University of California, Los Angeles
Department of Statistics

Statistics C183/C283 **Instructor: Nicolas Christou**

Accuracy of historical betas

Forecasting betas with accuracy is important because they affect the inputs for the portfolio analysis problem. The variance covariance matrix is based on the value of beta for each stock. Different techniques have been proposed for better estimation of betas.

a. **Unadjusted betas:**
These are the betas obtained by the regression of the returns of the stock on the returns of the market.

$$R_{it} = \alpha + \beta R_{Mt} + \epsilon_{it}.$$

b. **Blume's technique (1975):**
The betas are adjusted as follows. Let us assume that we want to forecast betas for the period is 2020-24. Then we will need two five-year periods, 2010-14 and 2015-19. First, we calculate the betas for all stocks of interest for period 2010-14. We then calculate the betas for the same stocks for period 2015-19. And then we run the regression of the betas in 2015-19 on the betas in 2010-14 to get the equation

$$\hat{\beta}_{i2} = \hat{\gamma}_0 + \hat{\gamma}_1 \hat{\beta}_{i1}.$$

We therefore estimate the relationship between betas in two successive periods. Assume now that we want to forecast the beta of a stock in 2020-24. Then we find its beta in the period 2015-19 and substitute it in the equation above. For example, if the equation that connects the betas in the two historical periods is

$$\hat{\beta}_{iadjusted} = 0.3 + 0.6 \hat{\beta}_{i1}.$$

and the beta for a stock in period 2 is equal to 2, then the forecasted beta for this stock for 2020-24 will be 1.5.

c. **Vasicek's technique (1973):**
Let $\bar{\beta}_1$ be the average beta for the sample of stocks in the historical period (2015-19), and $\sigma^2_{\bar{\beta}_1}$ be the variance of the betas for the sample of these stocks. Also, let β_{i1} be the beta of stock i in the historical period and $\sigma^2_{\beta_{i1}}$ be the variance of β_{i1}. Then the adjusted forecasted β_i for 2020-24 will be a weighted average of β_i and $\bar{\beta}_1$ in 2015-19.

$$\beta_{iadjusted} = \frac{\sigma^2_{\beta_{i1}}}{\sigma^2_{\bar{\beta}_1} + \sigma^2_{\beta_{i1}}} \bar{\beta}_1 + \frac{\sigma^2_{\bar{\beta}_1}}{\sigma^2_{\bar{\beta}_1} + \sigma^2_{\beta_{i1}}} \beta_{i1}.$$

Note: $\sigma^2_{\beta_{i1}} = \frac{\sigma^2_{\epsilon_i}}{\sum_{t=1}^m (R_{mt} - \bar{R}_m)2}$.

Note: In the paper "A Note on Using Cross-Sectional Information in Bayesian Estimation of Security Betas," by Oldrich Vasicek (1973) the formula for adjusting the betas is given by $b'' = \frac{\frac{b'}{S_b'^2}+\frac{b}{S_b^2}}{\frac{1}{S_b'^2}+\frac{1}{S_b^2}}$, which can be expresses as the previous equation.

Comparison:
Elton, Gruber, and Urich (1978) compared the following 4 models in terms of their ability to predict the correlation matrix of n stocks:

a. The historical correlation matrix.

b. The correlation matrix based on the unadjusted betas.

c. The correlation matrix based on the Blume's technique.

d. The correlation matrix based on the Vasicek's technique.

They found that the historical matrix perform the poorest predictions. Therefore, the single-index model not only reduces the amount of data input, but also produces better estimates of the variance-covaiance matrix. However, the comparison of the three beta techniques was more difficult, but in some cases the Blume technique was the winner, while in some other the Vasicek's technique was the winner.

