Clarifications on the problem of interest and the Bayesian solution.

Bertille Antoine

November 2, 2005

Assumption. *The (N,1)-vectors R_t of excess returns at date t, $t = 1, \cdots, T$ are independent and identically normally distributed with mean μ and variance Σ.*

The investor chooses the vector of weights θ so as to maximize the mean-variance objective function,

$$U(\theta) = E(R_{Pt}) - \frac{\eta}{2} Var(R_{Pt}) = \theta'\mu - \frac{\eta}{2}\theta'\Sigma\theta$$

where $R_{Pt} = \theta' R_t$ denotes the excess portfolio return at time t.
When the first two moments, μ and Σ, are known the optimal vector of weights is:

$$\theta^* = \frac{1}{\eta}\Sigma^{-1}\mu$$

In practice, μ and Σ are unknown.

1 Frequentist approach

Suppose the investor has T periods of observed returns data $\Phi_T = \{R_1, \cdots, R_T\}$ and wants to build the portfolio for period $(T+1)$.
Under the assumption about iid normal distribution for the returns, the MLE estimates for the mean and the variance are:

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})'$$

Portfolio can be ranked by using the following risk function,

$$U(\theta^*) - E(U(\hat{\theta}))$$

or equivalently using the expected utility $E(U(\hat{\theta}))$ where E is taken under the true distribution of returns across repeated random samples of Φ_T.

$\Rightarrow$ This is called the frequentist approach.

2 Standard Bayesian approach

When using a bayesian approach, the vector of weights is obtained by maximizing the expected utility based on the predictive distribution of R_{T+1} conditional on the historical returns (sample) Φ_T.
Note that by definition of the bayesian criterion for optimality, the rule based on the true parameters of the model (optimal unfeasible rule θ^*) is suboptimal. Hence to compare rules obtained in a classical (frequentist) framework with bayesian rules obtained with different priors, the risk function is a useful criterion.
Standard diffuse prior used to represent "noninformative" beliefs about the parameters of a multivariate normal distribution is,

$$p(\mu, \Sigma) \propto |\Sigma|^{-(N+1)/2}$$

(see Zellner and Chetty (65), Klein and Bawa (76) or Stambaugh (97) among others.)
Now use the likelihood unction along with the sample data Φ_T to form updated beliefs about μ and Σ represented by the posterior density,

$$p(\mu, \Sigma|\Phi_T) \propto p(\mu, \Sigma) \times p(R_1, \cdots, R_T|\mu, \Sigma)$$

To get the conditional density $p(R_{T+1}|\Phi_T$ known as the bayesian "predictive" density, the joint density is integrated with respect to μ and Σ,

$$\begin{aligned} p(R_{T+1}|\Phi_T) &= \int_\mu \int_\Sigma p(R_{T+1}, \mu, \Sigma|\Phi_T) d\mu d\Sigma \\ &= \int_\mu \int_\Sigma p(R_{T+1}|\mu, \Sigma, \Phi_T) p(\mu, \Sigma|\Phi_T) d\mu d\Sigma \end{aligned}$$

Given the noninformative prior, the likelihood function and the sample Φ_T, the first two moments of the predictive pdf for R_{T+1} are:

$$\mu_{pred} = \hat{\mu} \quad \text{and} \quad \Sigma_{pred} = \frac{T+1}{T-N-2}\hat{\Sigma}$$

where $\hat{\mu}$ and $\hat{\Sigma}$ are the MLE estimates.
This yields to the following 2-fund portfolio rule:

$$\hat{\theta}_{Bayes} = \frac{1}{\eta}\frac{T-N-2}{T+1}\hat{\Sigma}^{-1}\hat{\mu}$$