathbf{R}_{N\times1},\ \mathbf{C}_{N\times N})$,

$$E\left[\mathbf{X}|\mathbf{q}\right] = E\left[E\left[\mathbf{X}|\mathbf{R}\right]|\mathbf{q}\right] = E\left[\mathbf{R}|\mathbf{q}\right] = \boldsymbol{\pi} + \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{q} - \mathbf{P}\boldsymbol{\pi}),$$

$$\begin{aligned} Var(\mathbf{X}|\mathbf{q}) &= E\left[Var\left[\mathbf{X}|\mathbf{R}\right]|\mathbf{q}\right] + Var\left[E\left[\mathbf{X}|\mathbf{R}\right]|\mathbf{q}\right] \\ &= \mathbf{C} + Var(\mathbf{R}|\mathbf{q}) \\ &= \mathbf{C} + \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{P}\mathbf{\Sigma}), \end{aligned}$$

the optimal weights in the BL model is

$$w^* = \xi\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{R}},$$

where

$$\begin{aligned} \tilde{\mathbf{R}} &= E\left[\mathbf{X}|\mathbf{q}\right] = \boldsymbol{\pi} + \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{q} - \mathbf{P}\boldsymbol{\pi}), \\ \tilde{\mathbf{C}} &= Var(\mathbf{X}|\mathbf{q}) = \mathbf{C} + \mathbf{\Sigma} - \mathbf{\Sigma}\mathbf{P}^\top(\mathbf{P}\mathbf{\Sigma}\mathbf{P}^\top + \mathbf{\Omega})^{-1}(\mathbf{P}\mathbf{\Sigma}). \end{aligned}$$

To prove Theorem 1, we need the following result:

**Lemma 1** (Sherman-Morrison-Woodbury Formula, see Hager, 1989, SIAM Review). Suppose matrices $\mathbf{A}$, $\mathbf{D}$ and $\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}$ are invertible. Then

$$(\mathbf{A}+\mathbf{U}\mathbf{D}^{-1}\mathbf{V})^{-1}=\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1}.$$

*Proof.* We shall prove this by direct verification

$$\begin{aligned}
& \left(\mathbf{A}^{-1}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1}\right)(\mathbf{A}+\mathbf{U}\mathbf{D}^{-1}\mathbf{V})\\
= \; & \mathbf{I}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}+\mathbf{A}^{-1}\mathbf{U}\mathbf{D}^{-1}\mathbf{V}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\mathbf{D}^{-1}\mathbf{V}\\
= \; & \mathbf{I}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}+\mathbf{A}^{-1}\mathbf{U}\mathbf{D}^{-1}\mathbf{V}\\
& \quad -\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U})\mathbf{D}^{-1}\mathbf{V}+\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{D}\mathbf{D}^{-1}\mathbf{V}\\
= \; & \mathbf{I}-\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}+\mathbf{A}^{-1}\mathbf{U}\mathbf{D}^{-1}\mathbf{V}\\
& \quad -\mathbf{A}^{-1}\mathbf{U}\mathbf{D}^{-1}\mathbf{V}+\mathbf{A}^{-1}\mathbf{U}\left(\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\\
= \; & \mathbf{I},
\end{aligned}$$

from which the result is proved.

Lemma 1 is useful in two aspects. First, it show that the inverse of $\mathbf{A}+\mathbf{U}\mathbf{D}^{-1}\mathbf{V}$ exists. Secondly, it gives the inverse of $\mathbf{A}+\mathbf{U}\mathbf{D}^{-1}\mathbf{V}$ in terms of the inverse of $\mathbf{D}+\mathbf{V}\mathbf{A}^{-1}\mathbf{U}$, which is especially useful if the latter matrix is a much smaller matrix. In our proof of Theorem 1 below, the first matrix will have dimension $N\times N$, while the second matrix has the dimension $k\times k$, and the number of views $k$ is typically smaller than the number of assets $N$.

