Basic Stats Finance, Beta
*draft 2009-08-30 r.r

This note is to discuss the calculation and interpretation of the financial metric called "Beta", using ideas from statistics and geometry. My interpretation of the definition of *Beta*, for an individual stock, shows it can be estimated by the slope of a regression line relating the percentage fluctuation of a reference financial instrument, such as the S&P index, to the percentage fluctuations of the stock of interest. That is, if the S&P index fluctuations are taken as a baseline, then the B*eta* (slope) shows how that individual stock moves (percentage change) relative to the S&P. So, for example, a *Beta* of zero equates to a slope of zero, and so the stock of interest doesn't move relative to the S&P, that is, their average fluctuations are the same. If Beta is, say, 1.0, then for every point up that the S&P moves, the stock also moves up 1 percentage point. This is simply a consequence of a straight line having a slope. Similarly, if Beta is -1, then the stock moves one unit opposite to the S&P.
Note: What I am doing here is simply plotting the S&P percentage fluctuations on the horizontal axis and corresponding Y stock fluctuations plotted on the Y axis. Fitting the best line relating 'SP' and 'Y' gives the standard regression line $\hat{Y}$ = a + Beta * SP, a standard regression problem. *Beta* is simply this slope.
The reader is referred to another tutorial on this site for basic derivations and general insights regarding regression, correlation, and their geometric connections. (see *BasicStatsBivariateRegression.pdf).*

■ Cut to the Chase

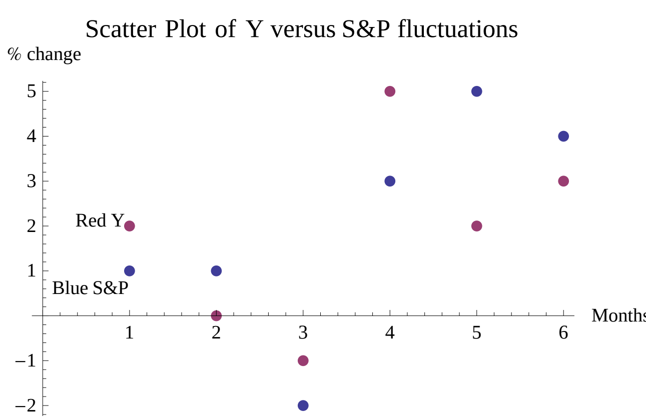
Let me use a simple example to illustrate how the *Beta* might be calculated and what its interpretation might be. Suppose I make up some S&P percentage fluctuations over a 6-month period. Similarly, over the same interval of time I make up data on the fluctuations of some hypothetical stock Y. Vectors of these fluctuations values might look like (the reader unfamiliar with vectors and their operations might wish to consult the tutorial *VectorOperationsQuickLook.pdf* on this site:

```
SP = {1., 1, -2, 3, 5, 4};(*S&P percentage fluctuations, over 6 months ( hypothetical) *)
```

```
Y = {2., 0, -1, 5, 2, 3};
(* stock Y percentage fluctuations over  six months (hypothetical)*)
```

■ Scatter Plot of Y versus SP

```
ListPlot[{SP, Y}, PlotStyle → PointSize[0.02],
 PlotLabel → "Scatter Plot of Y versus S&P fluctuations",
 AxesLabel → {"Months", "% change"}]
```



- **The answers are :**

Anticipating the results to be calculated below, we find that the slope of the regression line linking SP and Y is given by Beta = 0.594, that is, for every 1 percentage point rise of the S&P, the stock Y rises 0.594 percentage point. For a negative Beta value, such as -0.594, the stock would moves opposite to the S&P.
Correlation
r = correlation between SP and Y is : 0.703 (this is simply the Cosine of the angle between the SP and Y vectors, after centering)
That angle between (centered vectors of) SP and Y = 45 degrees
The square of the correlation, r^2 , is the proportion of the variability of the Y stock's fluctuations that are explained by the variability of the S&P fluctuations. That is:
Y variability = variability of the S&P index
Since $r^2 = .703^2 = 0.49$, this means that about 50% of the variability of the Y stock is explained by the variability of the S&P values.

