ESSEC

Master in Finance

Advanced Master in Financial Engineering (MSTF)

FINM32227

Financial Risk Management

CLASS HANDOUTS

SESSION 2

Peng Xu

Volatility, Correlations and Copulas

Outline

– Volatility

– Correlations and Copulas

I. Volatility

- A variable's volatility σ is defined as the standard deviation of the return provided by the variable per unit of time when the return is expressed using continuous compounding
  - For option pricing, the unit of time is usually one year
  - For risk management, the unit of time is usually one day
- In general, $\sigma\sqrt{T}$ is equal to the standard deviation of

$$\ln\frac{S_T}{S_0}$$

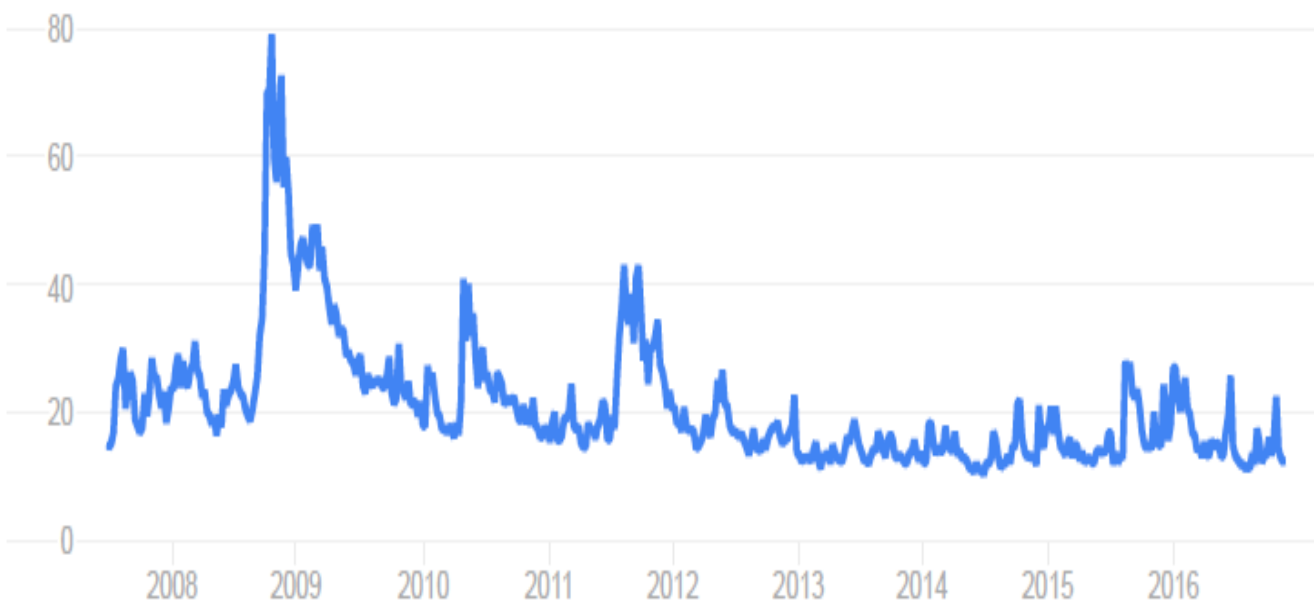
  where S_T is the value of the market variable at time T and S_0 is its value today.
  - If σ is per day, then T is measured in T days
  - If σ is per year, then T is measured in T years
- Normally days when markets are closed are ignored in volatility calculations, so the volatility per year is $\sqrt{252}$ times the daily volatility, i.e., $\sigma_{year} = \sqrt{252}\sigma_{day}$
- Risk managers often focus on the *variance rate*, which is defined as the square of the volatility.

  Note: The standard deviation of the return in time T increases with the square root of time, while the variance of this return increases linearly with time.

