Session 11: Estimating inputs to the pricing models

Susan Thomas

http://www.igidr.ac.in /~s usan t

susant@mayin.org

IGIDR

Bombay

Recap

- The Markowitz framework: efficient portfolios in $E(r), \sigma$ space. Given $E(r)$, choose the portfolio with the lowest σ.
- One fund separation theorem: assuming rational investors, and low transaction, any investor will choose a portfolio allocation of part r_f and part r_M, the market portfolio.
- CAPM: Given $E(r)$, choose the portfolio with the lowest β because the returns are only for systematic risk that cannot be diversified away.

We seem to have come to (a) an allocation solution and (b) a pricing solution. How do we implement this?

Estimation of input variables

List of input variables

- $E(r)$
- Σ, which is a set of
  - individual asset σ_i^2, and
  - covariance terms $\sigma_{i,j}$.
- $E(r_M), \sigma_M$
- β of individual assets with benchmark indexes

Expected returns for securities

Expected returns

Sources for estimates of $E(r)$ are:

- Historical returns data for the individual asset
- β times the historical expected returns on the index

Questions about the data:

- Over what time interval?
- How long is good enough?
- What is the standard error of our estimate?

Problems with historical estimates of $E(r)$

- We only have samples.
- Each sample gives an estimate, with a standard error.
- At high frequencies (such as daily data), the mean estimate tends to be smaller than it's σ.
- With more data, the σ tends to go down only by $\sqrt{n}$.
- Thus, the uncertainty of the estimate reduces very slowly with more data.

Example of $E(r)$ estimation problems

- Nifty daily returns has an average of around 15 bps, and $\sigma = 1.5\%$
  The 95% confidence interval is $\pm 4.5\%$. The uncertainty is much larger than the mean itself.
- This was calculated using the last three years of data (~ 750 trading days). How many more data points would we need to drop the range of this interval?
- If I had 10 years of data, the σ would drop by 50 times to around 30 bps.
- If I had 100 years of data, σ would drop to around 9 bps.
- With 1000 years of data, σ would drop to 4 bps which is beginning to be manageable.

$E(r)$ estimates using β and $E(r_m)$

Use $r_i = r_f + \beta(r_m - r_f) + \epsilon_i$ to estimate $E(r_i) = r_f + \beta(E(r_m) - r_f)$. This is called the "market model" estimation.

- A superior path: $E(r_m)$ tends to be far **less noisy** than $E(r)$ for individual stocks.
- Further, β is a covariance estimate.
  Covariances tend to be more stable estimators than averages. Therefore, σ of $E(r)$ calculated using the market model would tend to be **less noisy** than using historical data.

Generalised modelling approach to estimating $E(r)$

- In general, say $r_i = r_{\text{model}} + \epsilon_i$.
- In the market model $r_{\text{model}} = \beta(r_m - r_f)$.
- There are other models, such as the Arbitrage Pricing Theory, that model $r_i = \sum w_i f_i + \epsilon_i$. Here, f_i are some observable factors. For instance, the first factor in typical APT models for stocks is r_m.

Time series dependance as a source of "noise"

- Typically, the high level of σ in financial variables is caused by time series dependance – ie, yesterday's returns affect today's returns to a certain amount. If we account for this dependance, then σ tends to be lower.
- The time series dependance can be modelled to remove some of the noise in σ and therefore, $E(r)$.
- In the generalised model for r_i, this might be modelled as f_i being the previous day's returns or ϵ, or σ.

Estimating the variance–covariance matrix for a securities portfolio

Estimates of variances and covariances

1. Historical estimates
2. Market model estimates
3. Realised volatility
4. Implied volatility

Historical estimates of variances and covariances

- The most widely used estimates are based on historical data.
- For example, daily variance of A can be calculated by
  1. Calculating a vector of daily returns for $A, B, \vec{r}_{A_d}, \vec{r}_{B_d}$
  2. Daily $\sigma^2_{A_d} : \frac{1}{T}\sum_{i=1}^{T}(r_{A_{d_i}} - \mu_{r_d})^2$
  3. Daily $\sigma^2_{AB_d} : \frac{1}{T}\sum_{i=1}^{T}(r_{A_{d_i}} - \mu_{A_d})(r_{B_{d_i}} - \mu_{B_d})$

Scaling to variances at lower frequencies

Monthly $\sigma^2_{A_m}$ can be calculated in two ways:

1. Calculate a monthly returns vector, $\vec{r}_{A_m}$. Then we use the above formula to calculate $\sigma^2_{A_m}$.
2. Scale from daily to monthly variance: $\sigma^2_{A_m} = 25 * \sigma^2_{A_d}$ Here we make two assumptions:
   (a) There are 25 trading days in a month.
   (b) The returns are *independant* and *normally distributed.*
   Under these assumptions, we can formulate a general rule for scaling up from a one-day estimate to an $n-$day estimate as $\sigma^2_{A_n} = n * \sigma^2_{A_d}$.

Question: Which method do we choose for σ_m?

