Historical Correlation Estimation

Cliff

January 10, 2021

Abstract

The document discussed the difference between weekly esimation and daily estimation on historical correlation with assumption that assets were traded continously.

1 Assumptions

For simplicity, we assumed the interest rates satisfy normal distribution and their covariance matrix were constant. Let $\mathbf{X}, \mathbf{Y}$ be two normal random variables, and X, Y be their sample path from 1 to T, which they can also be the sample paths of interest rates.

$$\mathbf{X} \sim \mathcal{N}(\mu, \sigma^2), \quad \mathbf{Y} \sim \mathcal{N}(\mu', \sigma'^2) \tag{1}$$

And assmued there were N weeks in the history, then $T = 7N$, and

$$\begin{aligned} X &= (x_1, x_2, x_3, ..., x_{7N}) \\ Y &= (y_1, y_2, y_3, ..., y_{7N}) \end{aligned} \tag{2}$$

Since the interest rates in weekend are not visible, the daily raturn series can be expressed as below,

$$\begin{aligned} X^w &= (x_2, x_3, x_4, x_5, x_6 + x_7 + x_8, x_9, x_{10}, x_{11}, x_{12}, ..., x_{7N-2}) \\ Y^w &= (y_2, y_3, y_4, y_5, y_6 + y_7 + y_8, y_9, y_{10}, y_{11}, y_{12}, ..., y_{7N-2}) \end{aligned} \tag{3}$$

And the weekly return was

$$\begin{aligned} X^w &= (\sum_{m=0}^{6} x_{1+m}, \sum_{m=0}^{6} x_{2+m}, \sum_{m=0}^{6} x_{3+m}, ..., \sum_{m=0}^{6} x_{7N-6+m}) \\ Y^w &= (\sum_{m=0}^{6} y_{1+m}, \sum_{m=0}^{6} y_{2+m}, \sum_{m=0}^{6} y_{3+m}, ..., \sum_{m=0}^{6} y_{7N-6+m}) \end{aligned} \tag{4}$$

2 Sample Covariance

Since we know,

$$Correl(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

and

$$Var(X) = Cov(X,X)$$

we could directly compared the covariance between these methods, and the sample covariance was

$$Cov(X,Y) = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(Y_i - \bar{Y})}{N-1} = \frac{\sum_{i=1}^{N} X_i Y_i}{N-1} - \frac{N\bar{X}\bar{Y}}{N-1} \tag{5}$$

3 Analysis

3.1 Daily Estimation

Recall the formula (3), it can be generalized as

$$\begin{aligned} X^d &= (x_{n+1}, x_{n+2}, x_{n+3}, x_{n+4}, \sum_{m=0}^{2} x_{n+m+5})_{n=1}^{N} \\ Y^d &= (y_{n+1}, y_{n+2}, y_{n+3}, y_{n+4}, \sum_{m=0}^{2} y_{n+m+5})_{n=1}^{N} \end{aligned} \tag{6}$$

$$Cov(X^d, Y^d) = \mathbb{E}\{X^d Y^d\} - \mathbb{E}\{X^d\}\mathbb{E}\{Y^d\} \tag{7}$$

and its covariance is

$$\begin{aligned} \hat{Cov}(X^d, Y^d) = &\frac{\sum_{n=1}^{N}\left(\sum_{i=1}^{4} x_{n+i}y_{n+i} + (\sum_{m=0}^{2} x_{n+m+5})(\sum_{m=0}^{2} y_{n+m+5})\right)}{5N-1} \\ &- \frac{5N}{5N-1}\left(\frac{\sum_{j=1}^{7N} x_j}{5N}\right)\left(\frac{\sum_{k=1}^{7N} y_k}{5N}\right) \\ = &\frac{\sum_{n=1}^{N}\left(\sum_{i=1}^{4} x_{n+i}y_{n+i} + \sum_{m=0}^{2}\sum_{p=0}^{2} x_{n+m+5}y_{n+p+5}\right)}{5N-1} \\ &- \frac{5N}{5N-1}\left(\frac{\sum_{j=1}^{7N} x_j}{5N}\right)\left(\frac{\sum_{k=1}^{7N} y_k}{5N}\right) \end{aligned} \tag{8}$$

