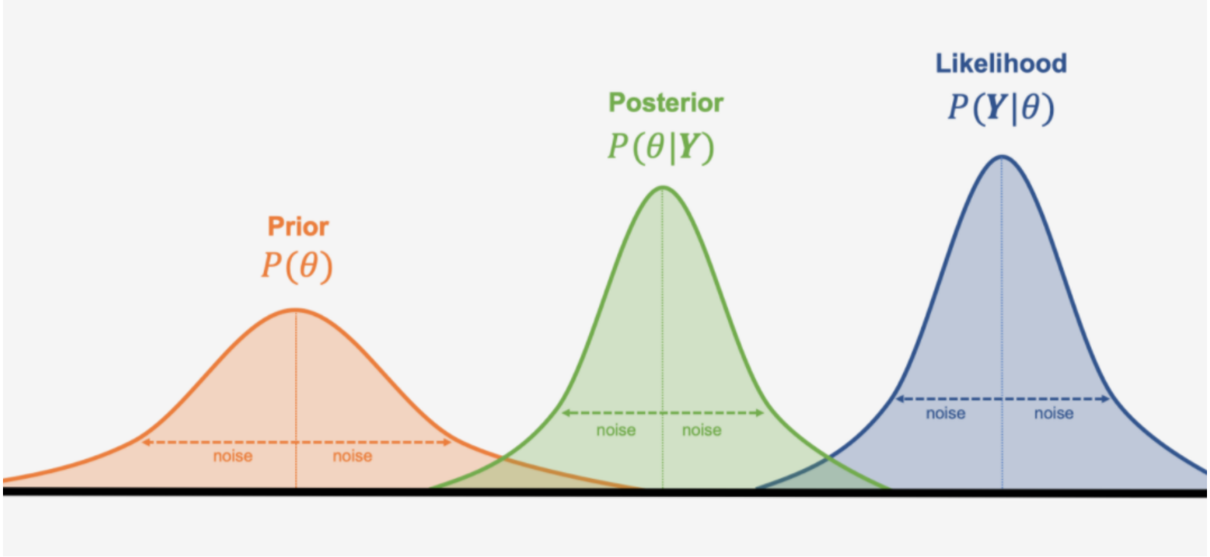


Bayesian Portfolio Optimization: Introducing the Black Litterman Model

BL model

- one of the many successfully used portfolio allocation models
- combines Capital Asset Pricing Theory (CAPM) with Bayesian statistics and Markowitz's modern portfolio theory (Mean-Variance Optimization) to produce efficient estimates of the portfolio weights

Why is it favored by practitioners in the industry?

- investors worked under the assumption that the risk and return relationship of a portfolio was linear, meaning that if an investor wanted higher returns, they would have to take on a higher level of risk.
- Harry Markowitz introduced Modern Portfolio Theory (MPT), which introduced the notion that the diversification of a portfolio can inherently decrease the risk of a portfolio.
- Simply put, this meant that investors could increase their returns while also reducing their risk.

- Markowitz's work on MPT was groundbreaking in the world of asset allocation, eventually earning him a Nobel prize for his work in 1990.
- even though it had a sound theory supporting it, MPT failed to produce favorable results in practice.
- The weights produced by the models did not match with the investors' knowledge and failed to account for a lot of unexpected market variables.

Shortcomings of MVT

- High Input Sensitivity: MPT relies on the past performance of the market (returns and covariance matrix) to get an estimate of the portfolio holdings in the current market conditions. A big problem with this approach is that the performance of the past never provides a guarantee for the market situations that arise in the future. Small estimation errors in the past data leads to highly erroneous mean-variance portfolios.
- Highly Concentrated Portfolios: The mean-variance procedure is found to be greedy and generates highly concentrated portfolios with only a few securities receiving the majority of allocations.
- Neglecting Investor Knowledge: Finally, mean-variance optimization does not take into account an investor's personal knowledge of the market conditions and intuition. In my opinion, this point can be considered to be the most important of the previously listed points. Rather than just relying on historical data, the models should also incorporate an investor's own views of the market which is a very important asset. MVO, instead, chooses to use only historical data with a neglect for valuable market knowledge.

