

Undergraduate Research Opportunity
Programme in Science

Value at Risk

Dai Bo

Supervisor: Dr. Arie Harel

Department of Mathematics
National University of Singapore
Academic year (2000/2001)

Summary

Value at Risk (VaR) is one of the most popular tools used to estimate exposure to market risks, and it measures the worst expected loss at a given confidence level. In this report, we explain the concept of VaR, and then describe in detail some methods of VaR computation. We then discuss some VaR tools that are particularly useful for risk management, including marginal VaR, incremental VaR and component VaR. The next consideration is the effect of time varying risk, which can be estimated by a moving average model or a GARCH process. Finally, we introduce some back testing methods to validate the use of VaR model.

All description, definitions, examples, results, proofs, tables, and remarks in this report are taken from the 2nd edition of the book of Philippe Jorion “Value at Risk” (Jorion 2001), unless otherwise indicated.

Table of contents

Cover page	I
Summary	II
Table of contents	III
Chapter 1 Motivation and Introduction	1
1.1 Motivation	1
1.2 Introduction	1
1.3 Overview of the report	2
Chapter 2 VaR computation	3
2.1 Definition of VaR	3
2.2 Measuring returns	3
2.3 Computation of VaR	4
2.4 VaR measurement over different parameters	9
2.5 Choice of parameters	10
Chapter 3 Portfolio VaR	12
3.1 Portfolio VaR for multiple assets	12
3.2 VaR reduction methods	14
3.3 VaR tools	15
Chapter 4 Modeling Time-varying Risk	20
4.1 The existence of time-varying risk	20
4.2 Moving average	21
4.3 GARCH estimation	23
4.4 RiskMetrics approach	27
Chapter 5 Back Testing VaR Models	29
5.1 Model back testing with exceptions	29
5.2 Model back testing with conditional coverage	32
Chapter 6 Conclusions	36
References	37

Chapter 1

Motivation and Introduction

1.1 Motivation

One may still remember the name Nicholas Leeson. He was an investment officer of Barings Bank London England and worked in the bank's Singapore office. Mr. Leeson lost 1.3 billion dollars because of risky derivative investments in the Japanese future market. This huge loss wiped out the firm's entire equity capital and resulted in the financial collapse of one of the world's largest banks.

This financial disaster revealed the absence of enforced risk management policies. In the later studies, it has been shown that under normal market conditions, the potential loss of Mr. Leeson's trading would exceed 835 million dollars 5 percent of the time. If these calculations had been in place, the parent company could provide some protection against rogue traders and other operational risks, and thus the bank disaster might be avoided.

Because Barings was viewed as a conservative bank, the bankruptcy served as a wakeup call for financial institutions all over the world. Thereafter, risk management has emerged and many methodologies were developed for the purpose of risk management.

1.2 Introduction

Value at Risk, abbreviated as VaR, was developed in 1993 in response to those famous financial disasters such as Barings's fall.

Initially, the purpose of VaR is to quantify financial risks by using standard statistical techniques. It measures the worst expected loss over a given horizon under normal market conditions at a given confidence level. For most users, VaR was simply a passive application of the risk quantification.

However, VaR is now being used to control and manage risk actively. By making use of VaR tools, institutions can decide how to allocate economic capital and how to trade off risk and return.

1.3 Overview of the Report

The body of this report consists of four parts, namely,

- VaR computation
- Portfolio VaR
- Modeling time-varying risk
- Back testing VaR models

The first part is about the VaR computation for a single asset. In this part, the definition of VaR is introduced and some basic methods of VaR computation are illustrated. Some discussions of choices of the parameters are also included.

The second part describes the computation of portfolio VaR, followed by some VaR tools that enable us to control the risk, such as marginal VaR and incremental VaR.

The third part discusses the effect of time-varying risk. In reality, the standard deviation of risk is not constant, thus some models like moving average and GARCH model are used to forecast the risk movement.

The last part is about back testing. In this part, back testing with exceptions and back testing with conditional coverage are used to validate the VaR model.

Chapter 2

VaR Computation for single asset

2.1 Definition of VaR

According to Philippe Jorion, “VaR measures the worst expected loss over a given horizon under normal market conditions at a given level of confidence.

For instance, a bank might say that the daily VaR of its trading portfolio is \$1 million at the 99 percent confidence level. In other words, under normal market conditions, only one percent of the time, the daily loss will exceed \$1 million.” (Jorion 2001, p. xxii)

More formally, VaR describes the quantile of the projected distribution of gains and losses over the target horizon. If c is the selected confidence level, VaR corresponds to the $1 - c$ lower-tail level.

2.2 Measuring returns

As the definition of VaR suggested, the VaR number is the portfolio return in the worst case, thus we first introduce the definition of portfolio return. The portfolio return ΔP is the difference between portfolio values, i.e. $\Delta P = P_{t+1} - P_t$, where P_t and P_{t+1} are the portfolio values at time t and $t + 1$ respectively.

The portfolio return can be described by the rate of return. However, there are two kinds of rates, namely, arithmetic and geometric. Arithmetic rate of return R_a is the

portfolio return divides the original portfolio value, i.e. $R_a = \frac{P_t - P_{t-1}}{P_{t-1}}$, and the

geometric rate of return R_g is the logarithm of the price ratio, i.e. $R_g = \ln \frac{P_t}{P_{t-1}}$.

Note that $R_g = \ln(P_t / P_{t-1}) = \ln(1 + R_a)$. If the time horizon is short, say one day, the daily arithmetic rate of return R_a is around zero, then by Taylor expansion, $R_g = R_a - \frac{R_a^2}{2} + \frac{R_a^3}{3} - \dots$, so $R_a \approx R_g$. Therefore, we shall treat arithmetic rate and geometric rate the same, and use R to denote both.

Let $R_{t,n}$ be the rate of return during the last n periods, by the means of geometric return, $R_{t,n} = \ln\left(\frac{P_t}{P_{t-n}}\right) = \ln\left(\frac{P_t}{P_{t-1}}\right) + \ln\left(\frac{P_{t-1}}{P_{t-2}}\right) + \dots + \ln\left(\frac{P_{t-n+1}}{P_{t-n}}\right) = R_t + R_{t-1} + \dots + R_{t-n+1}$, i.e. the rate of return during the last n periods is the sum of n previous rates. In particular, the portfolio return of two consecutive periods is $R_{t,2} = R_t + R_{t-1}$. (Jorion 2001, pp.99-101, 131-132)

2.3 Computation of VaR

In this section, we shall discuss four methods to compute the VaR number: distribution of portfolio return, rate of portfolio return, delta-normal valuation and delta-gamma method.

Method 1: Distribution of portfolio return.

Since the VaR number is the worst expected loss, the most straightforward method is to use the distribution of the portfolio return.

Let $f_{\Delta P}$ be the probability density function (pdf) of ΔP and c be the confidence interval, then the VaR number over some time horizon can be calculated from the following equation:

$$1 - c = \int_{-\infty}^{-VaR} f_{\Delta P}(x) dx \quad (2.1)$$

For a given portfolio, if the portfolio return is normally distributed with mean μ and standard deviation σ , the VaR number can be obtained as follows.

