1. Black-Litterman Model

Following relationship is deemed to be the most important assumptions about the Black-Litterman model and Portfolio Expectations and Covariances
.

$$PX = V + \epsilon$$

And, $X \sim N(\mu, \Sigma)$ is independent of $\epsilon \sim N(0, \Omega)$.
These assumptions give rise to the following formulation.

$$Cov(X, v) = Cov(X, PX - \epsilon) = \Sigma P^T$$
$$Cov(v, X) = Cov(PX - \epsilon, X) = P\Sigma$$
$$Cov(v, v) = Cov(PX - \epsilon, PX - \epsilon) = P\Sigma \mathrm{P}^{\mathrm{T}} + \Omega$$
$$\Rightarrow \begin{pmatrix} X \\ v \end{pmatrix} \sim N\left(\begin{pmatrix} \mu \\ P\mu \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma P^T \\ P\Sigma & P\Sigma P^T + \Omega \end{pmatrix}\right)$$

Lemma:

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$$
$$\Rightarrow \begin{pmatrix} \mathrm{I} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} \begin{pmatrix} \mathrm{I} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}^T = \begin{pmatrix} \mathrm{I} & 0 \\ 0 & \Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}$$
$$\begin{pmatrix} \mathrm{I} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix} \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{pmatrix} = \begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \end{pmatrix}$$

When jointly normally distributed random variables have zero covariances, they are independent. Therefore, $X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)$ is independent of $X_1 - \mu_1$. And, its distribution is $N(0, \Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$. This implies that

$$X_2|_{X_1 = x} = N(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1), \Sigma_{12} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}) \ \blacksquare$$

Therefore, the Black-Litterman model is as following:

$$\begin{pmatrix} v \\ X \end{pmatrix} \sim N\left(\begin{pmatrix} P\mu \\ \mu \end{pmatrix}, \begin{pmatrix} P\Sigma P^T + \Omega & P\Sigma \\ \Sigma P^T & \Sigma \end{pmatrix}\right)$$
$$X|_{V=v} \sim N(\mu + \Sigma \mathrm{P}^{\mathrm{T}}(P\Sigma P^T + \Omega)^{-1}(v - P\mu), \Sigma - \Sigma P^T(P\Sigma P^T + \Omega)^{-1} P\Sigma)$$

2. **The relationship between modified Black Litterman and confidence in the research paper, "Return Predictability and Dynamic Asset Allocations by Himanshu Almadi et al.**

$$PX = V + \epsilon$$

$X \sim N(\mu, \Sigma)$ is independent of $\epsilon \sim N(0, \Omega)$

$$\Omega = \begin{pmatrix} \omega_1 & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{pmatrix},$$

$$\omega_{\mathrm{i}} \equiv \left(\frac{1}{c_i} - 1\right) p_i \Sigma p_i', \qquad i \in \{1,2,3\}$$

$$P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix}$$

The case $c_i \to 0$ means that the investor's views have no impact, or in other words the investor is not trusted. The case $c_i \to 1$ means that the investor is trusted completely. The case $c_i \equiv \frac{1}{2}$ gives rise to the situation where the investor is trusted as much as the official market model.

- Case: $c_i \to 0$

The Black-Litterman model assumes the covariance matrix is as following:

$$\Sigma_{\mathrm{BL}} = \Sigma - \Sigma P'(P\Sigma P' + \Omega)^{-1} P\Sigma$$

$c_1 \to 0$ implies that $\Omega \to \begin{pmatrix} \infty & 0 & 0 \\ 0 & \omega_2 & 0 \\ 0 & 0 & \omega_3 \end{pmatrix}$

Denote $M \equiv (P\Sigma P' + \Omega)^{-1} = (M_1, M_2, M_3)$. For example, M_i is a i^{th} column vector of M.
Then, by definition,

$$MP\Sigma P' + M\Omega = \mathrm{M}(P\Sigma P' + \Omega) = I$$

Suppose there exists an element called $m_{i,1}$ of $M_1 = \begin{pmatrix} m_{1,1} \\ m_{2,1} \\ m_{3,1} \end{pmatrix}$ such that $m_{i,1} \neq 0$. Without loss of generality, assume $m_{1,1}, m_{3,1} = 0$ and $m_{2,1} \neq 0$. Then,

$$M\Omega = \begin{pmatrix} 0 & m_{1,2}\omega_2 & m_{1,3}\omega_3 \\ m_{2,1}\omega_2 & m_{2,2}\omega_2 & m_{2,3}\omega_3 \\ 0 & m_{3,2}\omega_2 & m_{3,3}\omega_3 \end{pmatrix} \to \begin{pmatrix} 0 & m_{1,2}\omega_2 & m_{1,3}\omega_3 \\ \pm\infty & m_{2,2}\omega_2 & m_{2,3}\omega_3 \\ 0 & m_{3,2}\omega_2 & m_{3,3}\omega_3 \end{pmatrix}$$