Behind the Scenes: Into the Math of Black-Litterman

Preliminary: Bayesian Theory

- classic Bayes formula equation: $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$
- $P(A) \sim$ prior distribution of A

- $P(B|A) \sim$ likelihood distribution of B given A

The Prior: Implied Excess Equilibrium Returns

The model uses market equilibrium returns as the prior. But what exactly are these “equilibrium” returns?

- CAPM Equation: $E(r_i) = r_f + \beta_i(E(r_{market}) - r_f)$
- Excessive Return: $\Pi = \beta_i(E(r_{market}) - r_f)$

Reverse Optimization

- We easily observe

$$w^i_{market} = \frac{MC_i}{\Sigma_{j=1}^{N} MC_j}$$

- an investor's risk utility

$$R = w^T_{market}\Pi - \frac{\delta}{2} w^T_{market}\Sigma w_{market}$$

- We know Σ is positive definite, then...
- $\frac{\partial^2 R}{\partial w^2_{market}} = -\delta\Sigma < 0$
- which also leads to the R being a concave function. Under no constraints, the risk utility will have a global maximum which will give us a least value of R.
- We differentiate R w.r.t w and equate it to 0

$$\frac{\partial R}{\partial w_{market}} = \Pi - \delta\Sigma w_{market} = 0$$

$$w_{market} = (\delta\Sigma)^{-1}\Pi$$

- Remember that we already have a way of calculating the market weights using the market-cap of the assets, which allows us to *reverse the above equation* and find the desired equilibrium returns...

$$\Pi = \delta \Sigma w_{market}$$

- This gives us our first distribution of the BL Bayesian model

$$\boxed{D_{prior} \equiv N(\Pi, \tau\Sigma)}$$

where Π and Σ carry their usual meanings and τ is a constant of proportionality. The value of τ has been a source of great confusion among practitioners since the original paper also does not clarify its exact significance and what values does it take. I will not go into a lot of details and you will find a lot of papers and articles which explore its significance. It suffices to say that its value is always close to 0 and different papers use different values depending on the use-case.

The Likelihood: Market Views

- The second step is to nail the likelihood distribution. In my opinion, this is one of the most important and ingenious aspects of the BL algorithm and you will soon see why. As we discussed earlier, one of the criticisms of the mean-variance theory was its inability to incorporate investor knowledge. This is important since it allows the weights to be influenced by the ever-changing market factors which an investor can observe and feed into the model.
- For the purpose of this section, let us assume that we have a portfolio comprising of the following assets – Apple, Google, Microsoft, and Tesla – and we have some important market views about some of them. There are 3 main components that will allow investors to incorporate them into the model:

Views Vector: Q

In the Black-Litterman world, there are two types of views:

- Absolute Views: These are views that correspond to the scenario where an investor has assumptions on specific values of the market factors. Considering our example portfolio, let’s say we observe the current market conditions and develop the following views: Apple will yield a monthly expected return of 10%, Microsoft will only give a monthly return of 2%.
- Relative Views: As the name suggests, these other type of views rely on relative performance comparisons between the portfolio assets. So, rather than specific estimates on the performance of Google and Tesla, we have the following view: Google will outperform Tesla (in terms of monthly expected returns) by 6%.

Pick Matrix: P

$$
\begin{array}{c c}
 & \begin{array}{cccc} \textbf{Apple} & \textbf{Google} & \textbf{Microsoft} & \textbf{Tesla} \end{array} \\
\begin{array}{c} \textbf{View - 1} \\ \textbf{View - 2} \\ \textbf{View - 3} \end{array} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}
\end{array}
$$