By standard normal table, there is a number α corresponding to the confidence level c . For example, if c is chosen to be 95 percent, the corresponding α is 1.65, and if c is 99 percent, α is 2.33. Since the VaR corresponds to the left tail, the actual cut line is $-\alpha$, as illustrated in the following graph.

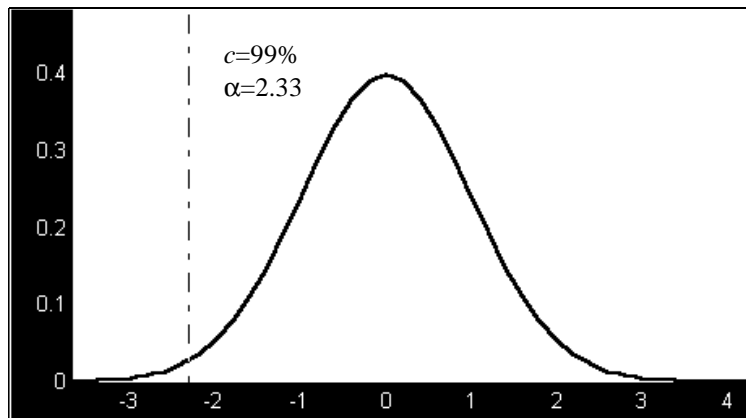


Fig 2.1. The standard normal curve with $c=99\%$.

To obtain the VaR number, we can do a standard transformation $\alpha = \frac{z - \mu}{\sigma}$ with $z = -VaR$ and $\alpha = -\alpha$, thus

$$VaR = \alpha\sigma - \mu \quad (2.2)$$

Furthermore, if $F_{\Delta P}(x)$ is the cumulative distribution function (cdf) of ΔP , the equation can be written as

$$1 - c = \int_{-\infty}^{-VaR} f_{\Delta P}(x) dx = F_{\Delta P}(-VaR) \quad (2.3)$$

Example 2.1

The daily return of a portfolio follows a normal distribution with mean \$1000 and standard deviation of \$500. The VaR number for this portfolio at the confidence level

99 percent can be computed directly from formula (2.2) by specifying the following factors:

The expected return $\mu = \$1000$;

The standard deviation of return $\sigma = \$500$;

The confidence level $c = 99\%$, which yields $\alpha = 2.33$.

Thus the VaR number can be reported as:

$$VaR = \alpha\sigma - \mu = 2.33 \times \$500 - \$1000 = \$165.$$

Method 2: Rate of portfolio return.

Suppose the current price of a portfolio is P_0 and the rate of return R for this portfolio is normally distributed with mean μ and standard deviation σ . Then the portfolio value at the end of the time horizon is $P_1 = P_0(1 + R)$ with mean $P_0(1 + \mu)$ and standard deviation $P_0\sigma$.

Firstly, we denote the lowest portfolio value at some confidence level c as $P_1^* = P_0(1 + R^*)$. Then the VaR number relative to the expected return is $VaR(mean) = E[P_1] - P_1^* = P_0(1 + \mu) - P_0(1 + R^*) = P_0(\mu - R^*)$. Assuming the expected return is zero, $VaR(zero) = -(P_1^* - P_0) = -P_0R^*$.

Next we can apply the same technique as in method 1, with the transformation equation $R^* = -(\alpha\sigma - \mu)$. By substituting R^* into the above equations, we have the following two formulas (Jorion 2001, pp.110-113):

$$VaR(zero) = P_0(\alpha\sigma - \mu) \tag{2.4}$$

$$VaR(mean) = P_0\alpha\sigma \tag{2.5}$$

Example 2.2

Assume that we need to measure the VaR number of a \$100 million equity portfolio over one year at the 99 percent confidence level; the following steps are required to compute VaR:

Evaluate the current portfolio: \$100 million;

Measure the standard deviation of the rate of return: 15 percent per annum;

Set the time horizon: one year;

Set the confidence level: 99 percent, which yields $\alpha = 2.33$;

Report the worst loss by processing all the preceding information:

$$VaR = P_0 \alpha \sigma = \$100M \times 2.33 \times 15\% = \$34.95M .$$

Method 3: Delta-normal valuation.

For simplicity, we will discuss the delta-normal valuation for an instrument whose value depends on a single risk factor – spot price S .

The first step is valuing the portfolio at the initial point. Let P be a pricing function of spot price S , then the price at the initial position is $P_0 = P(S_0)$. If Δ_0 is the

first partial derivative of function P evaluated at P_0 , we have $\Delta_0 = \left. \frac{\partial P}{\partial S} \right|_{S=S_0}$. Therefore,

$$dP = \Delta_0 \times dS .$$

Next, since Δ_0 is constant, the potential loss in value dP is a linear function of changes in S . Therefore, the worst loss for P can be obtained from the worst S , i.e. $VaR = |\Delta_0| \times VaR_S$. Assuming dS/S is a normal variable with mean zero, then by formula (2.4), we have (Jorion 2001, pp. 206-209)

$$VaR = |\Delta_0| \times (\alpha \sigma S_0) \tag{2.6}$$

Note that the α in (2.4) and (2.5) is the standard deviation of the return rate R , or $\Delta P / P$. Similarly, the σ in the above formula should be $\sigma(dS/S)$.

Example 2.3

Suppose there is a long call option for a given asset whose value is determined by $P = 0.2S$, where S is the spot price and P is the option value for the asset. The current price for the asset is \$100, and the VaR for the option value can be obtained through following steps.

Estimate Δ at initial position: $\Delta_0 = 0.2$;

Evaluate the current price: $S_0 = \$100$;

Measure the standard deviation of spot price: $\sigma(dS/S) = 1$ percent per day;

Set time horizon: daily;

Set the confidence level: $c = 99$ percent, which yields $\alpha = 2.33$;

Report the worst loss by processing all the predating information:

$$VaR = |\Delta_0| \times (\alpha \sigma S_0) = 0.2 \times 2.33 \times 0.01 \times \$100 = \$0.466.$$

Method 4: Delta-gamma method.

It is possible to increase the accuracy of delta-normal method. This purpose can be achieved by adding terms in the Taylor expansion of the valuation function:

$$dP = \frac{\partial P}{\partial S} dS + \frac{1}{2} \frac{\partial^2 P}{\partial S^2} dS^2 + \dots = \Delta dS + \frac{1}{2} \Gamma dS^2 + \dots \quad (2.7)$$

where Γ is the second derivative of the portfolio value.

Now, if the current spot price is S_0 , the worst loss around its mean is $\alpha \sigma S_0$. In such cases, the actual spot price will be $S_0 - \alpha \sigma S_0$, and its corresponding portfolio value is $P(S_0 - \alpha \sigma S_0)$, where P is the pricing function. Therefore, the portfolio VaR is

$$VaR = P(S_0) - P(S_0 - \alpha \sigma S_0). \quad (2.8)$$

Using formula (2.7), we can expand $P(S_0 - \alpha \sigma S_0) = P(S_0) + \Delta(-\alpha \sigma S) + \frac{1}{2} \Gamma(-\alpha \sigma S)^2 = P(S_0) - \Delta(\alpha \sigma S) + \frac{1}{2} \Gamma(\alpha \sigma S)^2$. Hence, the following formula can be obtained (Jorion 2001, pp.211-214):

$$VaR = |\Delta| \alpha \sigma S - \frac{1}{2} \Gamma(\alpha \sigma S)^2. \quad (2.9)$$

Example 2.4

Suppose there is a long call option for a given asset whose value is determined by $P = 0.2S + 0.02S^2$, where S is the spot price for the underlying asset and P is the option value. The current price for the asset is \$100, and the VaR for the option value with Delta-normal valuation is \$0.466 as in the previous example. Now using the Delta-gamma method, we can have a more accurate number. The procedures are illustrated in the following steps.