I. Volatility

- Volatility cannot be observed directly, but we can
  - imply volatilities from market prices, or
  - estimate volatility from historical data
- Implied volatilities are used extensively by traders.
  - E.g., calculating implied volatility from the Black-Scholes option pricing formula.
  - The CBOE publishes indices of implied volatility.

    e.g., SPX VIX is an index of the implied volatility of 30-day options on the S&P 500 calculated form a wide range of calls and puts



- Estimating Volatility from Historical data:
  - The value of the variable (e.g., stock price) is usually observed at fixed interval of time (e.g., day, week, or month)

    Define the time interval as τ

    Define $m+1$ as the number of observations

    Define S_i as the value of market variable at end of interval i $(i = 0, 1, \dots, m)$

    Define $u_i = \ln \frac{S_i}{S_{i-1}}$, which is the return during the ith interval
  - The estimate s of the standard deviation of u_i is given by

$$s = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} (u_i - \bar{u})^2}$$

    where $\bar{u}$ is the sample mean of u_i
  - Since the standard deviation of u_i is $\sigma\sqrt{\tau}$,

$$\hat{\sigma} = \frac{s}{\sqrt{\tau}}$$

    The standard error of this estimate is approximately $\frac{\hat{\sigma}}{\sqrt{2m}}$

– In practice, the volatility of asset prices is not constant, so it is important to monitor it on a daily basis.

- Define $\sigma_n$as the volatility of a market variable on day n, as estimated at the end of day $n-1$
- One approach to estimating σ_n is to use the most recent m days' data, i.e.,

$$\sigma_n^2 = \frac{1}{m-1}\sum_{i=1}^{m}(u_{n-i} - \bar{u})^2$$

- For risk management purpose, $u_i = \ln\frac{S_i}{S_{i-1}}$ is usually approximated by $u_i = \frac{S_i}{S_{i-1}} - 1 = \frac{S_i - S_{i-1}}{S_{i-1}}$, $\bar{u}$ is assumed to be 0, and $m-1$ is replaced by m. So the above formula is simplified to

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m}u_{n-i}^2$$

- Since our objective is to estimate σ_n, it makes sense to give more weight to recent data, e.g.,

$$\sigma_n^2 = \sum_{i=1}^{m}\alpha_i u_{n-i}^2$$

where $\alpha_i > \alpha_j$ when $i < j$, and $\sum_{i=1}^{m}\alpha_i = 1$

- We can also extend the idea by assuming that there is a long-run average variance rate and it should be given some weight, i.e.,

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2$$

and $\gamma + \sum_{i=1}^{m} \alpha_i = 1$. This is known as $\boldsymbol{ARCH(m)}$ model.

– If we assume $\gamma = 0$, and $\alpha_{i+1} = \lambda\alpha_i$ where $0 < \lambda < 1$, then the formula is simplified to

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2$$

This is the **exponentially weighted moving average (EWMA) model**.

Example:

Suppose that $\lambda = 0.9$, the volatility estimated for a market variable for day $n-1$ is 1% per day, and during day $n-1$ the market variable increased by 2%. What is your estimate of volatility for day n based on the EWMA model?

Attractions of EWMA:

- Relatively little data needs to be stored.

  We need only remember the current estimate of the variance rate and the most recent observation on the market variable

- Tracks volatility changes.

  λ governs how responsive the estimate of the daily volatility is to the most recent daily percentage change.

  RiskMetrics uses $\lambda = 0.94$ for daily volatility forecasting

– If we add a long-run average variance rate, V_L, to the EWMA model, i.e.,

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

where $\gamma + \alpha + \beta = 1$. This is **GARCH (1,1) model**.

- Example:

  Suppose that a GARCH (1,1) model is estimated from daily data as $\sigma_n^2 = 0.000002 + 0.13u_{n-1}^2 + 0.86\sigma_{n-1}^2$. What does this imply for long-run average variance rate? Suppose that the estimate of the volatility on day $n-1$ is 1.6% per day, and on day $n-1$ the market variable decreased by 1%. What is the estimate of the volatility on day n?

