Black-Litterman and Scenario Optimization

Koopa Hakimi

August 23, 2016
Overview of the Original Black Litterman Model

Overview of Copula Opinion Pooling

Applications of Copula Opinion Pooling

Comparison Between COP and Original BL
  CAPM-like equilibrium specification

Overview of the Scenario Optimization Theory

Portfolio Management within COP and Scenario Optimization
Overview of the Original Black Litterman Model

- include investor’s view
- include confidence level
- Starting point for expected return
- Bayesian approach to combine the subjective views with the market equilibrium to get the posterior distribution.
Overview of the Original Black Litterman Model

- The Black-Litterman Model was first published by Fischer Black and Robert Litterman in (1990).

\[ M \sim \mathcal{N}(\mu, \Sigma), \]

- Since BL acknowledged \( \mu \) cannot be known with certainty, they modelled it itself as a multivariate normal where its dispersion is proportional to that of the market:

\[ \mu \sim \mathcal{N}(\pi, \tau \Sigma). \]

- The most common method to calibrate \( \tau \) relies on falling back to basic statistics,

\[ \tau = \frac{1}{\text{number of samples}}. \]

- As in the BL framework, the views are a set of \( k \leq n \) statements on generic linear combinations of the market.
As an example we will examine some investors views. Let’s assume we have 4 securities and 2 views. First, a ”relative” view in which the investor believes that Asset 1 will outperform Asset 3 by 2% with confidence $\Omega_1$. Second, an ”absolute” view in which the investor believes that Asset 2 will return 3% with confidence $\Omega_2$. Therefore, our pick matrix, $q$ and $\Omega$ are as follows:

$$P = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \quad q = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \Omega = \begin{pmatrix} \Omega_1 & 0 \\ 0 & \Omega_2 \end{pmatrix}.$$ 

Note that in BL model, the views are expressed on the market parameters $\mu$, and are assumed to be normally distributed:

$$V \equiv P\mu \sim \mathcal{N}(q, \Omega)$$
By Bayes’ Law, the posterior distribution of the market mean conditional on these views follows the master formula:

\[ \mu | q; \Omega \sim \mathcal{N}(\mu_{BL}, \Sigma_{BL}), \]

\[ \mu_{BL} = \left( (\tau \Sigma)^{-1} + P^T \Omega^{-1} P \right)^{-1} \left( (\tau \Sigma)^{-1} \pi + P^T \Omega^{-1} q \right) \]

\[ \Sigma_{BL} = \left( (\tau \Sigma)^{-1} + P^T \Omega^{-1} P \right)^{-1} \]
Now to compute the distribution of $M$ note that since

$$M = \mu + Z, \quad Z \sim \mathcal{N}(0, \Sigma)$$

Therefore, assuming $\mu$ and $Z$ are independent we have

$$M|_{q;\Omega} = \mu|_{q;\Omega} + Z.$$

Therefore, assuming $\mu$ and $Z$ are independent we have

$$\Sigma_{BL} = \Sigma + \Sigma_{BL}^\mu,$$
Overview of Copula Opinion Pooling

- Normality Assumption as a first problem to BL approach which by its nature is more of a practical issue that can be overcome by applying numerical approaches.

- The second problem in original BL approach is more of a conceptual issue due to its Bayesian nature. Note that in the original BL portfolio managers express their views on the parameters that determine the market distribution; \( \mu \) the expected values. In reality, it is more natural to express views directly on the possible realizations of the market which is not evident under the normality assumption due to clear relationship between the parameters and the realizations.
Overview of Copula Opinion Pooling

\[ V \equiv PM, \]

the prior market distribution \( M \) induces a structure on the marginal distributions of \( V \):

\[ V_k \sim \phi_{V_k}(t) = \phi_{M}(tp_k) \quad t \in \mathbb{R}, \]

Also since each manager can express the subjective views by a univariate distribution

\[ f_{\hat{V}_k}, F_{\hat{V}_k}, \phi_{\hat{V}_k}, \]
Overview of Copula Opinion Pooling

- market induced distribution versus subjective views distribution

\[ \tilde{V}_k \sim F_{\tilde{V}_k} \equiv (1 - c_k)F_{V_k} + c_k F_{\hat{V}_k}, \]

- where the weight \( c_k \in [0, 1] \) can be interpreted as the confidence in the manager’s views.

\[ C \equiv (U_1, \ldots, U_k) = (F_{V_1}(V_1), \ldots, F_{V_n}(V_k)). \]

- Now by considering the reverse structure of copulas and the posterior margins \( (F_{\tilde{V}_k}) \), we get the joint posterior distribution

\[ \tilde{V} \equiv (F_{\tilde{V}_1}^{-1}(U_1), \ldots, F_{\tilde{V}_k}^{-1}(U_k)). \]
Let \( P^\perp \) be an arbitrary \((n - k) \times n\) matrix with rows corresponding to the basis of the null space of \( P \).