Estimate parameters: $\Delta = 0.2$ and $\Gamma = 0.04$;

Evaluate the current price: $S = \$100$;

Measure the standard deviation of spot price: $\sigma(dS/S) = 1$ percent per day;

Set time horizon: daily;

Set the confidence level: $c = 99$ percent, which yields $\alpha = 2.33$;

Using the above information and formula (2.9), the VaR number for this portfolio can be calculated as:

$$\begin{aligned} VaR &= |\Delta| \alpha \sigma S - \frac{1}{2} \Gamma (\alpha \sigma S)^2 \\ &= 0.2 \times 2.33 \times 0.01 \times \$100 - \frac{1}{2} \times 0.04 \times (2.33 \times 0.01 \times 100)^2 = \$0.358, \end{aligned}$$

which is smaller than \$0.466.

2.4 VaR measurement over different parameters

There are two key quantitative factors presented in the definition of VaR, namely, the time horizon τ and the confidence level c . In this section, we will discuss how to convert VaR numbers between different parameters.

Recall that returns of two consecutive periods $R_{t,2}$ is equal to $R_t + R_{t-1}$. Assume the risk factors are identically and independently distributed in any period, $\mu(R_t) = \mu(R_{t-1}) = \mu(R)$. Therefore, $\mu(R_{t,2}) = \mu(R_t) + \mu(R_{t-1}) = 2\mu(R)$. Also note that by independence, $\text{cov}(R_t, R_{t-1}) = 0$, which impels $\sigma(R_{t,2}) = \sqrt{2}R_t$. More generally, if the time period is Δt , $\mu(R_{t,\Delta t}) = \Delta t \times \mu(R_t)$, $\sigma(R_{t,\Delta t}) = \sqrt{\Delta t} \times \sigma(R_t)$. Therefore, a generalized formula for VaR computation is

$$VaR = P_0 \alpha \sigma \sqrt{\Delta t} \tag{2.10}$$

Example 2.5

Assume that we need to measure the VaR of a \$100 million equity portfolio over ten days at the 99 percent confidence level, and the annum standard deviation is 15 percent, the VaR should be computed using (2.10) as

$$VaR = P_0 \alpha \sigma = \$100M \times 2.33 \times 15\% \times \sqrt{\frac{10}{250}} = \$6.99M.$$

If the time horizons are Δt_1 and Δt_2 respectively, then the conversion formula for different time horizon is

$$\frac{VaR_{\Delta t_1}}{VaR_{\Delta t_2}} = \frac{\sqrt{\Delta t_1}}{\sqrt{\Delta t_2}}. \quad (2.11)$$

Similarly, if the confidence levels are c_1 and c_2 , each corresponding to α_1 and α_2 in the normal table, then the conversion formula for different confidence interval is:

$$\frac{VaR_{c_1}}{VaR_{c_2}} = \frac{\alpha_1}{\alpha_2}. \quad (2.12)$$

Example 2.6

As a simple application, the RiskMetrics(RM) risk use a 95 percent confidence interval ($\alpha = 1.65$) over one day, and the Basel Committee (BC) use a 99 percent confidence interval ($\alpha = 2.33$) over ten days. Then we can convert the VaR number from RM to BC as follows:

$$VaR_{BC} = VaR_{RM} \times \frac{\alpha_2 \sqrt{t_2}}{\alpha_1 \sqrt{t_1}} = VaR_{RM} \times \frac{2.33 \sqrt{10}}{1.65} = 4.45 VaR_{RM}.$$

Clearly, a longer time horizon or a greater confidence level will produce a larger VaR number.

2.5 Choice of parameters

Now we turn to the discussion of the choice of quantitative factors, including the time horizon τ and the confidence level c . VaR has many applications, and in different situations, there are different criteria to choose the quantitative factors (Jorion 2001, pp. 116-120). Consider the following applications of VaR:

If the VaR number is used to compare risks among different market, the choice of factors is arbitrary as long as consistency is maintained. For instance, Bankers Trust

uses a 99 percent confidence interval and one year horizon to compare risks over various units.

If it is used to measure potential loss, the factors should be determined by the nature of portfolio. For example, most banks use daily VaR due to rapid turnover of the portfolios. For this application, the choice of confidence interval is relatively arbitrary since VaR is only a probabilistic measure.

If the VaR number is used as equity capital measurement, the choice of factors is crucial. The equity capital measurement is used for regulatory purposes and to prevent a loss more than the whole equity of the institution. The time horizon should be chosen such that actions reducing risk could be taken in time, and the confidence level should reflect the degree of risk aversion. One illustration of the use of VaR as equity capital is the Basel Committee model, which uses a 10-business days (two weeks) horizon and a 99 percent confidence interval. However, in a 99 percent interval, a loss worse than VaR will happen 1 percent of the time, or once every four years. It is not possible for a big company to fail so often. So Basel Committee multiply the VaR number by a safety factor $k = 3$ as equity capital measurement.

Chapter 3

Portfolio VaR

3.1 Portfolio VaR for Multiple Assets

In the last chapter, we have introduced some methods to compute VaR for a single asset portfolio. Now, we shall extend the portfolio to multiple assets and continue the discussion of VaR computation.

Recall formula (2.5), $VaR = \alpha\sigma P_0$. If the confidence level is given and the portfolio value is known, the only variable is the standard deviation of the rate of return, which can be obtained from the following steps (Jorion 2001, pp.148-150).

Step 1: Determine the rate of return R_p

The portfolio rate of return for multiple assets is a linear function of rates for each underlying asset. Therefore, the portfolio rate of return $R_p = \sum_{i=1}^N w_i R_i$, where N is the number of assets, R_i is the rate of return for asset i , and w_i is the weight defined by $w_i = \frac{P_i}{P_0}$ with P_0 the portfolio value and P_i the value of asset i .

The portfolio rate of return R_p can also be expressed in matrix form:

$$R_p = [w_1 \quad w_2 \quad \dots \quad w_N] \begin{bmatrix} R_1 \\ R_2 \\ \dots \\ R_N \end{bmatrix}.$$

Let \mathbf{w} be the column vector of weight and \mathbf{R} be the column vector of rate of return, we have $R_p = \mathbf{w}^T \mathbf{R}$, where \mathbf{w}^T is the transpose of the weight matrix.

Step 2: Measure the standard deviation of return rate σ_p

Since the linear sum of normal variables is normally distributed, and each R_i is normal, the rate of return R_p is normally distributed. Thus, the expected portfolio rate of return $\mu_p = \sum_{i=1}^N w_i \mu_i$, and the variance of rate $\sigma_p^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}$, where σ_{ij} is the covariance matrix.