- If we continually substitute for σ_{n-i}^2, we get:

  $$\sigma_n^2 = (1 + \beta + \beta^2 + \cdots)\omega + \alpha(u_{n-1}^2 + \beta u_{n-2}^2 + \beta^2 u_{n-3}^2 + \cdots)$$

  β is the "decay rate". Similar to λ in EWMA model, it defines the relative importance of the u_i in determining the current variance rate.

  GARCH(1,1) model is the same as the EWMA model except that it also assigns some weight to the long-run average variance rate.

- GARCH(1,1) can be used to forecast future volatility

  $$\sigma_n^2 - V_L = \alpha(u_{n-1}^2 - V_L) + \beta(\sigma_{n-1}^2 - V_L)$$

  On day $n + t$ in the future, we have

  $$\sigma_{n+t}^2 - V_L = \alpha(u_{n+t-1}^2 - V_L) + \beta(\sigma_{n+t-1}^2 - V_L)$$

  then $\quad E(\sigma_{n+t}^2 - V_L) = (\alpha + \beta)E(\sigma_{n+t-1}^2 - V_L)$

  and $\quad E(\sigma_{n+t}^2 - V_L) = (\alpha + \beta)^t(\sigma_n^2 - V_L)$

  or $\quad E(\sigma_{n+t}^2) = V_L + (\alpha + \beta)^t(\sigma_n^2 - V_L)$

  Note: This equation forecasts volatility on day $n + t$ based on the information available at the end of day $n - 1$

When $\alpha + \beta < 1$ and $t \to \infty$, $E(\sigma_{n+t}^2) \to V_L$. This property is called *mean reversion*.

So for a stable GARCH (1,1) process, we require $\alpha + \beta < 1$. Otherwise the weight given to $V_L$is negative and the process is “mean fleeing” rather than “mean reverting”.

Example:

Suppose $\alpha + \beta = 0.9617$ and $V_L = 0.0000442$. Our estimate of the current variance rate per day is 0.00006. What is the expected variance rate in 10 days? How about 100 days?

- GARCH(1,1) can also be used to construct *volatility term structure,* which is the relationship between the volatility of options and their maturities

  Define $V(t) = E(\sigma_{n+t}^2)$, and let $\alpha + \beta = e^{-a}$, then

$$V(t) = V_L + e^{-at}(V(0) - V_L)$$

  Now assume t is continuous, then the average variance rate per day between today and time T is

$$\frac{1}{T}\int_0^T V(t)dt = V_L + \frac{1 - e^{-aT}}{aT}(V(0) - V_L)$$

If we define $\sigma(T)$ as the volatility per annum and assume there are 252 days per year, then

$$\sigma(T)^2 = 252\left\{V_L + \frac{1-e^{-aT}}{aT}\left(\frac{\sigma(0)^2}{252} - V_L\right)\right\}$$

Note: T is measured in days

When the current variance rate per day $V(0) = \frac{\sigma(0)^2}{252}$ is above the long-run average variance rate V_L, the GARCH(1,1) model estimates a downward-sloping volatility term structure, and vice versa.

Example:

Use the estimates from the previous example to construct the volatility term structure.

When $\sigma(0)$ changes by $\Delta\sigma(0), \sigma(T)$ changes by approximately

$$\frac{1-e^{-aT}}{aT}\frac{\sigma(0)}{\sigma(T)}\Delta\sigma(0)$$

Example:

Continue from the previous example. Suppose that there is an event that increases the volatility from 12.30% per year to 13.30 per year. Estimate by how much the event increases the volatilities used to price 30-day and 100-day options.

- The more general GARCH(p,q) :

$$\sigma_n^2 = \omega + \sum_{i=1}^{p} \alpha_i u_{n-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{n-j}^2$$

– Estimating Models:

The parameters of both EWMA and GARCH(1,1) model can be estimated using historical data with maximum likelihood estimation (MLE) method.

Example:

Suppose the observations $u_1, u_2, \cdots, u_m$ are normally distributed (conditional on the variance) with mean 0 and variance σ_i^2, where σ_i^2 can be modeled by both EWMA and GARCH(1,1) model.

