Asset Allocation with CER

Carlo Favero

2013-14

The Constant Expected Returns Model

- The CER model assumes that an asset's return over time is normally distributed with a constant mean and constant variance.
- The model allows for the returns on different assets to be contemporaneously correlated with correlations constant over time.
- Returns are independent over time both across assets and within the same asset.

$$
\begin{aligned}
r_{i,t} &= \mu_i + \sigma_i \epsilon_{it} \\
\epsilon_{it} &\sim NID\left(0,1\right) \\
cov\left(\epsilon_{it}, \epsilon_{js}\right) &= \begin{cases} \sigma_{ij} & t = s \\ 0 & t \neq s \end{cases}
\end{aligned}
$$

The Constant Expected Returns Model

- The CAPM provides a good measure of risk and thus a good explanation for why some stocks earn higher average returns than others according to the simple model

$$\boldsymbol{\mu} - r^f \mathbf{e} = \boldsymbol{\beta} \odot [(\mu_M - r^f)\mathbf{e}],$$

- Excess returns are close to unpredictable; any predictability is a statistical artifact or cannot be exploited after transaction costs are imputed to actual trades based on such alleged predictability, i.e., whatever is our information set $\mathcal{I}_t$, $E[\mathbf{r}_{t+1} - r^f \mathbf{e} | \mathcal{I}_t] = E[\mathbf{r}_{t+1} - r^f \mathbf{e}] = \boldsymbol{\mu} - r^f \mathbf{e}$, there is nothing to be learnt from $\mathcal{I}_t$ for practical purposes;
- Volatility and covariances are approximately constant over time, i.e., $\Sigma_t \equiv Var[\mathbf{r}_{t+1} - r^f \mathbf{e} | \mathcal{I}_t] = Var[\mathbf{r}_{t+1} - r^f \mathbf{e}] = \Sigma$.
- asset prices behave as a (log) random walk with drift

Insert Clicker 1 here

Regression Model Representation

$$\mathbf{y}^{+} = \mathbf{X}^{+}\delta^{+} + \mathbf{u}^{+}, \tag{1}$$

where $\mathbf{y}^{+}$ is a $(NT \times 1)$ vector, $\mathbf{X}^{+}$ is a $NT \times \sum_{i=1}^{N} K_i$ matrix (K_i is the number of regressors available at each point in time), δ^{+} is a $\sum_{i=1}^{N} K_i \times 1$ vector of unknown parameters, and $\mathbf{u}^{+}$ is a $(NT \times 1)$ vector of residuals:

$$\mathbf{y}^{+} = \begin{pmatrix} \mathbf{y}_1 \\ \vdots \\ \vdots \\ \mathbf{y}_N \end{pmatrix}, \quad \mathbf{X}^{+} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_2 & \cdots & \vdots \\ \mathbf{0} & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{X}_N \end{pmatrix},$$

$$\delta^{+} = \begin{pmatrix} \delta_1 \\ \vdots \\ \vdots \end{pmatrix}, \quad \mathbf{u}^{+} = \begin{pmatrix} \mathbf{u}_1 \\ \vdots \\ \vdots \end{pmatrix}.$$

Regression Model Representation

Each equation can be represented as follows:

$$\mathbf{y}_i = \mathbf{X}_i \delta_i + \mathbf{u}_i,$$

$i = 1, 2, ..., N$. (**??**) is then easy to use to derive inferences on means, variance and covariances:

$$\begin{aligned} E[\mathbf{y}^{+}] &= \mathbf{X}^{+} \delta^{+} \\ Var[\mathbf{y}^{+}] &= Var[\mathbf{u}^{+}] \end{aligned}$$