This contradicts the relationship that

$$M\Omega = I - MP\Sigma P'$$

Because every element on the RHS is finite.
Therefore, we must have

$$M_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Then, Letting $\Sigma P' = Q = (q_1, q_2, q_3)$
We have,

$$\Sigma P'(P\Sigma P' + \Omega)^{-1} P\Sigma = QMQ'$$

$$QMQ' = (q_1, q_2, q_3)(M_1, M_2, M_3)\begin{pmatrix} q_1' \\ q_2' \\ q_3' \end{pmatrix} = Q\sum_{i=1}^{3} M_i q_i' = Q(M_2 q_2' + M_3 q_3')$$

Since $\Sigma_{\text{BL}} = \Sigma - \text{QMQ}'$, the confidence of the first view c_1 does not influence Σ_{BL}.
In the case where $c_i \to 0\ \forall i \in \{1,2,3\}$, $\Sigma_{\text{BL}} = \Sigma$ indicating the investor view is not trusted.

- Case: $c_i \to 1$

Without loss of generality, assume $c_1 = 1$.
Then,

$$\begin{pmatrix} (PX)^1 \\ (PX)^2 \\ (PX)^3 \end{pmatrix} = \begin{pmatrix} V^1 \\ V^2 \\ V^3 \end{pmatrix} + \begin{pmatrix} 0 \\ \epsilon^2 \\ \epsilon^3 \end{pmatrix}$$

Therefore, it can be written as

$$(PX)^1 = V^1$$

$$\begin{pmatrix} (PX)^2 \\ (PX)^3 \end{pmatrix} = \begin{pmatrix} V^2 \\ V^3 \end{pmatrix} + \begin{pmatrix} \epsilon^2 \\ \epsilon^3 \end{pmatrix}$$

This implies $PX^1 = V^1$, which implies that the investor's first view is absolutely correct.
It implies that the probability density is infinitely peaked as can be frequently seen by the Dirac delta function. For example, the cumulative density function is almost surely the heaviside step function, which is zero below a threshold value and 1 above the threshold value, and probability density function is the Dirac delta function.

In the case where $c_i = 1\ \forall i \in \{1,2,3\}$, $PX = V$
And, X can be obtained by elementary matrix operations or, for notational convenience, using singular value decomposition $P = U\Lambda W$
Then, $X = W'\Lambda^{-1}U'V$.

- Case: $c_i \equiv \frac{1}{2}$

The illustration of the main idea is much clearer, if we suppose $c_1, c_2, c_3 = \frac{1}{2}$.

$$\Sigma_{\text{BL}} = \Sigma - \Sigma P'(P\Sigma P' + \Omega)^{-1} P\Sigma = (\Sigma^{-1} + P'\Omega^{-1}P)^{-1}$$

Proof of the last equality is presented below.
Since $P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ in the research paper, $\Sigma_{BL} = (\Sigma^{-1} + \Omega^{-1})^{-1}$.

Also, $\Omega^{-1} = \begin{pmatrix} p_1\Sigma p_1' & 0 & 0 \\ 0 & p_2\Sigma p_2' & 0 \\ 0 & 0 & p_3\Sigma p_3' \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_{1,1} & 0 & 0 \\ 0 & \Sigma_{2,2} & 0 \\ 0 & 0 & \Sigma_{3,3} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\Sigma_{1,1}} & 0 & 0 \\ 0 & \frac{1}{\Sigma_{2,2}} & 0 \\ 0 & 0 & \frac{1}{\Sigma_{3,3}} \end{pmatrix}$

Which is roughly equal to Σ^{-1} when covariance terms (non-diagonal terms) are small in magnitudes. Then,

$$\Sigma_{BL} = (\Sigma^{-1} + \Omega^{-1})^{-1}$$

It tells us that information from the market Σ^{-1} accounts for roughly the same amount of trust in comparison to the information from investor views Ω^{-1}.

$\Sigma - \Sigma P'(P\Sigma P' + \Omega)^{-1}P\Sigma = (\Sigma^{-1} + P'\Omega^{-1}P)^{-1}$ can be derived from straightforward matrix computations called the matrix inversion lemma. I have not come across following derivation from other references yet. However, I speculate that this is how matrix inversion lemma was originally derived because this proof does not assume prior knowledge of the solution.