The likelihood of the observation is

$$\prod_{i=1}^{m}\left[\frac{1}{\sqrt{2\pi\sigma_i^2}}\exp\left(\frac{-u_i^2}{2\sigma_i^2}\right)\right]$$

Using the MLE method, the best estimates of the parameters are the ones that maximize the likelihood of the observations.

This is equivalent to maximizing the logarithm of the likelihood:

$$\sum_{i=1}^{m}\left(-\ln\left(\sigma_i^2\right)-\frac{u_i^2}{\sigma_i^2}\right)$$

– Choosing between EWMA and GARCH(1,1):

In practice, variance rates do tend to be pulled back to a long-run average level, which is known as *mean reversion*.

GARCH (1,1) is theoretically more appealing since it incorporates mean reversion.

If the parameter ω is 0, the GARCH(1,1) reduces to EWMA. If the estimate of ω is negative, then GARCH(1,1) is not stable and should switch to EWMA

We can also test whether GARCH model is working well in explaining data by looking at the autocorrelation structure of $\frac{u_i^2}{\sigma_i^2}$.

– The correlation coefficient between two variables V_1 and V_2 is

$$\rho = \frac{COV(V_1, V_2)}{SD(V_1)SD(V_2)} = \frac{E(V_1, V_2) - E(V_1)E(V_2)}{SD(V_1)SD(V_2)}$$

– Two variables are statistically independent if knowledge about one of them does not affect the probability distribution for the other. I.e., $V_1$and V_2 are independent if $f(V_2|V_1 = x) = f(V_2)$, where $f(\cdot)$ denotes the probability density function.

– Independence is <u>Not</u> the same as zero correlation.

Correlation coefficient only measures linear dependence between two variables.

– Suppose that X_i and Y_i are the values of two variables X and Y at the end of day i. The returns on the variables on day i are

$$x_i = \frac{X_i - X_{i-1}}{X_{i-1}},\ y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}.$$

The covariance rate between X and Y on day n is

$$cov_n = E(x_n y_n) - E(x_n)E(y_n)$$

Risk managers assume the expected daily returns are zero, so

$$cov_n = E(x_n y_n)$$

– Using equal weights for the last m observations on x_i and y_i gives the estimate

$$cov_n = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}y_{n-i}$$

Similarly,

$$var_{x,n} = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}^2, \qquad var_{y,n} = \frac{1}{m}\sum_{i=1}^{m} y_{n-i}^2$$

The correlation estimate on day n is $\frac{cov_n}{var_{x,n}var_{y,n}}$

– EWMA model for covariance

$$cov_n = \lambda cov_{n-1} + (1-\lambda)x_{n-1}y_{n-1}$$

– GARCH model can also be used for updating covariance rate estimates and forecasting the future level of covariance rates. E.g., the GARCH(1,1) is

$$cov_n = \omega + \alpha x_{n-1}y_{n-1} + \beta cov_{n-1}$$

So the long-run average covariance rate is $\omega/(1-\alpha-\beta)$

– Once variance and covariance rates have been calculated for a set of N variables, an $N \times N$ variance-covariance matrix, $\boldsymbol{\Omega}$, can be constructed.

A variance-covariance matrix, $\mathbf{\Omega}$, is internally consistent if the positive semi-definite condition

$$w^T \mathbf{\Omega}\, w \geq 0$$

holds for all $N \times 1$ vectors w.

To ensure that a positive-semidefinite matrix is produced, variance and covariance should be calculated consistently.

- Multivariate Normal Distribution:
  - Suppose V_1 and V_2 are bivariate normal. Conditional on V_1, V_2 is normal with mean

$$\mu_2 + \rho\sigma_2 \frac{V_1 - \mu_1}{\sigma_1}$$

and standard deviation

$$\sigma_2\sqrt{1-\rho^2}$$

where μ_1 and μ_2 are unconditional mean of $V_1$and V_2, σ_1 and σ_2 are their unconditional standard deviations, and ρ is the correlation coefficient between $V_1$and V_2.