\[
\tilde{M} = \begin{pmatrix} P \\ P^\perp \end{pmatrix}^{-1} \begin{pmatrix} \tilde{V} \\ P^\perp M \end{pmatrix}.
\]

The final result is independent of choice of \( P^\perp \).
- iShares MSCI Canada ETF (EWC)
- iShares Trust -1-3 YR TR BD ETF (SHY)
- SPDR Dow Jones International Real Estate ETF (RWX)
- Vanguard Charlotte Funds (VBMFX)
- Vanguard FTSE Emerging Markets (BNDX)
- Vanguard FTSE Europe ETF (VFV.TO)
- Vanguard FTSE Pacific ETF (VWO)
- Vanguard SnP 500 ETF (VGK)
- Vanguard Total Bond Market Index Inv (VPL)
Clearly, since the scatter plot shows some both upper and lower tail dependence, which would suggest to model this features by means of the skew t distribution, see Azzalini (2005):

\[ M \sim ST(\xi, \Omega, \alpha, \nu), \]

where \( \xi \) is a numeric vector of the location parameter of the distribution, \( \alpha \) is a numeric vector which regulates the slant of the density, \( \Omega \) is a symmetric positive-definite matrix, and \( \nu \) is a positive number.

The skew t distribution coincides with the Student t distribution when the shape parameter \( \alpha \) is null.
\[ \nu = 8.109906 \]
\[ \alpha = 10^{-7} \begin{pmatrix} -3.062 \\ -0.0003 \\ -0.675 \\ -0.0252 \\ -0.057 \\ 0.455 \\ -2.762 \\ -1.92 \\ -0.788 \end{pmatrix} \]
Comparison Between COP and Original BL

- Views are made on realizations of the market
- Any multivariate distribution can be used to model the market
- Views can be modeled by any distribution. We are no longer restricted to normality
- The parameters in the model have more realistic meanings
- We can no longer use closed-form expressions
the original BL we start with CAPM model to form portfolio weights.

The CAPM was originally derived under the assumption of either quadratic utility functions or that asset returns are normally distributed.

However, Berk (1997) showed that we can also derive CAPM under the assumption of elliptically symmetric return distributions, see also Ingersoll (1987).

\[ \mu = \delta \Sigma w_{eq}, \]
One of the advantages of COP model is that views can be modelled by any distributions. For example one can model them as uniformly distributed on given ranges; \( U(a, b) \). Assuming the manager is bearish on the SHY (seconed asset in our portfolio), \( U(-0.0005, 0.0005) \) within a day.

The GARCH(1,1) is the simplest and most robust of the family of volatility models. An exponentially weighted moving average of squared returns (integrated GARCH) is quite popular amongst practitioners, most notably RiskMetrics, for forecasting.
Kernel density estimates of posterior and prior
Therefore, in order to evaluate the riskiness of the portfolio in highly non-normal markets, we need a better risk measure which could represent the potential asymmetries and also could capture the extreme events (tail sensitive).

\[
\text{CVaR}(w) = -\mathbb{E}\{w^T R | w^T R \leq F_R^{-1}(1 - \alpha)\},
\]

\[
w(r) \equiv \text{argmin} \{ \text{CVaR}(w) \}
\]

subject to:
\[
w^T \mathbb{E}[M] \geq r,
\]
\[
\sum_{i=1}^{n} w_i = 1,
\]
\[
w_i \geq 0
\]
Rockafellar and Uryasev (2000) show that the mean-CVaR optimization problem can be restated equivalently as follows:

$$w(r) \equiv \arg\min \{ \alpha + \frac{1}{1 - \gamma} \int [-w^T \mathbf{M} - \alpha]^+ f_{\tilde{M}}(x) dx \}$$

subject to:

$$w^T \mathbb{E}[\mathbf{M}] \geq r,$$

$$\sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0$$

$$w_i \geq 0$$
In the case of portfolio optimization, the uncertainty in the optimization process comes from the uncertainty of returns. One can model such uncertainty by distribution which would not resolve the uncertainty. Another way to solve this problem (in the sense of addressing uncertainty, not estimation error) is to solve a very large-scale deterministic program instead, where a large number of scenarios $S$ try to capture randomness by representing the possible future path or paths of the underlying process.
In general following Dembo(1991) a deterministic linear programming is as follow:

Minimize \( R^T x \)
subject to \( Ax = b \)
\( x \geq 0 \)
According to Dembo (1991) we define a scenario as a particular realization of the uncertain data represented by $R_s, A_s$ and $b_s$. Therefore, for each scenario $s \in S$ our problem reduces to the deterministic problem

Minimize $R_s^T x$
subject to $A_s x = b_s,$
$A_d x = b_d$ deterministic constraints
\[ L^1 \text{-norm coordination model:} \]

\[ \min_x \sum_s p_s |R_s^T x - v_s| + \sum_s p_s |A_s x - b_s| \]

subject to \[ A_d x = b_d \]