By our assumptions, the mean μ_i of each asset is zero, so the covariance $\sigma_{ij} = \text{cov}(R_i, R_j) = E[(R_i - \mu_i)(R_j - \mu_j)] = E(R_i R_j)$. Therefore, $\sigma_{ij} = \sigma_{ji}$, and $\sigma_{ii} = \sigma_i^2$. Thus, the standard deviation can be computed from $N(N-1)/2$ different terms using

$$\text{formula } \sigma_p^2 = \sum_{i=1}^N w_i^2 \sigma_i^2 + 2 \sum_{i=1}^N \sum_{j=1, j < i}^N w_i w_j \sigma_{ij}.$$

The variance of rate of return σ_p^2 also has a matrix representation:

$$\sigma_p^2 = \begin{bmatrix} w_1 & \dots & w_N \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1N} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2N} \\ \dots & \dots & \dots & \dots \\ \sigma_{N1} & \sigma_{N2} & \dots & \sigma_{NN} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \dots \\ w_N \end{bmatrix}.$$

By defining Σ as the covariance matrix, the variance of the portfolio rate of return can be written more compactly as $\sigma_p^2 = \mathbf{w}^T \Sigma \mathbf{w}$.

Step 3: Compute the VaR number

Using the above information, we are ready to compute the VaR number for a multiple assets portfolio.

Multiply equation $\sigma_p^2 = \mathbf{w}^T \Sigma \mathbf{w}$ both sides by P_0^2 , we have $\sigma_p^2 P_0^2 = \mathbf{x}^T \Sigma \mathbf{x}$, where \mathbf{x} is the column vector of asset values. Substituting to formula (2.4), the VaR number can be obtained by the following formula:

$$VaR_p = \alpha \sigma_p P_0 = \alpha \sqrt{\mathbf{x}^T \Sigma \mathbf{x}}. \quad (3.1)$$

3.2 VaR Reduction Methods

The VaR number of a portfolio depends on the covariance of the variables and the number of assets. Lower portfolio risk can be achieved through low correlation or large number of assets. (Jorion 2001, p.150)

Correlation:

Consider the formula (3.1), $VaR_p = \alpha \sqrt{\mathbf{x}^T \Sigma \mathbf{x}}$. The covariance matrix Σ plays an important role in determining the VaR. Assuming other factors are unchanged, a smaller covariance matrix leads to a smaller VaR number.

Covariance measures the extent to which two variables move linearly together. If the two variables are independent, the covariance is zero. A positive covariance means the two variables move in the same direction and a negative covariance means they move in the opposite direction.

The magnitude of covariance depends on the variances of the individual components, so the correlation defined by $\rho_{12} = \sigma_{12} / \sigma_1 \sigma_2$ is more convenient compared to covariance.

Note that ρ is range from -1 to 1 . When ρ is unity, the two variables are perfectly correlated and when ρ is zero, they are uncorrelated.

Number of assets:

To see the effect of N , assume a portfolio consists of N assets and has value 1. Also assume that the standard deviation of return rate of any asset is σ_1 and the covariance of any two assets is σ_2 , then the standard deviation for return rate of the

portfolio is $\sigma_p^2 = N\left(\frac{1}{N}\right)^2 \sigma_1^2 + N(N-1)\left(\frac{1}{N}\right)^2 \sigma_2 = \frac{\sigma_1^2}{N} + \left(1 - \frac{1}{N}\right)\sigma_2$. Since the correlation

between any two assets is $\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2} = \frac{\sigma_2}{\sigma_1^2}$, therefore, $\sigma_p = \sigma_1 \sqrt{\frac{1}{N}(1-\rho) + \rho}$. We can

easily see that as N increases, σ_p becomes smaller, so does the portfolio risk.

Based on these discussions, we introduce two definitions: diversified VaR and undiversified VaR. A diversified VaR is the portfolio VaR taking into account diversification benefits between components, whereas an undiversified VaR is the portfolio VaR with all its components highly correlated.

Example 3.1

Assume a portfolio P with two components, each carries weight w_1 and w_2 . The standard deviation for each of its component is σ_1 and σ_2 respectively, and the correlation is ρ_{12} . Then the VaR for this portfolio can be carried on as follows:

$$\text{The variance: } \sigma_p^2 = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\sigma_{12} = w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2;$$

$$\text{The portfolio VaR: } VaR_p = \alpha\sigma_p P = \alpha\sqrt{w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + 2w_1w_2\rho_{12}\sigma_1\sigma_2} P.$$

When the correlation ρ is -1 , the VaR is minimized and we have

$$VaR_p = \sqrt{VaR_1^2 + VaR_2^2 - 2VaR_1 \times VaR_2} = |VaR_1 - VaR_2|;$$

when the correlation ρ is zero, the VaR for the portfolio is

$$VaR_p = \sqrt{(\alpha w_1 P \sigma_1)^2 + (\alpha w_2 P \sigma_2)^2} = \sqrt{VaR_1^2 + VaR_2^2};$$

and when the two components are perfectly correlated, i.e. $\rho = 1$, the VaR is maximized:

$$VaR_p = \sqrt{VaR_1^2 + VaR_2^2 + 2VaR_1 \times VaR_2} = VaR_1 + VaR_2.$$

Note that the VaR for the portfolio is always less than or equal to the sum of individual VaR, i.e. $VaR_p \leq VaR_1 + VaR_2$. (Jorion 2001, p.152)

3.3 VaR Tools

One important purpose for VaR method is to control and manage the risk. Therefore, besides the computation of VaR for the entire portfolio, we also want to

know which asset contributes most to the total risk, what the effect is if delete or add an asset, and so on. In this section, a detailed analysis of VaR tools is introduced to control and manage the portfolio risk.

Marginal VaR

The first tool for risk management is the marginal VaR, which is defined as the partial derivative with respect to the component weight. It measures the change in portfolio VaR resulting from adding additional dollar to a component.

Take the partial derivative of the variance with respect to w_i :

$$\frac{\partial \sigma_p^2}{\partial w_i} = 2w_i\sigma_i^2 + 2 \sum_{j=1, j \neq i}^N w_j\sigma_{ij} = 2\text{cov}(R_i, w_iR_i + \sum_{j \neq i}^N w_jR_j) = 2\text{cov}(R_i, R_p)$$

Since $\partial \sigma_p^2 = 2\sigma_p \partial \sigma$, the above equation is equivalent to $\frac{\partial \sigma_p}{\partial w_i} = \frac{\text{cov}(R_i, R_p)}{\sigma_p}$,

therefore, $\Delta VaR_i = \frac{\partial \alpha \sigma_p P}{\partial x_i} = \alpha \frac{\partial \sigma_p}{\partial w_i} P = \alpha \frac{\partial \sigma_p}{\partial w_i} = \alpha \frac{\text{cov}(R_i, R_p)}{\sigma_p} = \frac{\alpha \sigma_{ip}}{\sigma_p}$. Thus, the marginal VaR for i^{th} component is (Jorion 2001, pp.154-155)

$$\Delta VaR_i = \frac{\alpha \sigma_{ip}}{\sigma_p} \quad (3.2)$$

The marginal VaR is closely related to vector $\boldsymbol{\beta}$, which has its i^{th} component defined by $\beta_i = \frac{\text{cov}(R_i, R_p)}{\sigma_p^2} = \frac{\sigma_{ip}}{\sigma_p^2}$.