*Proof of Theorem 1.* Since

$$\mathbf{P}\mathbf{R}=\mathbf{q}+\boldsymbol{\varepsilon},$$

we have

$$\mathbf{q}|\mathbf{R}\sim N(\mathbf{P}\mathbf{R},\boldsymbol{\Omega})$$

The posterior density of $\mathbf{R}|\mathbf{q}$ is given by

$$\begin{aligned}
& f(\mathbf{R}|\mathbf{q})\\
\propto \; & f(\mathbf{q}|\mathbf{R})f(\mathbf{R})\\
\propto \; & \exp\left\{-\frac{1}{2}(\mathbf{q}-\mathbf{P}\mathbf{R})^\top\boldsymbol{\Omega}^{-1}(\mathbf{q}-\mathbf{P}\mathbf{R}))\right\}\cdot\exp\left\{-\frac{1}{2}(\mathbf{R}-\boldsymbol{\pi})^\top\boldsymbol{\Sigma}^{-1}(\mathbf{R}-\boldsymbol{\pi}))\right\}.
\end{aligned}$$

Some algebra leads to

$$\begin{aligned}
& -\frac{1}{2}(\mathbf{q}-\mathbf{PR})^\top \mathbf{\Omega}^{-1}(\mathbf{q}-\mathbf{PR})) - \frac{1}{2}(\mathbf{R}-\boldsymbol{\pi})^\top \mathbf{\Sigma}^{-1}(\mathbf{R}-\boldsymbol{\pi})) \\
= & -\frac{1}{2}\left\{\mathbf{q}^\top \mathbf{\Omega}^{-1}\mathbf{q} - 2(\mathbf{PR})^\top \mathbf{\Omega}^{-1}\mathbf{q} + (\mathbf{PR})^\top \mathbf{\Omega}^{-1}(\mathbf{PR})\right\} \\
& \quad -\frac{1}{2}\left\{\mathbf{R}^\top \mathbf{\Sigma}^{-1}\mathbf{R} - 2\mathbf{R}^\top \mathbf{\Sigma}^{-1}\boldsymbol{\pi} + \boldsymbol{\pi}^\top \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right\} \\
= & -\frac{1}{2}\left\{\mathbf{q}^\top \mathbf{\Omega}^{-1}\mathbf{q} + (\mathbf{PR})^\top \mathbf{\Omega}^{-1}(\mathbf{PR})) + \mathbf{R}^\top \mathbf{\Sigma}^{-1}\mathbf{R} + \boldsymbol{\pi}^\top \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right\} \\
& \quad +(\mathbf{PR})^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{R}^\top \mathbf{\Sigma}^{-1}\boldsymbol{\pi} \\
\propto & -\frac{1}{2}\left\{(\mathbf{PR})^\top \mathbf{\Omega}^{-1}(\mathbf{PR})) + \mathbf{R}^\top \mathbf{\Sigma}^{-1}\mathbf{R}\right\} \\
& \quad +(\mathbf{PR})^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{R}^\top \mathbf{\Sigma}^{-1}\boldsymbol{\pi} \\
= & -\frac{1}{2}\mathbf{R}^\top \left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{\Sigma}^{-1}\right)\mathbf{R} + \mathbf{R}^\top \left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right).
\end{aligned}$$

Note that $\mathbf{\Omega} + \mathbf{P\Sigma P}^\top$ is positive definite (hence invertible), because $\mathbf{\Omega}$ is positive definite,

$$\mathbf{P\Sigma P}^\top = \mathbf{P\Lambda\Lambda}^\top \mathbf{P}^\top,$$

where $\mathbf{\Sigma} = \mathbf{\Lambda\Lambda}^\top$ by Cholesky decomposition, and $\mathbf{P}$ is of full rank.

Using Lemma 1 we have

$$(\mathbf{\Sigma}^{-1} + \mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P})^{-1} = \mathbf{\Sigma} - \mathbf{\Sigma P}^\top \left(\mathbf{\Omega} + \mathbf{P\Sigma P}^\top\right)^{-1} \mathbf{P\Sigma},$$

because $\mathbf{\Sigma}^{-1}$ are $\mathbf{\Omega} + \mathbf{P\Sigma P}^\top$ are invertible. Thus,

$$f(\mathbf{R}|\mathbf{q}) \propto \exp\left\{-\frac{1}{2}\mathbf{R}^\top \mathbf{A}^{-1}\mathbf{R} + \mathbf{R}^\top \mathbf{A}^{-1} \cdot \mathbf{A}\left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right)\right\},$$

where

$$\mathbf{A}^{-1} = \mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{\Sigma}^{-1}.$$