Regression Model Representation

If $\mathbf{u}_i$ is assumed to have standard white noise properties, then the following properties hold for $\mathbf{u}^+$:

$$
\begin{aligned}
E\left(\mathbf{u}^+\right) &= \mathbf{0}_{NT} \\
E[\mathbf{u}^+(\mathbf{u}^+)'] &= \begin{pmatrix} E\left(\mathbf{u}_1\mathbf{u}_1'\right) & E\left(\mathbf{u}_1\mathbf{u}_2'\right) & \cdots & E\left(\mathbf{u}_1\mathbf{u}_N'\right) \\ E\left(\mathbf{u}_2\mathbf{u}_1'\right) & E\left(\mathbf{u}_2\mathbf{u}_2'\right) & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ E\left(\mathbf{u}_N\mathbf{u}_1'\right) & \cdots & \cdots & E\left(\mathbf{u}_N\mathbf{u}_N'\right) \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{11}\mathbf{I}_T & \sigma_{12}\mathbf{I}_T & \cdots & \sigma_{1N}\mathbf{I}_T \\ \sigma_{21}\mathbf{I}_T & \sigma_{22}\mathbf{I}_T & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ \sigma_{N1}\mathbf{I}_T & \cdots & \cdots & \sigma_{NN}\mathbf{I}_T \end{pmatrix} = {}^\circ \otimes \mathbf{I}_T.
\end{aligned}
$$

where each block of the covariance matrix $E[\mathbf{u}^+(\mathbf{u}^+)']$ is $T \times T$ by construction. Here $\otimes$ denotes a standard Kronecker product. $^\circ$ is non-singular covariance matrix.

Regression Model Representation

The simplest case of the CER also assumes that all residuals are both contemporaneously and serially *un*correlated, with diagonal covariance matrix $\sum_d \equiv diag\{\sigma_{11}, \sigma_{22}, ..., \sigma_{NN}\}$. Then, because the diagonal structure of $^\circ{}_d$, classical OLS equation by equation can be applied. Consider for example the observations on the ith return:

$$\mathbf{y}_i = \mathbf{e}_T \delta_i + \mathbf{u}_i,$$

where

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad \mathbf{X}_i = \mathbf{e}_T = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

The OLS estimates of the relevant parameters are then simply

$$\hat{\delta}_i = \frac{1}{T} \sum_{t=1}^{T} r_{it} = \bar{r}_i \quad \hat{\sigma}_{11} = \hat{\sigma}_1^2 = \frac{1}{T} \sum_{t=1}^{T} (r_{it} - \bar{r}_i)^2,$$

which are sample mean and sample variance.

Insert clicker 2 here

Insert Clicker 2

A Static Asset Allocation Problem with Constant Expected Returns

Let's denote with $\mathbf{r}$ the random vector of linear total returns from time t to time T from a given menu of N risky assets for interval $[t, T]$,
$\mathbf{r} \sim \mathcal{D}(\boldsymbol{\mu}, \Sigma)$
Given a degree of risk aversion λ, a standard *mean-variance* description of this allocation problem is the following:

$$\max_{\mathbf{w}} \left(1 - \mathbf{w}'\mathbf{e}\right) r^f + \mathbf{w}'\boldsymbol{\mu} - \frac{1}{2}\lambda(\mathbf{w}'\Sigma\mathbf{w})$$

where $E[\mathbf{r}] = \left(1 - \mathbf{w}'\mathbf{e}\right) r^f + \mathbf{w}'\boldsymbol{\mu} = r^f + \mathbf{w}'\left(\boldsymbol{\mu} - r^f\mathbf{e}\right)$ and $Var[\mathbf{r}] = \mathbf{w}'\Sigma\mathbf{w}$.

A Static Asset Allocation Problem with Constant Expected Returns

first-order conditions (FOCs) are necessary and sufficient and define the following system of N linear equations in N unknowns, the portfolio weights $\mathbf{w} \in \mathcal{R}^N$:

$$\left(\boldsymbol{\mu} - r^f \mathbf{e}\right) - \lambda \Sigma \mathbf{w} = \mathbf{0}.$$

Solving the FOCs yields:

$$\mathbf{w} = \frac{1}{\lambda} \Sigma^{-1} \left(\boldsymbol{\mu} - r^f \mathbf{e}\right),$$