Recall that $\sigma_p^2 = \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$ and $\sigma_{ip} = \sum \mathbf{w}$, the vector $\boldsymbol{\beta}$ can be expressed in matrix notation as $\boldsymbol{\beta} = \frac{\boldsymbol{\Sigma} \mathbf{w}}{\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}}$. Since $\rho_{12} = \sigma_{12} / \sigma_1 \sigma_2$, we have $\beta_i = \frac{\rho_{ip} \sigma_i \sigma_p}{\sigma_p^2} = \rho_{ip} \frac{\sigma_i}{\sigma_p}$.

Thus the relationship between ΔVaR_i and β_i is

$$\Delta VaR_i = \alpha (\beta_i \times \sigma_p) = \frac{VaR}{P} \times \beta_i. \quad (3.3)$$

Incremental VaR

Another tool for risk management is the incremental VaR, which measures the change in VaR due to a new position on the portfolio.

Let \mathbf{a} be the new position added, and a_i is the amount invested on asset i . Then intuitively the incremental VaR can be defined by the difference between the new VaR and original VaR, i.e.

$$\text{Incremental VaR} = \text{VaR}_{p+a} - \text{VaR}_p. \quad (3.4)$$

However, to calculate the VaR for the new portfolio, we need to compute the new covariance matrix, which might be time consuming. Therefore, the following approximation sometimes is used to shorten the computation time.

Expanding VaR_{p+a} , we have $\text{VaR}_{p+a} = \text{VaR}_p + (\Delta \text{VaR})^T \times \mathbf{a} + \dots$. Thus, when a is small relative to P , (Jorion 2001, p.155-159)

$$\text{Incremental VaR} \approx (\Delta \text{VaR})^T \times a. \quad (3.5)$$

Therefore, we can simultaneously compute ΔVaR_p and VaR_p . When a new trade is added to the portfolio, the approximation of incremental VaR can be immediately known by formula (3.5).

If only one asset is added to the portfolio, we can choose the amount to invest so that the risk is minimized. This action is also called best hedge.

Suppose amount a_i is invested on asset i , then the new portfolio $P_N = P + a_i$, and the variance of returns for P_N is $\sigma_N^2 P_N^2 = \sigma_p^2 P^2 + 2a_i P \sigma_{ip}^2 + a_i^2 \sigma_i^2$. Differentiating with

respect to a_i , we get $\frac{\partial \sigma_N^2 P_N^2}{\partial a_i} = 2P \sigma_{ip}^2 + 2a_i \sigma_i^2$. Thus the best hedge occurs when the

equation equals to zero, or

$$a^* = -P \frac{\sigma_{ip}^2}{\sigma_i^2}. \quad (3.6)$$

Recall the definition of β , the optimal of a_i can also be computed by the following formula:

$$a_i^* = -P\beta_i \frac{\sigma_p^2}{\sigma_i^2} \quad (3.7)$$

Component VaR

The other tool that is extremely useful to manage risk is the component VaR, which is a partition of the portfolio VaR that indicates the change of VaR if a given component was deleted. We can use it to have a risk decomposition of the current portfolio.

As discussed before, the sum of individual VaRs is not so useful since it discards the diversification effects. Thus, we define the component VaR in term of marginal VaR as follows (Jorion 2001, pp.159-161):

$$\text{Component VaR} = (\Delta VaR_i) \times w_i P = VaR \beta_i w_i . \quad (3.8)$$

Note that the sum of all component VaRs (CVaR) is the VaR_p for the entire portfolio,

$$\text{i.e. } \sum_{i=1}^N CVaR_i = VaR \left(\sum_{i=1}^N w_i \beta_i \right) = VaR_p .$$

Recall the definition of β , we can obtain the following formula as well:

$$CVaR_i = VaR_p w_i \beta_i = (\alpha \sigma_i P) w_i \beta_i = VaR_i \rho_i . \quad (3.9)$$

And the percentage of each asset carries is $p(i) = \frac{CVaR_i}{VaR} = w_i \beta_i$.

Example 3.2

Consider a portfolio with two foreign currencies, the Canadian dollar (CAD) and the euro (EUR). Assume that these two currencies are uncorrelated and have a standard deviation of 5 and 12 percent respectively. The portfolio has \$2 million invested in CAD and \$1 million in EUR.

1. The portfolio VaR at a 95 confidence interval

Firstly we will compute the variance for the portfolio

$$\sigma_p^2 P^2 = x^T \sum x = \begin{bmatrix} \$2 & \$1 \end{bmatrix} \begin{bmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{bmatrix} \begin{bmatrix} \$2 \\ \$1 \end{bmatrix} = 0.0244$$

using $\alpha = 1.65$, $VaR_p = 1.65 \times \sqrt{0.0244} \times 10^6 = \$257,738$

2. The individual VaR

The individual VaR is

$$\begin{bmatrix} VaR_1 \\ VaR_2 \end{bmatrix} = \begin{bmatrix} 1.65 \times 0.05 \times \$2 \text{ million} \\ 1.65 \times 0.12 \times \$1 \text{ million} \end{bmatrix} = \begin{bmatrix} \$165000 \\ \$198000 \end{bmatrix}$$

3. The marginal VaR

First, we can compute β for the original portfolio.

$$\beta = \frac{\text{cov}(R_i, R_p)}{\sigma_p^2} = \begin{bmatrix} 0.0050 \\ 0.0144 \end{bmatrix} / 0.0244 = \begin{bmatrix} 0.205 \\ 0.590 \end{bmatrix}$$

and the marginal VaR is

$$\Delta VaR = \alpha \beta \sigma_p = 1.65 \times \begin{bmatrix} 0.205 \\ 0.590 \end{bmatrix} \times 0.156 = \begin{bmatrix} 0.0528 \\ 0.1521 \end{bmatrix}$$

4. The incremental VaR if increasing the CAD position by \$10,000

The incremental VaR by approximation method is

$$(\Delta VaR)^T \times a = \begin{bmatrix} 0.0528 & 0.1521 \end{bmatrix} \begin{bmatrix} 10000 \\ 0 \end{bmatrix} = \$528,$$

the actual value by the full revaluation is

$$\sigma_{p+a}^2 P^2 = \begin{bmatrix} \$2.01 & \$1 \end{bmatrix} \begin{bmatrix} 0.05^2 & 0 \\ 0 & 0.12^2 \end{bmatrix} \begin{bmatrix} \$2.01 \\ \$1 \end{bmatrix} = 0.0245$$

and the VaR for the new portfolio is

$$VaR_{p+a} = \alpha \sigma_{p+a} P = 1.65 \times \$0.1565 \times 10^6 = \$258267.$$

Then the incremental VaR is $VaR_{p+a} - VaR_p = \$258,267 - \$257,738 = \$529$.

The difference between the two incremental VaR is only one dollar, thus our approximation method is acceptably accurate.