Comparing this with the multivariate normal density, the posterior density is again normal with mean

$$\begin{aligned}
E(\mathbf{R}|\mathbf{q}) &= \mathbf{A}\left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right) \\
&= \left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{\Sigma}^{-1}\right)^{-1}\left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{q} + \mathbf{\Sigma}^{-1}\boldsymbol{\pi}\right),
\end{aligned}$$

and covariance

$$Var(\mathbf{R}|\mathbf{q}) = \mathbf{A} = \left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{\Sigma}^{-1}\right)^{-1}.$$

Thus, the covariance can be simplified as

$$Var(\mathbf{R}|\mathbf{q}) = \left(\mathbf{P}^\top \mathbf{\Omega}^{-1}\mathbf{P} + \mathbf{\Sigma}^{-1}\right)^{-1} = \mathbf{\Sigma} - \mathbf{\Sigma P}^\top \left(\mathbf{P\Sigma P}^\top + \mathbf{\Omega}\right)^{-1} \mathbf{P\Sigma},$$

which is what is stated in the theorem. As for the simplification of $E(\mathbf{R}|\mathbf{q})$, we have

$$\begin{aligned}
& \left(\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{P}+\boldsymbol{\Sigma}^{-1}\right)^{-1}\left(\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi}\right) \\
= & \left(\boldsymbol{\Sigma}-\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\right)\left(\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}+\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi}\right) \\
= & \boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}-\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q} \\
& \quad +\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi}-\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\boldsymbol{\Sigma}^{-1}\boldsymbol{\pi} \\
= & \boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}-\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q} \\
& \quad +\boldsymbol{\pi}-\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\mathbf{P}\boldsymbol{\pi} \\
= & \boldsymbol{\pi}+\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\left[\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)\boldsymbol{\Omega}^{-1}\mathbf{q}-\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}-\mathbf{P}\boldsymbol{\pi}\right] \\
= & \boldsymbol{\pi}+\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\left[\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}+\mathbf{q}-\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top\boldsymbol{\Omega}^{-1}\mathbf{q}-\mathbf{P}\boldsymbol{\pi}\right] \\
= & \boldsymbol{\pi}+\boldsymbol{\Sigma}\mathbf{P}^\top\left(\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}^\top+\boldsymbol{\Omega}\right)^{-1}\left[\mathbf{q}-\mathbf{P}\boldsymbol{\pi}\right],
\end{aligned}$$

and the proof is finished.

## 4.4 The 1/N Rule

DeMiguel et al. (2007, Review of Financial Studies) evaluate the out-of-sample performance of the portfolio policy from the sample-based mean-variance portfolio model and the various extensions of this model (including the James-Stein estimators, Shrinkage towards the minimal variance, shrinkage of the covariance matrix), designed to reduce the impact of estimation error relative to the benchmark strategy of investing a fraction $1/N$ of wealth in each of the N assets available. Of the fourteen models of optimal portfolio choice that they evaluate across seven empirical data sets, they find that none is consistently better than the $1/N$ rule in terms of Sharpe ratio, certainty-equivalent return, or turnover. This finding indicates that, out of sample, the gain from optimal diversification is more than offset by estimation error.

To gauge the severity of estimation error, they derive analytically the length of the estimation window needed for the sample-based mean-variance strategy to outperform the $1/N$ benchmark based on normal distributions; for parameters calibrated to U.S. stock market data, and they find that for a portfolio with only 25 assets, the estimation window needed is more than 3,000 months, and for a portfolio with 50 assets, it is more than 6,000 months, although in practice these parameters are estimated using 120 months of data. Using simulated data, they further document that even the various extensions to the sample-based mean-variance model designed to deal with estimation error reduce only moderately the estimation window needed to outperform the naive 1/N benchmark.

This suggests that there are still many miles to go before the gains promised by optimal portfolio choice can actually be realized out of sample.

For example, an open problem is how to incorporate leptokurtic and dependence features into the estimation of the optimal portfolio weights, as James-Steins estimator and various shrinkage are all derived under i.i.d. normal distribution assumption.

In short, modern finance only starts in 1950's and still in its infancy.

HWK:

Using the data in this lecture notes, compute the optimal asset allocation in the Black-Litterman model for US, Euro, and Japanese stocks for $\tau = 0.1$, 1, and 10.