Using elementary gaussian elimination to turn a block matrix into diagonal block matrix,

$$\begin{pmatrix} I & 0 \\ -BA^{-1} & I \end{pmatrix}\begin{pmatrix} A & C \\ B & D \end{pmatrix}\begin{pmatrix} I & -A^{-1}C \\ 0 & I \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & D - BA^{-1}C \end{pmatrix}$$

The gaussian elimination gives rise to the following properties.

$$\begin{pmatrix} I & 0 \\ -BA^{-1} & I \end{pmatrix}^{-1} = \begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix}$$

$$\begin{pmatrix} A & 0 \\ 0 & D - BA^{-1}C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & 0 \\ 0 & (D - BA^{-1}C)^{-1} \end{pmatrix}$$

$$\begin{pmatrix} I & -A^{-1}C \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} I & A^{-1}C \\ 0 & I \end{pmatrix}$$

Therefore,

$$\begin{aligned}\begin{pmatrix} A & C \\ B & D \end{pmatrix}^{-1} &= \begin{pmatrix} I & A^{-1}C \\ 0 & I \end{pmatrix}\begin{pmatrix} A^{-1} & 0 \\ 0 & (D - BA^{-1}C)^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ BA^{-1} & I \end{pmatrix} \\ &= \begin{pmatrix} A^{-1} + A^{-1}C(D - BA^{-1}C)^{-1}BA^{-1} & A^{-1}C(D - BA^{-1}C)^{-1} \\ (D - BA^{-1}C)^{-1}BA^{-1} & (D - BA^{-1}C)^{-1} \end{pmatrix}\end{aligned}$$

Suppose

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix}^{-1} \equiv \begin{pmatrix} X & Y \\ Z & W \end{pmatrix}$$

This, for example, says $X = A^{-1} + A^{-1}C(D - BA^{-1}C)^{-1}BA^{-1}$
Then we have following result,

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix}\begin{pmatrix} X & Y \\ Z & W \end{pmatrix} = \begin{pmatrix} AX + CZ & AY + CW \\ BX + DZ & BY + DW \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

This implies that

$$\begin{pmatrix} D & B \\ C & A \end{pmatrix}\begin{pmatrix} W & Z \\ Y & X \end{pmatrix} = \begin{pmatrix} BY + DW & BX + DZ \\ AY + CW & AX + CZ \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} D & B \\ C & A \end{pmatrix}^{-1} = \begin{pmatrix} W & Z \\ Y & X \end{pmatrix}$$

However, the gaussian elimination would yield

$$\begin{aligned}\begin{pmatrix} D & B \\ C & A \end{pmatrix}^{-1} &= \begin{pmatrix} I & D^{-1}B \\ 0 & I \end{pmatrix}\begin{pmatrix} D^{-1} & 0 \\ 0 & (A - CD^{-1}B)^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ CD^{-1} & I \end{pmatrix} \\ &= \begin{pmatrix} D^{-1} + D^{-1}B(A - CD^{-1}B)^{-1}CD^{-1} & D^{-1}B(A - CD^{-1}B)^{-1} \\ (A - CD^{-1}B)^{-1}CD^{-1} & (A - CD^{-1}B)^{-1} \end{pmatrix}\end{aligned}$$

This means that

$$(A - CD^{-1}B)^{-1} = X = A^{-1} + A^{-1}C(D - BA^{-1}C)^{-1}BA^{-1}$$

Although many letters make the derivation look complicated, only ideas used were gaussian elimination and the fact that inverse of linear map is unique.

3. Understanding PCA: The dimension reduction transformation that tries to retain original variance as much as possible.

Using singular decomposition, the matrix of interest X can be represented as $X = U\Sigma V'$.
For simplicity of illustration, I assume $X \in \mathbb{R}^{n\times 3}$. And, we want to reduce its dimension to $\mathbb{R}^{n\times 2}$.
$U = (u_1, u_2, u_3)$ where $u_i \in \mathbb{R}^{n\times 1}$ is a column vector. Similarly, $V = (v_1, v_2, v_3)$ where $v_i \in \mathbb{R}^{3\times 1}$.

