1 Factor Models

The Markowitz mean-variance framework requires having access to many parameters: If there are n risky assets, with rates of return $r_i,\ i = 1, 2, \ldots, n$, then we must know all the n means ($\overline{r}_i$), n variances (σ_i^2) and $n(n-1)/2$ covariances (σ_{ij}) for a total of $2n + n(n-1)/2$ parameters. If for example $n = 100$ we would need 4750 parameters, and if $n = 1000$ we would need $501,500$ parameters! At best we could try to estimate these, but how? In fact, it is easy to see that trying to estimate the means, for example, to a workable level of accuracy is almost impossible using historical (e.g., past) data over time.[1] What happens is that the standard deviation of our estimate is too large (for example larger than the estimate itself), thus rendering the estimate worthless. One can bring the standard deviation down only by increasing the data to go back to (say) over a hundred years!

To see this: If we want to estimate expected rate of return over a typical 1-month period we could take n monthly data points $r(1), \ldots, r(n)$ denoting the rate of return over individual months in the past, and then average

$$\frac{1}{n}\sum_{j=1}^{n} r(j). \tag{1}$$

This estimate (assuming independent and identically distributed (iid) returns over months) has a mean and standard deviation given by

$\overline{r},\ \sigma/\sqrt{n}$ respectively, where $\overline{r}$ and σ denote the true mean and standard deviation over 1-month. If, for example, a stock's yearly expected rate of return is 16%, then the monthly such rate is $\overline{r} = 16/12 = 1.333\% = 0.0133$. Moreover, suppose that the monthly standard deviation is $\sigma = 0.05$ (e.g., a variance of 0.0025). Using (1) for our estimate with $n = 12$ yields a standard deviation for the estimate as $0.05/\sqrt{12} = 0.0144$, which is larger than the mean! (e.g., confidence intervals would be worthless.) Using $n = 60$ (5 years of data), helps to lower the standard deviation for the estimate to $0.05/\sqrt{60} = 0.00645$, which is a little below half of what it was, hardly a great improvement. To get the standard deviation down to about one-tenth of the mean would require that $0.05/\sqrt{n} = 0.00133$ or about $n = 1413$ corresponding to data for 117 years!

As such, it would seem imperative to derive simpler models that are not so data intensive, and which capture enough of reality to make them useful.

1.1 Factor models

A (linear) factor model assumes that the rate of return of an asset is given by

$$r = a + b_1 f_1 + \cdots + b_k f_k + e, \tag{2}$$

where the $f_j,\ j = 1, \ldots, k$, are $k \geq 1$ random variables (rvs) called *factors*, a and the b_j are constants and e is a mean zero "error" term rv assumed uncorrelated with the factors, $E(e) = 0$ and $E(ef_j) = E(e)E(f_j) = 0,\ j = 1, \ldots k$. The factors themselves are allowed to be correlated and are meant to simplify and reduce the amount of randomness required in an analysis of our assets. When $k = 1$ we call the model a single-factor model

[1] This is called the *historical blur* problem.

and when $k \geq 2$ we call it a multi-factor model. We use the notation $\overline{f}_j = E(f_j)$, $\sigma^2_{f_j} = var(f_j)$, $\sigma^2_{e_j} = var(e_j)$ throughout our discussion.

The factors are chosen by the modeler and depend upon the type of assets being considered. For example, for stocks, factors might be selected from among the stock market average, dividend yield of the $S\&P$ 500's Composite common stock, a measure of the risk of corporate bonds, interest rate variables, and various macroeconomic factors that capture the state of the economy such as employment rate, monthly growth rate in industrial production, monthly change in inflation rate, monthly growth rate in consumption and in disposable income. (See for example Ericsson and Karlsson, (2003) "Choosing factors in a multi-factor pricing model", Stockholm School of Economics, http://econpapers.hhs.se/paper/hhshastef/0524.html.)

When there are n risky assets indexed by $1 = 1, 2 \ldots, n$ (stocks say), then there are n equations for the model, one for each asset;

$$r_i = a_i + b_{1,i} f_1 + \cdots + b_{k,i} f_k + e_i, \quad i = 1, \ldots, n.$$

The factors are the same for each asset (that is what makes them correlated), but it is assumed that the error terms are uncorrelated between assets, $E(e_i e_j) = 0,\ i \neq j$.