5. The component VaR for the original portfolio

The component portfolio can be computed using $CVaR_i = \Delta VaR_i x_i$

$$\begin{bmatrix} CVaR_1 \\ CVaR_2 \end{bmatrix} = \begin{bmatrix} 0.0528 \times \$2 \text{ million} \\ 0.1521 \times \$1 \text{ million} \end{bmatrix} = \begin{bmatrix} \$105,630 \\ \$152,108 \end{bmatrix}$$

Chapter 4

Modeling Time-Varying Risk

4.1 The existence of time-varying risk

In the last chapter, we assume that the risk factors are normally distributed with fixed mean μ and standard deviation σ . However, the risk of basic financial variables such as interest rate, exchange rate, appears to change over time.

As an illustration, let's consider the exchange rate between US dollars and British pound (\$/BP). The following graph plots the changes for the exchange rate from 1971 to 2000.

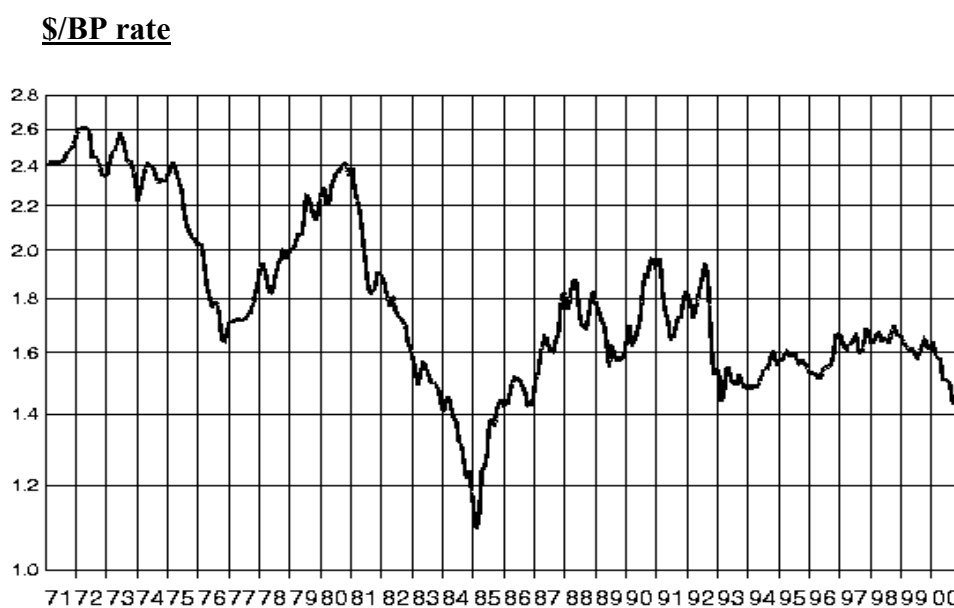


Fig 4.1. The exchange rates between US dollar and British pound.

Source: Pacific exchange rate service.

From the graph, we can see that the exchange between US dollars and British pound is quite stable before 1974, and it oscillates in a very small range between 2.2

to 2.6. However, after 1974, the exchange rate fluctuates intensely and displays more volatility.

Recall formula (2.5), $VaR = P_0 \alpha \sigma$, an increased standard deviation results in a larger VaR number. Thus, if the risk factor is not constant over time, forecasting time-varying risk is necessary to the VaR computation.

In this chapter, we will introduce three methods to estimate time-varying risk: moving average, GARCH estimation, and RiskMetrics approach.

4.2 Moving Average

Moving average (MA) is one of the most popular and easy to use tools available to measure time-varying risk. By using an average of prices, moving average provides a smooth trend line, which can be used to predict future changes in the risk factor.

Suppose we have the data of returns r_t over n days and we choose to use an M -day average. Then the day M is the first day possible to compute an average, and the average variance can be obtained by $\sigma_M^2 = \frac{1}{M} \sum_{t=1}^M r_t^2$. The variance for the day $M + 1$ can be obtained by adding the newest data r_{M+1} and dropping the earliest data r_1 .

Continue the process in this way: each day, the variance is updated by adding the most recent day's information and dropping the information M days ago, and divide the sum by M . The general formula for average computation is as follows:

$$\sigma_t^2 = \frac{1}{M} \sum_{i=0}^{M-1} r_{t-i}^2. \quad (4.1)$$

When we use up all n days of data, we can fit these points with a smooth line, and this line can indicate the trend of changes. (Jorion 2001, pp.186-187)

There is a parameter M for moving average. The choice of M is arbitrary, but usually it is an integral number such as 30 days or 60 days. We will see the effect of different M in the following example.

Let's consider the exchange rate between US dollars and British pound (\$/BP) again. Now we will use only a portion of the data between 1992 to 1993, and plot the graph with 30-day average, abbreviated as MA(30).

MA(30) for \$/BP



Fig 4.2 Moving average with $M = 30$ for exchange rate \$/BP.

Source: Pacific exchange rate service.

We can see that the MA(30) line ignores the small oscillation and closely follows the direction of actual standard deviation. If we use a 60-day average, we will get a different graph.

MA(60) for \$/BP

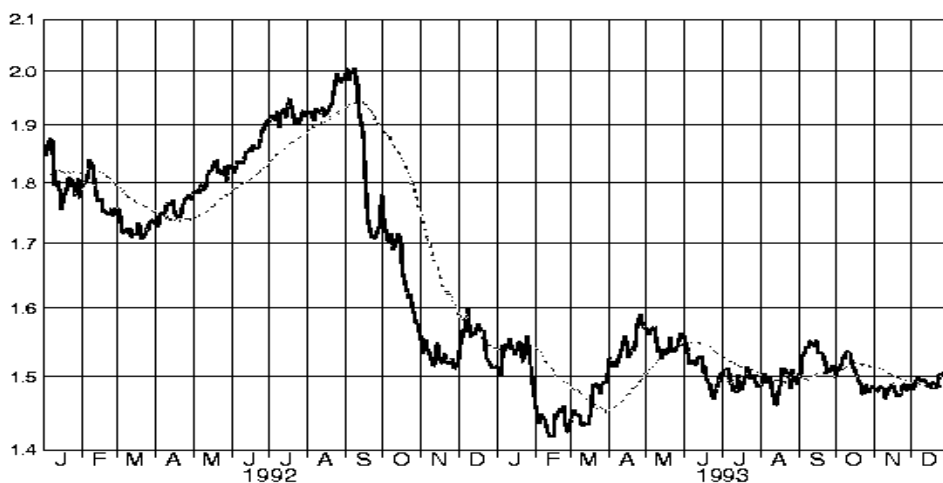


Fig 4.3 Moving average with $M = 60$ for exchange rate \$/BP.

Source: Pacific exchange rate service.

It is not difficult to see that the 60-day average is more stable since it decreases the weight of any single day.

Another observation from the graph is that, when the \$/BP exchange rates is rising between April to July, 1992, the MA line is always below the actual rate; when the exchange rates is going down between October to November, 1992, the MA line is always above the actual rate. This phenomenon indicates that moving averages are lagging indicators and will always be behind the price. When the price is rising, the moving average, which is based on the previous M -days of data, remains below the price. And if the price were trending down, the moving average most likely is above.