Then, PCA of X denoted as $\hat{X}$ is given by

$$\hat{X} = X(v_1, v_2) = U\Sigma \begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix} (v_1, v_2) = U\Sigma \begin{pmatrix} v_1'v_1 & v_1'v_2 \\ v_2'v_1 & v_1'v_2 \\ v_3'v_1 & v_1'v_2 \end{pmatrix} = U\Sigma \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The last equality used the property of SVD that both U and V are orthogonal.

$$\hat{X} = U\Sigma \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = U \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = (u_1, u_2, u_3) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{pmatrix} = (\sigma_1 u_1, \sigma_2 u_2)$$

Then, as the variance of $\hat{X}$ is approximated using mean of squared deviations, variance $\approx \hat{X}'\hat{X}$.

$$\hat{X}'\hat{X} = \begin{pmatrix} \sigma_1 u_1' \\ \sigma_2 u_2' \end{pmatrix} (\sigma_1 u_1, \sigma_2 u_2) = \begin{pmatrix} \sigma_1^2 u_1' u_1 & \sigma_1\sigma_2 u_1' u_2 \\ \sigma_1\sigma_2 u_2' u_1 & \sigma_2^2 u_2' u_2 \end{pmatrix} = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

The last equality corresponds to the orthogonality property.
I hope to compare this the result above with $X'X$.

$$X'X = V\Sigma^2 V' = (v_1, v_2, v_3) \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix} \begin{pmatrix} v_1' \\ v_2' \\ v_3' \end{pmatrix} = \sigma_1^2 v_1 v_1' + \sigma_2^2 v_2 v_2' + \sigma_3^2 v_3 v_3'$$

We know that $\mathrm{trace}(v_\mathrm{i} v_j') = v_i' v_\mathrm{j} = \begin{cases} 0 \text{ if } i \neq j \\ 1 \text{ if } \mathrm{i} = \mathrm{j} \end{cases}$ by orthogonality.
And, if we let $v_\mathrm{i} = (v_{1,i}, v_{2,\mathrm{i}}, v_{3,\mathrm{i}})$, then $1 = v_\mathrm{i}' v_\mathrm{i} = v_{1,i}^2 + v_{2,\mathrm{i}}^2 + v_{3,\mathrm{i}}^2$ also by orthogonality.
This indicates that each element of v_i is small in magnitudes.
Therefore, assuming that X is normalized so each σ_i^2 is not excessively large,

$$X'X \approx \begin{pmatrix} \sigma_1^2 v_{1,1}^2 & 0 & 0 \\ 0 & \sigma_1^2 v_{2,1}^2 & 0 \\ 0 & 0 & \sigma_1^2 v_{3,1}^2 \end{pmatrix} + \begin{pmatrix} \sigma_2^2 v_{1,2}^2 & 0 & 0 \\ 0 & \sigma_2^2 v_{2,2}^2 & 0 \\ 0 & 0 & \sigma_2^2 v_{3,2}^2 \end{pmatrix} + \begin{pmatrix} \sigma_3^2 v_{1,3}^2 & 0 & 0 \\ 0 & \sigma_3^2 v_{2,3}^2 & 0 \\ 0 & 0 & \sigma_3^2 v_{3,3}^2 \end{pmatrix}$$

In the case where σ_i^2 is large, non-diagonal terms would not be close to 0.

Recalling from linear algebra that orthogonal matrices consist of rotations or reflections, consider a situation where $V \approx \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. This corresponds to the rotation by $0^\circ, 360^\circ$ and so forth. Then,

$$X'X \approx \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$$

Therefore, PCA dimension reduction tries to reduce dimension of the original matrix in a way that the original variance is maintained as much as possible. However, the result depends on the assumption that

- the original matrix is normalized
- rotation/reflection of the $X'X$ is modest (close to identity matrix).

4. Expected Portfolio Returns and Covariance of Portfolio Returns

Expected Holding Period Return of asset i is $E\left[\frac{P_i(T)}{P_i(0)} - 1\right] = E[e^{X_i}] - 1$ where $X = [X_1, \dots, X_n]$.
Since X is normally distributed,

$$E\left[e^{w^T X}\right] = e^{w^T \mu_{BL} + \frac{1}{2} w^T \Sigma_{BL} w}$$

Let $w_i = [0, \dots, \underbrace{1}_{i^{th}}, \dots, 0]^T$

Then, $E[e^{X_i}] = E\left[e^{w_i^T X}\right] = e^{\mu_i + \frac{1}{2}\Sigma_{ii}}$

This implies that $E[e^{X_i} e^{X_j}] = E[e^{X_i + X_j}] = E\left[e^{\left(w_i^T + w_j^T\right) X}\right] = e^{\mu_i + \mu_j + \frac{1}{2}\Sigma_{ii} + \frac{1}{2}\Sigma_{jj} + \Sigma_{ij}}$

Therefore, covariance of the portfolio return is

$$E[e^{X_i} e^{X_j}] - E[e^{X_i}]E[e^{X_j}] = e^{\mu_i + \mu_j + \frac{1}{2}\Sigma_{ii} + \frac{1}{2}\Sigma_{jj}}\left(e^{\Sigma_{ij}} - 1\right)$$