Consider now the special case in which $\mathbf{w}'\mathbf{e} = 1$, that is no investment in the riskfree bond is allowed. The optimal portfolio in this case is the famous *tangency portfolio*:

$$\mathbf{e}'\mathbf{w} = \frac{1}{\lambda} \mathbf{e}' \Sigma^{-1} \left(\boldsymbol{\mu} - r^f \mathbf{e}\right) = 1 \Longrightarrow \lambda = \mathbf{e}' \Sigma^{-1} \left(\boldsymbol{\mu} - r^f \mathbf{e}\right)$$

$$\mathbf{w}^T = \frac{\Sigma^{-1} \left(\boldsymbol{\mu} - r^f \mathbf{e}\right)}{\mathbf{e}' \Sigma^{-1} \left(\boldsymbol{\mu} - r^f \mathbf{e}\right)},$$

A Static Asset Allocation Problem with Constant Expected Returns

(1) The weights in the tangency portfolio do not depend on the risk aversion parameter λ.
(2) Given that the optimal risky portfolio is uniquely determined, the tangency portfolio must then coincide with the market portfolio. Agents maximize their utility by taking a linear combination of the market portfolio and and the risk-free securities. Note that in this case we can express the return on any portfolio in the following way:

$$
\begin{aligned}
r^p &= (1-\beta)\, r^f + \beta r^M \\
r^p - r^f &= \beta \left(r^p - r^M \right)
\end{aligned}
$$

and the CAPM holds.

A Static Asset Allocation Problem with Constant Expected Returns

(3) Efficient portfolios are those with the highest expected return for a given level of risk.If we summarize the expected return-risk (mean-variance) properties of the feasible portfolios in a plot with portfolio expected return, $\boldsymbol{\mu}^P$, on the vertical axis and portfolio standard-deviation, σ^P, on the horizontal axis, then all efficient portfolios can be represented as points in the space $\left(\sigma^P, \boldsymbol{\mu}^P\right)$ and the efficient frontier is the line that connects all these points. given the properties of the tangency portfolios, weights for all portfolios on the efficient frontier are obtained by inputing different values for the risk-free rate in the expression for optimal weights
(4) the CER implies that the tangency portfolios and the efficient frontier do not depend on the horizon at which returns are defined.

Insert Clicker 3 here

Insert Clicker 3 here

What Happens in Practice ?

- First, estimates/forecasts of μ and Σ are not constant over time: think of an investor who as avalaible a sample from time t_0 to time t_T to decide optimal asset allcation over period the $T+1...T+k$. The estimates of μ and Σ based on the sample t_0 ...t_T are usually very different from those based on the sample $T+1...T+k$ and optimal asset allocation ex-ante does not coincide with the optimal asset allocation ex-post.
- Moreover, the CER approach to portfolio allocation can lead to dramatic swings in optimal portfolio weights for small changes in investment views and conditions, as given by the estimates/forecasts of μ and Σ. There is a simple reason for these common findings: too much sampling error in the estimation of the vector of expected returns and, due to this, an asset allocation which is idiosyncratic to the specific estimation sample.

Solutions

The are several solutions to the above problem

- explicitly allow for sample uncertainty
- stabilize weights by combining a portfolio based on views with a market portfolio
- Abandon the CER andwork with more complicated predictive models

The resampled optimal mean-variance portfolio

Implement *bootstrap methods* to derive the optimal portfolio allocation. Consider the estimation of a simple multivariate model, in which the only regressor is a constant for the returns r_t^i on N assets, $i = 1, 2, ..., N$:

$$
\begin{aligned}
r_{1t} &= \hat{\mu}_1 + \hat{u}_{1t} \\
r_{2t} &= \hat{\mu}_2 + \hat{u}_{2t} \\
&\dots \\
r_{Nt} &= \hat{\mu}_N + \hat{u}_{Nt}
\end{aligned}
$$

$$
\begin{bmatrix} \hat{u}_{1t} & \hat{u}_{2t} & \dots & \hat{u}_{Nt} \end{bmatrix}' \sim \mathcal{N}\left(\mathbf{0}, \hat{\Sigma}\right).
$$