Suppose there is a large standard deviation some day, then this observation will increase the moving average immediately. If after that day the standard deviation returns to normal, this large average will remains until we drop this observation M -days later. As a result, the moving average measures of standard deviation tend to be plateaus of width M . Since recent information is more important, we want to put on more weight on the recent data. That is why we introduce the GARCH estimation.

4.3 GARCH Estimation

GARCH stands for *generalized autoregressive heteroskedastic*, and heteroskedastic means the lack of stationary volatility, i.e. the presence of periods of large standard deviation alternating with periods of small standard deviation.

The original version of GARCH model is ARCH. In an ARCH(q) process, the standard deviation at time t is a function of observed data at $t-1, t-2, \dots, t-q$. In 1986, Bollerslev introduced generalized ARCH or GARCH. In a GARCH(p, q) process, the standard deviation at time t depends on observed data at $t-1, t-2, \dots, t-q$, as well as on the standard deviation at $t-1, t-2, \dots, t-p$.

Define h_t as the conditional variance, i.e. the forecast of the variance of a time series at time t based on previous data. In GARCH model, h_t is a function of previous

conditional variance up to time $t - p$, and previous returns up to time $t - q$, and the value of h_t in the GARCH(p, q) process is

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_p r_{t-p}^2 + \beta_1 h_{t-1} + \beta_2 h_{t-2} + \dots + \beta_q h_{t-q}, \quad (4.2)$$

where r_t is the return on day t and h_t is the conditional variance on day t .

Here we focus our attention to the simplest case, which is GARCH(1,1) process. In GARCH (1, 1) model, the value of h_t can be calculated from the return r_{t-1} and the conditional variance h_{t-1} by the following formula (Jorion, 2001, pp.187-189):

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta h_{t-1}. \quad (4.3)$$

Computation of average standard deviation:

Since the return r_t is a normal variable with mean zero and variance h_t , we have

$$r_t = \sqrt{h_t} \varepsilon_t \text{ with } \varepsilon_t \sim N(0,1). \quad (4.4)$$

Then the mean of return $E[r_t] = E[\sigma_t] \times E[\varepsilon_t] = 0$, and the second moment of r_t $E[r_t^2] = E[\sigma_t^2] \times E[\varepsilon_t^2] = h_t$. Moreover, we can assume $h_t = h_{t-1} = h$, where h is the average variance. Therefore, we have $h = \alpha_0 + \alpha_1 h + \beta h$. Solving the equation for h , we get a general formula:

$$h = \frac{\alpha_0}{1 - \alpha_1 - \beta}. \quad (4.5)$$

Note that for this formula making sense, the sum of parameters α_1 and β must be less than unity.

Calculation of new standard deviation at $t + 1$:

Suppose we have the value for α_0, α_1 and β , and data of returns r_1 to r_t . In order to calculate the standard deviation σ_{t+1} , we can apply formula (4.3) with the replacement equation $h_t = \sigma_t^2$. Thus we have

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_1 r_t^2 + \beta \sigma_t^2,$$

and replacing σ_t by its value in function of σ_{t-1} and r_{t-1} , we get

$$\sigma_{t+1}^2 = \alpha_0 + \alpha_0\beta + \alpha_1\beta r_{t-1}^2 + \beta^2\sigma_{t-1}^2.$$

Replacing again the value of σ_{t-1} and repeating this operation up to the initial condition σ_0 , we can reach the following expression:

$$\sigma_{t+1}^2 = \alpha_0 \sum_{k=0}^t \beta^k + \alpha_1 \sum_{k=0}^t \beta^k r_{t-k}^2 + \beta^{k+1} \sigma_0^2 \quad (4.6)$$

This expression can be further simplified if assuming $\beta < 1$ and $t \ll 1$:

$$\sigma_{t+1}^2 = \frac{\alpha_0}{1-\beta} + \alpha_1 \sum_{k=0}^t \beta^k r_{t-k}^2 \quad (4.7)$$

Determination of parameters:

The most difficult part of GARCH model is the determination of the parameters. Since the model is nonlinear, we must do a numerical optimization. Here a maximum-likelihood method (ML) will be interpreted to estimate the parameters.

Note that $r_t = \varepsilon_t \sqrt{h_t}$ is approximately normally distributed with mean zero and standard deviation σ_t , and is *i.i.d.* If we have T observations, their joint density function is the product of the densities for each time period t , i.e. $\prod_{t=1}^T f(r_t)$ with f

being the normal density function defined by $f(r_t) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{r_t^2}{2\sigma^2}}$.

Thus the parameters can be obtained by maximization of the logarithm of the likelihood function:

$$\max F(\alpha_0, \alpha_1, \beta | r) = \sum_{t=1}^T \ln f(r_t) = \sum_{t=1}^T \left(\ln \frac{1}{\sqrt{2\pi}\sigma_t} - \frac{r_t^2}{2\sigma_t^2} \right) \quad (4.8)$$

Example 4.1

Once the model has been estimated, one can use the parameters α_0, α_1 and β , along with returns r_1 to r_t , to generate standard deviation σ_k from time 1 to $t + 1$. Since we use the data until r_t , only σ_{t+1} is the predicted value. However, the whole series of σ_k can be considered as forecasting of standard deviation. Here two sets of parameters will be given to illustrate the pattern of variance.

The exchange rate between US dollar and British pound with annual standard deviation 11.33% has the parameters as follows:

$$\alpha_0 = 0.00299, \alpha_1 = 0.0379 \text{ and } \beta = 0.9529.$$

The graph for its forecasting standard deviation is shown below:

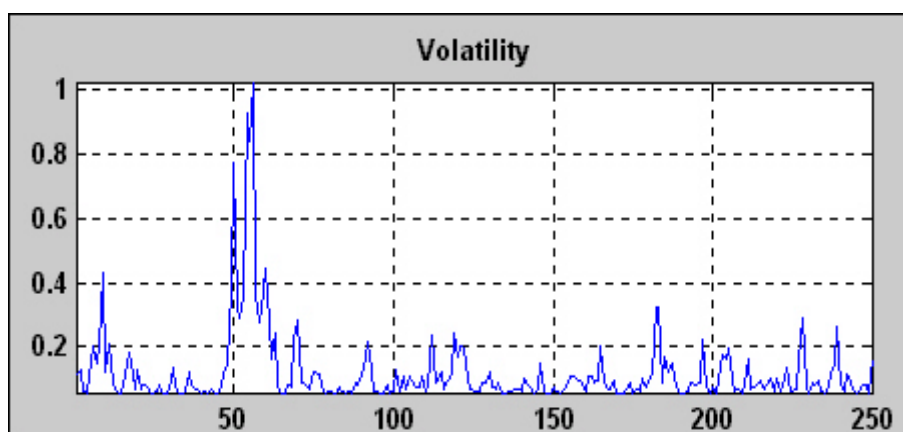


Fig 4.6. The standard deviation for exchange rate (\$/BP).

And the GARCH(1,1) parameters for exchange rate of Yen/\$ with annual standard deviation 11.78 percent is

$$\alpha_0 = 0.0104; \alpha_1 = 0.0528 \text{ and } \beta = 0.9284.$$

The graph below is the forecasting standard deviation for this exchange rate.