- An obvious choice is to assign +1 to Apple-Microsoft, -1 to Google-Tesla, and then divide it equally among them. So, individually, Apple and Microsoft get +0.5 while Google and Tesla receive -0.5. The problem is there is no universally accepted method for specifying the pick values and different researchers have proposed different ways of doing so.
- Litterman (2003, pg 82) uses a percentage value for the pick matrix while Satchell and Scowcroft (2000) employ an equal weighting scheme which is what I just described and also shown in the image above. However, I prefer the market-capitalization based assignment proposed in Idzorek (2004) – the weightings are assigned in proportion to the market-cap of the asset

divided by the total market-cap of the sector/group. Based on this method, we get the following pick values: (I have considered the approximate market-cap of these assets at the time of writing)

$$p_{AAPL} = \frac{MC_{AAPL}}{MC_{AAPL} + MC_{MSFT}} = 0.541$$

$$p_{MSFT} = \frac{MC_{MSFT}}{MC_{AAPL} + MC_{MSFT}} = 0.459$$

$$p_{GOOGL} = \frac{MC_{GOOGL}}{MC_{GOOGL} + MC_{TSLA}} = 0.745$$

$$p_{TSLA} = \frac{MC_{TSLA}}{MC_{GOOGL} + MC_{TSLA}} = 0.255$$

Omega Matrix: Ω

$$\Omega = \begin{bmatrix} \omega_1 & 0 & 0 & . & . & 0 \\ 0 & \omega_2 & 0 & . & . & 0 \\ 0 & 0 & . & . & . & 0 \\ 0 & . & . & . & . & 0 \\ 0 & . & . & . & . & 0 \\ 0 & . & . & . & . & \omega_k \end{bmatrix}$$

- Thus one can think of the investor's expected return vector $E_{investor}(r)$ to be described by Q and an error component ϵ

$$E_{investor}(r) \sim Q + \epsilon$$
$$\epsilon \sim N(0, \Omega)$$

- The error in views has 0 mean indicating no bias in an investor's assumptions towards any particular assets in the portfolio. Another assumption is that all the error terms are uncorrelated and independent of each other – the reason for having a diagonal matrix.

$$\Omega = diag(P\tau\Sigma P^T)$$

- where, $diag()$ indicates creating a diagonal matrix from the diagonal elements of the target matrix and τ is a constant of proportionality whose values range from 0 to 1 (In the paper the authors use a τ = 0.05). The assumption here is that the variance in the view errors will be proportional to prior variance and thus can be estimated only from the prior.

$$\boxed{D_{likelihood} \equiv N(Q, \Omega)}$$

- The BL posterior portfolio always oscillates between two extremes – one scenario where the investor has 100% confidence in the views ($\Omega = 0$) and the other where he/she has no confidence in the specified views ($\Omega = \infty$).

$$\begin{tikzcd}
& E[R] \sim N(\mu_{BL}, \Sigma_{BL}) \arrow[dl, "\Omega = \infty"'] \arrow[dr, "\Omega = 0"] & \\
E[R] \sim N(\Pi, \tau\Sigma) & & E[R] \sim N(\mu|Q, \Sigma|Q) \\
\text{Prior} & & \text{Views}
\end{tikzcd}$$

The Posterior: The Master Formula

$$\boxed{\begin{aligned}
&D_{BL} \equiv N(\mu_{BL}, \Sigma_{BL}) \\
&\mu_{BL} = ((\tau\Sigma)^{-1} + P^T\Omega^{-1}P)^{-1}((\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q) \\
&\Sigma_{BL} = ((\tau\Sigma)^{-1} + P^T\Omega^{-1}P)^{-1}
\end{aligned}}$$