Cautionary note

- We use less information when we use monthly returns. T data points of daily data means only $T/25$ points of monthly return.
- We need independance when using σ_d to calculate σ_m.
- However, most asset returns have some amount of *serial dependance*: there is some amount of information that flows from the previous day's returns to the current returns.
  1. If the serial dependance is *insignificant*, then scaling from σ_d to σ_m is effi cient.
  2. If not, scaling using σ_d will give us an *over-estimate* of the true σ_m.

Market model estimates

- The market model can be applied here as well. For example,

$$\begin{aligned} \sigma_A^2 &= \beta_A^2 \sigma_m^2 + \sigma_{\epsilon_A}^2 \\ \sigma_{AB} &= \beta_A \beta_B \sigma_m^2 + \rho_{\epsilon_A \epsilon_B} \sigma_{\epsilon_A} \sigma_{\epsilon_B} \end{aligned}$$

  have been shown to be more robust estimates of σ_{AB} than those estimated from historical data.
- Compared with historical data estimates of $E(r)$, these tend to be less noisy.

Realised volatility

- Use intra-day price data to calculate high–frequency returns for stock A.
  For example, we can calculate returns at every 5 minutes for Reliance.
- If we generate 5-minute returns, they can be used to calculate $\sigma^2_{A_{5mins}}$.
- $\sigma^2_{A_d} = 66 * \sigma^2_{A_{5mins}}$
  This is called the daily realised volatility of A.
- Note that this can also be used to get a good estimate of the volatility of any *single* day.

Pros and cons of using realised volatility

- Realised volatility has been found to be a better estimator of daily volatility than using daily data.
- It is found to be a better estimator of a single day's volatility than even the high-low estimator.
- The problems of serial correlation is higher for intra-day data than for daily data.
  The effects of serial correlation has to be removed before this estimator can be used.

Implied volatility

- If a liquid options market exists for a stock, then the option price can be used to estimate the annualised volatility of the stock – called the *implied volatility*.
- This is calculated using a recursive search algorithm (like the Newton–Raphson) on the Black-Scholes equation, with the option premium from the market as input.
- It is one of the best direct estimators of *future* volatility.
- However, the observed behaviour is that options with different maturities and different strikes give *different* estimates of the implied volatility of a stock. There is as yet no consisent framework to identify which is the best one.

Expected returns on the market

Estimating the expected market returns

- $E(r_M)$ is an important variable in pricing securities as well as portfolio optimisation.
- The most simple estimate used is the historical average of market index returns.
- Two problems with this estimate:
  1. It is a historical average – it need not be repeated into the future.
  2. The current historical average might be lower than the current risk free rate, r_f!
     This is not consistent with our notion that risky assets have to have a positive risk premium.

Equity premium

- The alternative: Use the equity premium added to the risk–free rate.
  1. The equity premium is the returns of the market over the risk-free rate.
  2. It is calculated as the average of the excess return of the market over a period of time.
  3. For every date, t, calculate $\tilde{e}_{M,t} = (R_{M,t} - r_{f,t})$.

     The equity premium $E_{eqprem} = \sum_{i=1}^{T} e_{M,t}/T$

  Then, $E(r_M) = E_{eqprem} + r_f$

Problems with the above approach

- There is no adjustment for a risk factor in the above equation.
- Merton, 1980 “On estimating the expected return on the market”, published in Journal of Financial Economics
- Several models incorporate various economic factors to accurately capture the risk factor in equity investment.
- This is ongoing research.

Estimating the stock β

The market model

- A stock β is estimated using a market model.
- For a particular j, the market model is:

$$r_{j,t} = \alpha_j + \beta_j(r_{M,t} - r_{f,t}) + \epsilon_t$$

  Here β_j can be estimated using a linear regression solution.
- It can also be worked out that $\beta_j = cov(r_j, r_M)/\sigma^2_{r_M}$ Once $cov(r_j, r_M)$ and $\sigma^2_{r_M}$ is estimated, β_j can be simply calculated.

Pros and cons

- β_j captures the leverage in a firm.
- If the firm's leverage changes, β_j changes.
- Empirically, it is observed that β_j is not a constant.
- Need to build models for estimating the time series behaviour of β_j.

Problems

Q1: Returns at different intervals

Take three stock prices returns for the last five years. Additionally, get the three-month interest rates as well as the INR-USD returns over the same five year period.

- For each of these series, calculate the historical μ and σ. Do this for daily data.
- Using the daily returns μ_d and σ_d, compute what the monthly μ_m and σ_m for each of these series ought to be. Next, calculate monthly data for the five price series and calculate the monthly μ_{md} and σ_{md} from the monthly returns data. How do they compare – are they the same? Why not?
- If you were to create monthly $E(r), \Sigma$ for a portfolio containing these five assets, which path would you choose – calculate monthly estimates from monthly data or from daily

Q2: Returns at different intervals

For the three stock prices returns used in the previous question,

- Calculate the variance-covariance matrix using historical data for the last one year, the last two years, the last three years.
  Are these numbers stable over time, or do they fluctuate significantly?
- Calculate the variance-covariance matrix using the market model over the same time periods.
  How do these estimates perform in comparison to those estimated using historical data? Which do you think are better?
  (NOTE: You can get estimates of β for most stocks either from the CMIE Prowess database, or if they are from the Nifty/Nifty Jr. set, from the NSE website.)