- ## *End Cut to the Chase*

- ## Calculations and Comments

All statistical calculations benefit from 'centering' of variables and *correlation* and *beta* calculations are not affected by this operation. Centering a variable subtracts off its constant part. That constant part is equal to its mean or average value. (For those who have studied the tutorial, *BasicStatsBivariateRegression.pdf*, you will recognize that centering subtracts off the component of the vector projected onto the constant space, which is spanned by the equiangular vector **1**= {1,1,1, . . 1}, leaving only its variability in the complementary space. So, centered variables are those with their constant part subtracted out and consist of pure variability.

```
SPBar = Mean[SP](* the average fluctuation is '2 percentage points per month'*)
```

```
2.
```

```
YBar = Mean[Y](* the average fluctuation is '1.83 percentage pointsw per month'*)
```

```
1.83333
```

I will use lower case to indicate the centered variables and so, subtracting off the constant components of the **SP** and **Y** vectors leaves me:

```
sp = SP - SPBar
```

```
{-1., -1., -4., 1., 3., 2.}
```

```
y = Y - YBar
```

```
{0.166667, -1.83333, -2.83333, 3.16667, 0.166667, 1.16667}
```

Remember, calculations involving the slope of regression lines and the correlations between vectors, are unaffected by this transformation, but the calculations are easier.

- **Calculating the Line Linking sp and y (the regression line)**

Given that I want to see how well I can predict my stock fluctuations **y**, given the reference fluctuations **sp**, I calculate the best line possible by using least squares. Note, this line that I am going to calculate uses the centered variables.

Ok, back to the centered variable vectors. Geometrically this means that I find a Beta such that Beta * **sp** is as close as possible to **y**. (See the diagram below). Beta is some scalar that multiplies the vector **sp** by in order to get as close as possible to **y**.

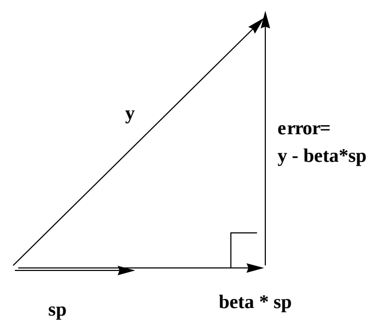
The diagram below is enough to determine Beta since it is that multiple of **sp** that gets as close as possible to the tip of **y**. This simply means that the *error vector* is perpendicular to **sp**. So, the dot product of sp with (y - Beta * sp) must be zero. For this to happen the dot product, which is proportional to the cosine between vectors must be zero.So, when the dot product is zero, the angle must be 90° since Cosine [90°] = 0, the signature of perpendicularity)

sp . (y - Beta * sp) == 0

After distributing the dot product through the parentheses I can write beta as:

Beta = sp.y / (sp.sp)

0.59375



- **The Correlation Angle between these vectors**

First I calculate the correlation coefficient between these two vectors which is;

```
correlation = Correlation[sp, y] // N
```

```
0.702901
```

and since correlation simply is the cosine of the angle between these two vectors in 6 - space, I can see what the angle actually is by finding the angle whose cosine is 0.702901 (I have to convert from radians to degrees).

```
angle = ArcCos[correlation] 180. / Pi
```

```
45.3398
```

So, the angle between these two vectors is 45 degrees.

- **The Connection between Correlation and Beta**

Since beta = y. sp / (sp.sp) and

$r = \text{correlation} = \text{y.sp} \Big/ \left(\sqrt{(\text{sp.sp})} * \sqrt{\text{y.y}}\right)$

beta = r | y | /| sp | = r σ_y / σ_{sp}

Summary

If I have interpreted the Beta definition correctly, then we can use simple regression techniques to determine stock movement and volatility. Since Beta figures prominently in the CAPM calculations for example, it might help to have a simple geometric interpretation and calculation method.