Implement the following algorithm.
1) Collect of the residuals from estimation in the following $T \times N$ matrix:

$$\hat{U} \equiv \begin{bmatrix} \hat{u}_{11} & \hat{u}_{21} & \ldots & \hat{u}_{N1} \\ \hat{u}_{12} & \hat{u}_{22} & \ldots & \hat{u}_{N2} \\ \vdots & \ldots & \ddots & \vdots \\ \hat{u}_{1T} & \hat{u}_{2T} & \ldots & \hat{u}_{NT} \end{bmatrix}.$$

At this point, draw a new sample of size T of residuals by extracting randomly T rows from $\hat{U}$.
2) Given these new, re-sampled residuals collected in a vector $\hat{u}_t^1$ $(t = 1, 2, \ldots, T)$ and the estimates $\hat{\boldsymbol{\mu}}$, we proceed to generate a new artificial sample of returns using

$$\mathbf{r}_t^1 = \hat{\boldsymbol{\mu}} + \hat{u}_t^1,$$

where the subscript "1" alludes to the fact that this represents the first iteration of the algorithm. At this point, a new OLS estimation of the model is performed on this artificial data, obtaining as an outcome a pair of new, bootstrapped estimates, $\hat{\mu}^1$ and $\hat{\Sigma}^1$ and, using the classical formula, $\hat{w}^1$.

3) Iterate the algorithm B times, where B is in general a large number (let's say 5,000 or 10,000 times), using the fact that at the bth iteration one simply draws a new sample of size T of residuals by extracting randomly T rows from $\hat{\mathbf{U}}$, generate a new artificial sample of returns using

$$\mathbf{r}_t^b = \hat{\boldsymbol{\mu}} + \hat{u}_t^b,$$

perform OLS estimation of $\hat{\boldsymbol{\mu}}^b$ and $\hat{\boldsymbol{\Lambda}}^b$ to obtain $\mathbf{w}^b$, for $b = 1, 2, ..., B$. This total number B of replications of this procedure will generate B optimal portfolio allocations $\{\hat{w}^b\}_{b=1}^B$.
The desidered vector of re-sampled, optimized portfolio weights may be represented by the average, across the B bootstraps, of the weights in $\{\hat{w}^b\}_{b=1}^B$:

$$\breve{w}^{boot} = \frac{1}{B} \sum_{b=1}^{B} \hat{w}^b.$$

Insert Clicker 4 here

Insert Clicker 4 here

Black and Litterman's approach

Given the knowledge of the market portfolio weights $\mathbf{w}_{mkt}$ and some estimates of the variance-covariance matrix of returns, we can use the optimal portfolio allocation condition to derive the expected returns consistent with the market capitalization:

$$\mathbf{w}_{mkt} = \frac{1}{\lambda}\Sigma^{-1}\left(\boldsymbol{\mu} - r^f \mathbf{e}\right) \Longrightarrow \boldsymbol{\mu}_{mkt} = \lambda \Sigma \mathbf{w}_{mkt} + r^f \mathbf{e}.$$

Assume now that the portfolio manager holds some (normally distributed, for simplicity) views on a subset of size $Q \leq N$ of the N expected returns included in the market portfolio:

$$P\mu_r \sim N_Q\left(v, \Gamma\right),$$

where $\mathcal{N}_Q$ denotes a q-variate multivariate normal distribution, $\boldsymbol{\mu}_r$ is a vector of N expected returns, and $\mathbf{P}$ is an appropriate $Q \times N$ selection matrix that selects the subset of returns on which there are subjective views expressed by the investor.