Take an expectation,

$$
\begin{aligned}
\mathbb{E}\{\hat{Cov}(X^d, Y^d)\} &= \mathbb{E}\{\frac{\sum_{n=1}^{N}\left(\sum_{i=1}^{4} x_{n+i}y_{n+i} + \sum_{m=0}^{2}\sum_{p=0}^{2} x_{n+m+5}y_{n+p+5}\right)}{5N-1}\} \\
&= \frac{\mathbb{E}\{\sum_{n=1}^{N}\left(\sum_{i=1}^{4} x_{n+i}y_{n+i} + \sum_{m=0}^{2} x_{n+m+5}y_{n+m+5}\right)\}}{5N-1} \\
&+ \frac{\mathbb{E}\{\sum_{m=0}^{2}\sum_{p=0,p\neq m}^{2} x_{n+m+5}y_{n+p+5}\}}{5N-1} \\
&= \frac{7N}{5N-1} Cov(x,y)
\end{aligned}
\tag{9}
$$

Since

$$
\hat{Cov}(X,Y) = \frac{\sum_{i=1}^{7N}\mathbb{E}(x_i y_i)}{7N-1} - \frac{7N}{7N-1}\left(\frac{\sum_{i=1}^{7N}\mathbb{E}(x_i)}{7N}\right)\left(\frac{\sum_{i=1}^{7N}\mathbb{E}(y_i)}{7N}\right) \tag{10}
$$

3.2 Weekly Estimation

Recall the formula (4),

$$
\begin{aligned}
X^w &= (\sum_{m=0}^{6} x_{n+m})_{n=1}^{N} \\
Y^w &= (\sum_{m=0}^{6} y_{n+m})_{n=1}^{N}
\end{aligned}
\tag{11}
$$

and its covariance is

$$
\begin{aligned}
\hat{Cov}(X^w, Y^w) &= \frac{\sum_{7N-6}^{n=1}\left(\sum_{m=0}^{6} x_{m+n}\sum_{m=0}^{6} y_{m+n}\right)}{7N-7} \\
- \frac{7N-6}{7N-7}&\left(\frac{\sum_{n=1}^{7N-6}\sum_{i=n}^{n+6} x_i}{7N-6}\right)\left(\frac{\sum_{n=1}^{7N-6}\sum_{i=n}^{n+6} y_i}{7N-6}\right)
\end{aligned}
\tag{12}
$$

Take and expectation

$$
\begin{aligned}
\mathbb{E}\{\hat{Cov}(X^w, Y^w)\} &= \mathbb{E}\{\frac{\sum_{7N-6}^{n=1}\left(\sum_{m=0}^{6} x_{m+n}\sum_{m=0}^{6} y_{m+n}\right)}{7N-7}\} \\
&- \mathbb{E}\left(\frac{7N-6}{7N-7}\left(\frac{\sum_{n=1}^{7N-6}\sum_{i=n}^{n+6} x_i}{7N-6}\right)\left(\frac{\sum_{n=1}^{7N-6}\sum_{i=n}^{n+6} y_i}{7N-6}\right)\right) \\
&= \mathbb{E}\{\frac{\sum_{7N-6}^{n=1}\left(\sum_{m=0}^{6}\sum_{p=0}^{6} x_{m+n}y_{p+n}\right)}{7N-7}\} - \frac{36(7N-6)}{7N-7}\mathbb{E}\left(x\right)\mathbb{E}\left(y\right) \\
&= \mathbb{E}\{\frac{\sum_{7N-6}^{n=1}\left(\sum_{m=0}^{6} x_{m+n}y_{m+n} + \sum_{m=0}^{6}\sum_{p=0,p\neq m}^{6} x_{m+n}y_{p+n}\right)}{7N-7}\} \\
&- \frac{36(7N-6)}{7N-7}\mathbb{E}\left(x\right)\mathbb{E}\left(y\right) \\
&= \frac{\sum_{n=1}^{7N-6}\sum_{m=0}^{6}\mathbb{E}(xy) + 0}{7N-7} - \frac{6(7N-6)^2}{7N-7}\mathbb{E}\left(x\right)\mathbb{E}\left(y\right) \\
&= \frac{\sum_{n=1}^{7N-6}\sum_{m=0}^{6}(Cov(x,y) + \mathbb{E}(x)\mathbb{E}(y))}{7N-7} - \frac{36(7N-6)}{7N-7}\mathbb{E}\left(x\right)\mathbb{E}\left(y\right) \\
&= \frac{6(7N-6)}{7N-7}(Cov(x,y) + \mathbb{E}(x)\mathbb{E}(y)) - \frac{36(7N-6)}{7N-7}\mathbb{E}\left(x\right)\mathbb{E}\left(y\right)
\end{aligned}
\tag{13}
$$