From now on let \( R = M \sim (r_1, \ldots, r_n) \) be our market consisting of \( n \) securities. Assuming we have \( m \) scenarios, we get a matrix of \( n \times m \) scenarios. Thus in the absent of the scenario constraints; \( A_s = 0, b_s = 0 \)
Mean Absolute Deviation (MAD)

Konno and Yamazaki (1991) introduced Mean Absolute Deviation as an improvement to the Markowitz model. According to Konno and Yamazaki (1991) by considering the $L^1$ risk model as a special case of the piecewise linear risk model they removed most of the difficulties of Markowitz’s model while maintaining its advantages over equilibrium models.

$$\min_{x} \sum_{j=1}^{m} \frac{1}{m} \left| \sum_{i=1}^{n} w_i (r_{i,j} - \nu_i) \right|$$

subject to: $$\sum_{i=1}^{n} \nu_i w_i = C$$

$$\sum_{i=1}^{n} w_i = 1$$
Young (1998) introduced a principle for choosing portfolios based on historical returns data; the optimal portfolio based on this principle is the solution to a simple linear programming problem. This principle uses minimum return rather than variance as a measure of risk. In particular, the portfolio is chosen that minimizes the maximum loss over all past observation periods, for a given level of return:

$$\min \max \left( \sum_{j=0}^{n} -r_{ij}w_j, \forall i = 1, \ldots, m \right),$$
Minimizing Conditional Value at Risk (CVaR)

Rockafellar and Uryasev (2000) show that the mean-CVaR optimization problem can be restated equivalently as follows:

\[ \mathbf{w}(r) \equiv \text{argmin} \left\{ \alpha + \frac{1}{1 - \gamma} \int \left[ -\mathbf{w}^T \mathbf{M} - \alpha \right]^+ f_M(x) dx \right\} \]

subject to:

\[ \mathbf{w}^T \mathbb{E}[\mathbf{M}] \geq r, \]
\[ \sum_{i=1}^{n} w_i = 1, \]
\[ w_i \geq 0 \]
Note that MAD, Standard Deviation and MiniMax measures belong to the general $L^p$ function space with $p = 1$, $p = 2$ and $p = \infty$ respectively.

MiniMax risk measure can be seen as an extreme case of CVaR.
In Scenario Optimization we no longer interested in the distributions of any kind we can assume that the view matrix $P$ expresses the views directly on the target vector $\nu$.

$$V = P\nu.$$ 

By the same argument introduced in COP, let $P^\perp$ be the matrix with each row corresponding to a basis of the null space of $P$. Therefore, we get

$$\begin{pmatrix} V & * \end{pmatrix}_{n \times n} = \begin{pmatrix} P & P^\perp \end{pmatrix}_{n \times n} \nu.$$
By defining the posterior view vector as a weighted average of target and prior view vector,

\[ \tilde{V} = \left( (1 - c_i)\nu_i + c_i v_i \right)_{k \times 1}, \]

we get our posterior forecast:

\[ \bar{\nu} = \left( P P_{\bot} \right)^{-1} \left( \tilde{V} \right)_{n \times n}. \]
Efficient Frontier

Target Risk [CVaR]

Target Return [mean]

Posterior COP
Scenario Optimization With Views
Scenario Optimization No Views

EWP
SHY
BNDX
VBMFX
VFW.TO
VGK
RWX
EWC
VPL
VPL
VWO
VWO
VFV.TO
VFW.TO
According to Sarykalin, Serraino and Uryasev (2008):

- CVaR has superior mathematical properties versus VaR. CVaR is a coherent measure of risk in the basic sense and it is an averse measure of risk. CVaR of a portfolio is a continuous and convex function with respect to optimization variables.

- Risk control using VaR may lead to paradoxical results for skewed distribution. Minimization of VaR may lead to a stretch of the tail of the distribution exceeding VaR. For instance, minimization of 99%-VaR Deviation leads to 13% increase in 99%-CVaR compared to 99%-CVaR of optimal 99%-CVaR Deviation portfolio.

- VaR does not control scenarios exceeding VaR. The indifference of VaR risk measure to extreme tails may be a good property if poor models are used for building distributions. Also, the indifference of VaR to extreme tails may be quite an undesirable property, allowing to take high uncontrollable risks. For instance, so-called ”naked” option positions involve a very small chance of extremely high losses; these rare losses may not be picked by VaR.
CVaR accounts for losses exceeding VaR. CVaR provides an adequate picture of risks reflected in extreme tails. This is a very important property if the extreme tail losses are correctly estimated. CVaR may have a relatively poor out-of-sample performance compared with VaR if tails are not modelled correctly. In this case, mixed CVaR can be a good alternative that gives different weights for different parts of the distribution (rather than penalizing only extreme tail losses).

CVaR Deviation is a strong competitor to the Standard Deviation. For instance, in finance, a CVaR deviation can be used in the following concepts: the Sharpe ratio, portfolio beta, one-fund theorem (i.e., optimal portfolio is a mixture of a risk-free asset and a master fund), market equilibrium with one or multiple deviation measures, and so on.