Optimal Portfolio Weights

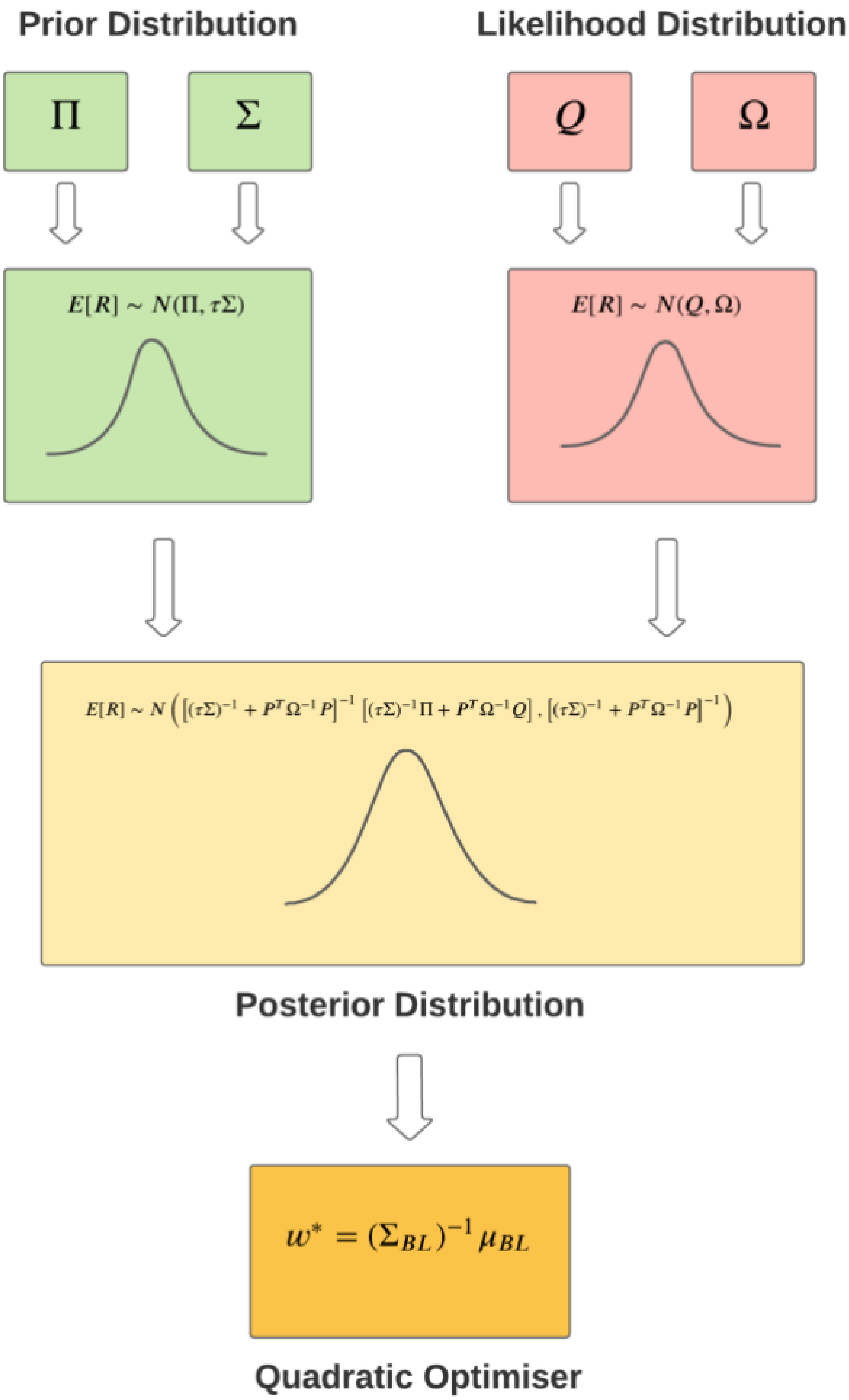
Optimal Portfolio Weights

The final weights can be found by simply plugging the posterior mean and covariance into a standard mean-variance solver. Under no constraints, the following equation holds true.

$$w^* = (\Sigma_{BL})^{-1}\mu_{BL}$$

Note that the new mean and covariance can be passed into any custom optimisation problem and the above equation is true only for a no constraints problem.

Summarize


Prior Distribution
Likelihood Distribution
Π
Σ
Q
Ω
$E[R] \sim N(\Pi, \tau\Sigma)$
$E[R] \sim N(Q, \Omega)$
$E[R] \sim N\left(\left[(\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right]^{-1}\left[(\tau\Sigma)^{-1}\Pi + P^T\Omega^{-1}Q\right], \left[(\tau\Sigma)^{-1} + P^T\Omega^{-1}P\right]^{-1}\right)$
Posterior Distribution
$w^* = (\Sigma_{BL})^{-1}\mu_{BL}$
Quadratic Optimiser