- When there are N variables, V_i $(i = 1, 2, .. N)$, in a multivariate normal distribution there are $N(N-1)/2$ correlations

  We can reduce the number of correlation parameters that have to be estimated to N with a one-factor model

  Suppose that $U_1, U_2, \cdots, U_N$ have standard normal distribution.

  In a one-factor model,

$$U_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

  where F and Z_i have a standard normal distribution, Z_i are uncorrelated with each other and uncorrelated with F, and a_i is a constant between -1 and $+1$.

  In this model, all the correlation between U_i and U_j arises from their dependence on the common factor F, and the correlation coefficient is $a_i a_j$.

  The m-factor model is

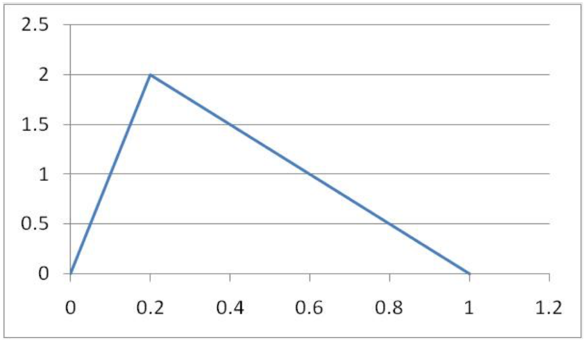
$$U_i = a_{i1}F_1 + a_{i2}F_2 + \cdots + a_{im}F_m + \sqrt{1 - a_{i1}^2 - a_{i2}^2 - \cdots - a_{im}^2} Z_i$$

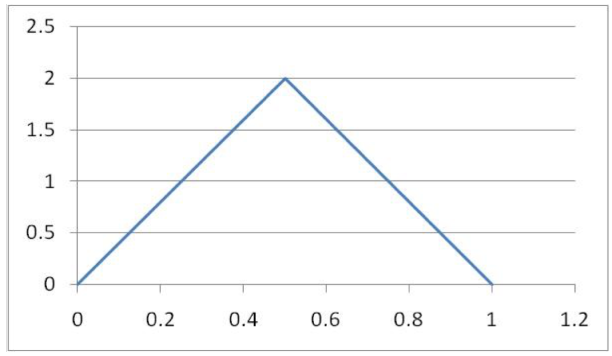
– Gaussian Copula Models:

Creating a correlation structure for variables that are not normally distributed

- Suppose we wish to define a correlation structure between two variable V_1 and V_2 that do not have normal distributions
- We transform the variable V_1 to a new variable U_1 that has a standard normal distribution on a “percentile-to-percentile” basis.
- We transform the variable V_2 to a new variable U_2 that has a standard normal distribution on a “percentile-to-percentile” basis.
- U_1 and U_2 are assumed to have a bivariate normal distribution with correlation ρ
- The correlation structure between V_1 and V_2 is defined by that between U_1 and U_2

Example:





V_1 V_2

V_1 mapping to U_1		
V_1	Percentile	U_1
0.2	20	-0.84
0.4	55	0.13
0.6	80	0.84
0.8	95	1.64

V_2 mapping to U_2		
V_2	Percentile	U_2
0.2	8	−1.41
0.4	32	−0.47
0.6	68	0.47
0.8	92	1.41

Assume correlation between U_1 and U_2 (copula correlation) is .5.

The Probability that V_1 and V_2 are both less than 0.2 is the probability that $U_1 < -0.84$ and $U_2 < -1.41$

When copula correlation is 0.5, this is $M(-0.84, -1.41, 0.5) = 0.043$, where M is the cumulative distribution function for the bivariate normal distribution

- The key property of a copula model is that it preserves the marginal distribution of V_1 and V_2 while defining a correlation structure between them.
- Instead of a bivariate normal distribution for U_1 and U_2 we can assume any other joint distribution

  One possibility is the bivariate Student t distribution, which has a higher tail correlation than the bivariate normal distribution.
- Copulas can be used to define a correlation structure between V_1, V_2,…V_n

  E.g., multivariate Gaussian copula:
  - We transform each variable V_i to a new variable U_i that has a standard normal distribution on a “percentile-to-percentile” basis.
  - The *U’s* are assumed to have a multivariate normal distribution

    In a factor copula model the correlation structure between the *U*’s is generated by assuming one or more factors.