If we form a portfolio of the n assets, defined by the weights $(\alpha_1, \ldots, \alpha_n)$, then in fact this portfolio is itself determined by a factor model, that is, the rate of return $r = \sum_{i=1}^n \alpha_i r_i$ of the portfolio satisfies (2) with

$$\begin{aligned} a &= \sum_{i=1}^n \alpha_i a_i \\ b_j &= \sum_{i=1}^n \alpha_i b_{j,i} \\ e &= \sum_{i=1}^n \alpha_i e_i. \end{aligned}$$

1.2 Single-factor models: CAPM revisited

The simplest case is when there is only one factor being considered;

$$r_i = a_i + b_i f + e_i.$$

All the mean-variance parameters can be computed directly in terms of the model parameters:

$$\begin{aligned} \overline{r}_i &= a_i + b_i \overline{f} \\ \sigma_i^2 &= b_i^2 \sigma_f^2 + \sigma_{e_i}^2 \\ \sigma_{ij} &= b_i b_j \sigma_f^2. \end{aligned}$$

Unlike the original Markowitz framework, where $2n + n(n-1)/2$ parameters must be estimated, here we need only estimate $3n + 2$ parameters; the $2n$ of a_i, b_i, one $\overline{f}$, one σ_f^2, and n of $\sigma_{e_i}^2$, a significant savings of work.

Note also that $Cov(r_i, f) = Cov(a_i + b_i f + e_i, f) = b_i \sigma_f^2$ and so we conclude that b_i is given by

$$b_i = Cov(r_i, f)/\sigma_f^2 \tag{3}$$

In general, once a factor is chosen, the a (intercept) and b (slope) constants are chosen in a least-squares sense: choose them so as to minimize the expected squared distance of plotted data points to the line $a + bf$:

Plot many independent pairs of realizations of (r, f) and try to find the "best" line through them. From (3) however, we see that only the intercept a needs to be chosen since the slope b is determined from the factor f and could be estimated.

Derivation of CAPM as a one-factor model

To see that even the one-factor model is not trivial we will here derive CAPM as a special case of such a model. (Consider n stocks as our risky assets.)

To this end, we use as our one factor, the market rate of return r_M, and for convenience express the linear model for each stock via

$$r_i - r_f = \alpha_i + \beta_i (r_M - r_f) + e_i,$$

where α_i and β_i are apriori unknown constants (effectively, we are taking as our factor $f = r_M - r_f$). Taking expected value yields

$$\overline{r}_i - r_f = \alpha_i + \beta_i (\overline{r}_M - r_f).$$

We can compute the slope β_i as in (3) yielding $\beta_i = \sigma_{i,M}/\sigma_M^2$ which indeed is the beta from the CAPM formula. This is CAPM exactly if $\alpha_i = 0$ which is what CAPM claims yielding the intercept $a = r_f$ (the risk-free interest rate).

Thus CAPM is a special case of a one-factor model.

1.3 Markowitz mean-variance or CAPM?

As we pointed out earlier, estimating parameters for the Markowitz mean-variance problem is not possible by using historical (past) data only. Estimates using both past and future prospects of the assets, however, can be carried out and thus all is not hopeless[2]. But even if this is done, since they are only estimates, the optimal solution obtained might be unreasonably far off from the actual solution. CAPM on the other hand uses an "equilibrium" approach to determine the market portfolio and in the end does not need to solve an optimization at all; it just uses the existing capitalization weights. Neither approach is completely satisfying: the first has estimation error while the second assumes apriori an equilibrium situation. Finally, both of our approaches assumes a 1-period time frame as opposed to a multi-period or continuous one. What should an investor do?

There are a variety of ways to combine and modify approaches, and there are even multi-period approaches that have been developed (we will study some later); but we should recognize that the CAPM and Markowitz mean-variance are very important first steps towards a serious quantitative analysis; financial engineering owes a lot to them.

[2] Such future information might include future plans of the company, including new products to be releaseed, etc., or information from the media