Black and Litterman's approach

The views are expressed as a vector of mean expected returns **V** and a diagonal variance-covariance matrix Γ, expressing the confidence on the views. Such subjective views have to be balanced against the distribution of returns implied by the market portfolio:

$$\boldsymbol{\mu}_r \sim N\left(\boldsymbol{\mu}_{mkt}, \tau\Sigma\right),$$

where τ is a scalar smaller than one (and conventionally set to 1/3 by Black and Littermann and most of the subsequent literature) to filter out of the estimated covariance matrix of returns the impact of their random variation (i.e., to take into account the effect of noise in small samples).
Black and Littermann's approach aims then at generating a value for the expected return vector $\boldsymbol{\mu}_{BL}$ by optimally combining the distribution of returns implied in the market capitalization and the subjective views of the portfolio manager.

Black and Litterman's approach

This is obtained by solving the following optimization problem:

$$\boldsymbol{\mu}_{BL} = \arg\min_{\boldsymbol{\mu}} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_{mkt}\right)' \left(\tau\Sigma\right)^{-1} \left(\boldsymbol{\mu} - \boldsymbol{\mu}_{mkt}\right) + \left(P\mu - v\right)' \Gamma^{-1} \left(P\mu - v\right).$$

This is a weighted least squares problems, where the weights depend on covariance matrix.
When the diagonal elements of Γ all approach zero, that is, when there is infinite confidence in the subjective views by the investor, the problem becomes a constrained least squares problem where the relevant constraint is $P\mu_{BL} = v$. On the other hand, when Γ has diagonal elements diverging to infinity (no confidence in the views), the solution to the problem is simply $\boldsymbol{\mu}_{BL} = \boldsymbol{\mu}_{mkt}$.

Black and Litterman's approach

The first order conditions for the solution of the problem can be written as follows:

$$2\left(\tau\Sigma\right)^{-1}\left(\boldsymbol{\mu}_{BL}-\boldsymbol{\mu}_{mkt}\right)+2P'\Gamma^{-1}\left(P\hat{\mu}_{BL}-v\right)=\mathbf{0}$$

from which we can derive:

$$\begin{aligned}\hat{\mu}_{BL} &= \left(\left(\tau\Sigma\right)^{-1}+P'\Gamma^{-1}P\right)^{-1}\left(\left(\tau\Sigma\right)^{-1}\boldsymbol{\mu}_{mkt}+P'\Gamma^{-1}v\right)\\ &= \left[\underbrace{\left(\left(\tau\Sigma\right)^{-1}+P'\Gamma^{-1}P\right)^{-1}\left(\tau\Sigma\right)^{-1}}_{=\Psi}\right]\boldsymbol{\mu}_{mkt}+\left[\underbrace{\left(\left(\tau\Sigma\right)^{-1}+P'\Gamma^{-1}P\right)^{-1}}_{=I_N-\Psi}\right.\end{aligned}$$

This expression emphasizes that $\boldsymbol{\mu}_{BL}$ is obtained by optimally combining market views ($\boldsymbol{\mu}_{mkt}$) with the investor's views ($\mathbf{v}$), through a rather complex weighting matrices given by Ψ and $I_N-\Psi$, respectively.

Black and Litterman's approach

Also note that μ_{BL} can be equivalently written as:

$$\hat{\mu}_{BL} = \mu_{mkt} + K\left(v - P\mu_{mkt}\right) \qquad K = \left(\tau\Sigma\right)P'\left(P\tau\Sigma P' + \Gamma\right)^{-1}.$$

At this point, given μ_{BL} the optimal BL portfolio weights are obtained by the usual formula:

$$\hat{w}_{BL} = \frac{1}{\lambda}\hat{\Sigma}^{-1}\left(\hat{\mu}_{BL} - r^f\mathbf{e}\right) \text{ or } \hat{w}_{BL}^T = \frac{\Sigma^{-1}\left(\hat{\mu}_{BL} - r^f\mathbf{e}\right)}{e'\Sigma^{-1}\left(\hat{\mu}_{BL} - r^f\mathbf{e}\right)}.$$

Similarly to how optimal portfolio weights in the tangency portfolio are computed, the BL efficient frontier can also be computed using $\hat{\mu}_{BL}$ and $\hat{\Sigma}$ as inputs.

Insert Clicker 5 here

Insert Clicker 5 here