- Application of the one-factor Gaussian copula to loan portfolio:

  Vasicek's model

  - Define T_i as the time when company i default
  - Define PD_i as the probability that company i will default by time T:

$$PD_i = Prob(T_i < T)$$

  - We map the time when company i default, T_i, to a new variable U_i and assume

$$U_i = a_i F + \sqrt{1 - a_i^2} Z_i$$

    where F and the Z_i have independent standard normal distributions

  - The mappings imply

$$Prob(U_i < U) = Prob(T_i < T)$$

    when

$$U = N^{-1}[PD_i]$$

  - From the one-factor model,

$$Prob(U_i < U|F) = Prob\left(Z_i < \frac{U - a_i F}{\sqrt{1 - a_i^2}}\middle| F\right) = N\left[\frac{U - a_i F}{\sqrt{1 - a_i^2}}\right]$$

$$Prob(T_i < T|F) = N\left[\frac{N^{-1}[\,PD_i] - a_i F}{\sqrt{1 - a_i^2}}\right]$$

Assuming PD_i and a_i are the same for all i, i.e., $PD_i = PD$ and $a_i = \sqrt{\rho}$

$$Prob(T_i < T|F) = N\left[\frac{N^{-1}[\,PD] - \sqrt{\rho}F}{\sqrt{1 - \rho}}\right]$$

This is the default rate conditional on F.

As F decreases, the default rate increases.

The probability of $F < N^{-1}(Y)$ is Y, so there is probability Y that the default rate will be greater than

$$N\left[\frac{N^{-1}[\,PD] - \sqrt{\rho}N^{-1}(Y)}{\sqrt{1 - \rho}}\right]$$

Or there is probability $X = 1 - Y$ that the default rate will be less than

$$N\left[\frac{N^{-1}[\,PD]\;+\sqrt{\rho}N^{-1}(X)}{\sqrt{1-\rho}}\right]$$

This is WCDR(T,X), the worst-case default rate for time horizon T and a confidence level X.

The VaR for this time horizon and confidence limit is

$$\mathrm{VaR(T,X) = L{\times}(1 - R){\times}WCDR(T,X)}$$

where *L* is the dollar size of the loan portfolio and *R* is recovery rate

Example:

Suppose that a bank has a total of $100 million of retail exposures. The one-year probability of default averages 2% and the recovery rate averages 60%. The copula correlation parameter is estimated as 0.1. What is the VaR with one year horizon and 99.9% confidence level?

- Estimating PD and ρ:

The MLE method can be used to estimate PD and ρ from historical data on default rates.

Define DR as the default rate and $G(DR)$ is the cumulative probability distribution function for DR, we have

$$DR = N\left[\frac{N^{-1}[\,PD] + \sqrt{\rho}N^{-1}(G(DR))}{\sqrt{1-\rho}}\right]$$

This implies:

$$G(DR) = N\left(\frac{\sqrt{1-\rho}N^{-1}(DR) - N^{-1}[\,PD]}{\sqrt{\rho}}\right)$$

Differentiating this, the probability density function for the default rate is

$$g(DR) = \sqrt{\frac{1-\rho}{\rho}}\,exp\left\{\frac{1}{2}\left[\left(N^{-1}(DR)\right)^2 - \left(\frac{\sqrt{1-\rho}N^{-1}(DR) - N^{-1}[\,PD]}{\sqrt{\rho}}\right)^2\right]\right\}$$

MLE: Use Solver in Excel to search for the values of PD and ρ that maximize of the log likelihood of the observations.