Example

```
#Stocks used:
^GSPC,AAPL,AMZN,BA,COST,CVX,DIS,ELY,FL,GILD,GS,HD,JCP,LMT,M,MMM,PG,S,
SBUX,SHOO,SMRT,TM,YUM,K,C,XOM,IBM,F,CAT,MCD,COKE.

#Three periods:
1.  2003-12-31 to 2008-12-31
2.  2009-12-31 to 2014-12-31
3.  2015-01-31 to 2019-12-31

a1 <- read.csv("stockData48.csv", sep=",", header=TRUE)

a2 <- read.csv("stockData49.csv", sep=",", header=TRUE)

a3 <- read.csv("stockData50.csv", sep=",", header=TRUE)


#Convert adjusted close prices into returns:
r1 <- (a1[-1,3:ncol(a1)]-a1[-nrow(a1),3:ncol(a1)])/a1[-nrow(a1),3:ncol(a1)]

r2 <- (a2[-1,3:ncol(a2)]-a2[-nrow(a2),3:ncol(a2)])/a2[-nrow(a2),3:ncol(a2)]

r3 <- (a3[-1,3:ncol(a3)]-a3[-nrow(a3),3:ncol(a3)])/a3[-nrow(a3),3:ncol(a3)]


#Compute the variance covariance matrix of the returns for each period:
covmat1 <- var(r1)
covmat2 <- var(r2)
covmat3 <- var(r3)


#Compute the betas in each period:
beta1 <- covmat1[1,-1] / covmat1[1,1]
beta2 <- covmat2[1,-1] / covmat2[1,1]
beta3 <- covmat3[1,-1] / covmat3[1,1]

#Here is the plot of the betas in period 2 against the betas in perod 1:
plot(beta1, beta2)

#Correlation between the betas in the two periods:
cor(beta1, beta2)
```

```
#Adjust betas using the Blume's technique:
q1 <- lm(beta2 ~ beta1)

beta3adj_blume <- q1$coef[1] + q1$coef[2]*beta2

#If we use beta2 as our forecasts for the betas in period 3 then our
PRESS (prediction sum of squares) is:
PRESS1 <- sum((beta2-beta3)^2) / 30

#If we use beta3adj_blume (the adusted betas) as our forecasts for the
betas in period 3 then our PRESS (prediction sum of squares) is:
PRESS2 <- sum((beta3adj_blume-beta3)^2) / 30




#====================================
#====================================
#Adjusting the betas using the Vasicek's technique:
#The adjustment is a weighted average of the average beta and
individual beta_i's.  The weights are the variance of the betas of the
30 stocks and the variance of beta_i_hat.
We need only one historical period.
We will use historical period 2 to forecast the betas in period 3.

#Initialize the vectors:
beta2 <- rep(0,30)
alpha2 <- rep(0,30)
sigma_e2 <- rep(0,30)
var_beta2 <- rep(0,30)

for(i in 1:30){
q <- lm(data=r2, formula=r2[,i+1] ~ r2[,1])
beta2[i] <- q$coefficients[2]
alpha2[i] <- q$coefficients[1]
sigma_e2[i] <- summary(q)$sigma^2
var_beta2[i] <- vcov(q)[2,2]
}


#Adjusting the betas using the Vasicek's technique:
beta3adj_vasicek <- var_beta2*mean(beta2)/(var(beta2)+var_beta2) +
var(beta2)*beta2/(var(beta2)+var_beta2)

#===============================================
#===============================================
```

```
#Now let's compare:
#Note:
beta3:  Actual betas in period 3.
beta2:  Betas in period 2 that can be used as forecasts for period 3.
beta3adj_blume:  Adjusted betas (Blume) that can be used as forecast
                 for period 3.
beta3adj_vasicek:  Adjusted betas (Vasicek) that can be used as forecast
                   for period 3.


cbind(beta3, beta2, beta3adj_blume, beta3adj_vasicek)

PRESS1 <- sum((beta2-beta3)^2) / 30     #unadusted


PRESS2 <- sum((beta3adj_blume-beta3)^2) / 30   #Blume


PRESS3 <- sum((beta3adj_vasicek-beta3)^2) / 30   #Vasicek
```