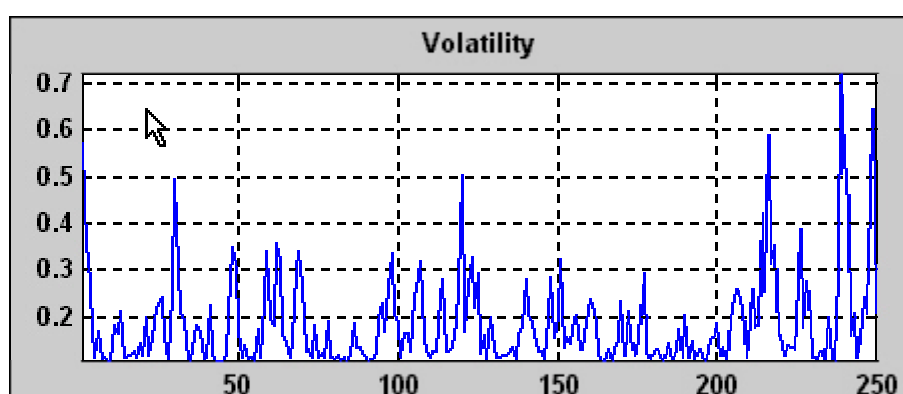


Fig 4.7. The standard deviation for exchange rate (Yen/\$).

4.4 RiskMetrics Approach

RiskMetrics approach was developed in 1994 and it uses exponential weighted moving average (EWMA) to forecast the time-varying risk.

Formally, the forecast variance for time t is a weighted average of the previous forecast, using weight λ , and of the latest squared innovation, using weight $(1 - \lambda)$. Thus the variance can be computed from (Jorion 2001, p.193)

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2. \quad (4.9)$$

Here, the parameter λ is called the decay factor and must be less than unity.

Weights of previous returns:

For any day t , the standard deviation σ_t can be computed from σ_{t-1} and r_{t-1} by formula (4.9). Replacing σ_{t-1} by its value in function of σ_{t-2} and r_{t-2} , we have

$$\sigma_t^2 = (1 - \lambda)r_{t-1}^2 + \lambda(1 - \lambda)r_{t-2}^2 + \lambda^2\sigma_{t-2}^2$$

Replacing σ_{t-2} again and recursively continue the process until reach the initial position. Then we obtain the following formula:

$$\sigma_t^2 = (1 - \lambda)(r_{t-1}^2 + \lambda r_{t-2}^2 + \lambda^2 r_{t-3}^2 + \dots) \quad (4.10)$$

Since the decay factor $\lambda < 1$, the weight of past observations declines geometrically. Therefore, RiskMetrics approach assigns greater importance to recent data of returns instead of simple average in moving average method.

Comparison of RiskMetrics and GARCH:

The RiskMetrics model and GARCH model are closely related. As we shall see that RiskMetrics is actually a special case of GARCH(1, 1) with some constraints on choice of parameters.

Suppose the decay factor λ is known and we want to calculate σ_{t+1} . By formula (4.10), we have

$$\sigma_{t+1}^2 = (1-\lambda)(r_t^2 + \lambda r_{t-1}^2 + \lambda^2 r_{t-2}^2 + \dots)$$

Comparing it with (4.7), if $\alpha_0 = 0$, $\alpha_1 + \beta = 1$, and decay factor $\lambda = \beta$,

$$\begin{aligned}\sigma_{t+1}^2 &= \frac{\alpha_0}{1-\beta} + \alpha_1 \sum_{k=0}^t \beta^k r_{t-k}^2 = 0 + (1-\beta)(\beta^0 r_t^2 + \beta^1 r_{t-1}^2 + \beta^2 r_{t-2}^2 + \dots) \\ &= (1-\lambda)(r_t^2 + \lambda r_{t-1}^2 + \lambda^2 r_{t-2}^2 + \dots)\end{aligned}$$

The two formulas are exactly the same. However, since there is only one parameter λ in RiskMetrics, it is particularly easy to implement.

Determination of decay factor:

We can interpret the ML estimation as an algorithm to estimate optimal decay factors. The optimal λ is obtained by minimizing

$$F(\lambda) = \sum_{k=1}^t (r_k^2 - \sigma_k^2)^2.$$

The minimum is reached at $r_k^2 = \sigma_k^2$, which is the same as in ML estimation for GARCH. However, ML for RiskMetrics puts equal weights on every data, i.e. it tries to get as close as we can σ_k^2 to r_k^2 , no matter the value of r_k^2 . On the other hand, ML for GARCH model puts more weights on large r_k^2 , i.e. the bigger value of r_k^2 is, the more the algorithm tries to get close σ_k^2 to r_k^2 . As a result, the ML for GARCH is more consistent than for RiskMetrics.

Chapter 5

Back testing VaR models

5.1 Model Back Testing with Exceptions

In this chapter, some of the validation methods will be discussed. Model validation is the general process of checking whether a model is adequate.

The application of VaR models should always be accompanied by validation. One of VaR validation methods is back testing, which verifies the actual losses are inline with the VaR measurement.

The first concern of back testing is the number of exceptions. Recall the definition of VaR. Suppose a VaR number is reported at the confidence interval c , then an exception occurs if its losses exceed the VaR number. Therefore, the expected number of exceptions N in a total of T observations is $T(1-c)$. Certainly, the number of exceptions will not exactly be $T(1-c)$; instead, it could swing in a small range.

In the back testing method, the range for N will be calculated and hence the VaR model can be accepted or rejected.

Number of exceptions

To illustrate the procedure of decision making, suppose following conditions are given:

Number of exceptions: N

VaR number: VaR

Total observations: T

VaR confidence level: c

Test confidence level: p

The actual daily loss exceeds VaR or not is a sequence of success or failure with probability $1-c$, thus assuming all the observations are independent, this is a

Bernoulli process, and follows a binomial distribution. The pdf for this binomial distribution is given by

$$f(x) = \binom{T}{x} (1-c)^x c^{T-x} \text{ for } x = 0, 1, 2, \dots \quad (5.1)$$

Note that for a binomial distribution, $E(x) = T(1-c)$ and $V(x) = Tc(1-c)$.

If the sample size T is large enough, we can apply the central limit theorem, and approximate the binomial distribution by a normal distribution.

Let $z = \frac{x - \mu}{\sigma} = \frac{x - T(1-c)}{\sqrt{Tc(1-c)}}$, then by central limit theorem, z follows a standard

normal distribution $\mathbf{N}(0, 1)$ with cdf $\phi(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$.

Therefore, given a test confidence level p , there is a range for z , say $|z| < \alpha$, where α is the number in the standard normal table corresponding to p . Then the range for x can be calculated as

$$-\alpha\sqrt{Tc(1-c)} + T(1-c) < x < \alpha\sqrt{Tc(1-c)} + T(1-c). \quad (5.2)$$

If the number of exceptions N is within the range, we accept the model, and if N is out of the range, we reject the model.

One observation from formula (5.2) is that, the interval for exceptions is dependent on the test confidence level p . A larger p leads to a smaller value of α and thus a smaller interval for N , and makes it easier to reject the current VaR model. On the other hand, a smaller p leads to a larger interval for N , and makes it easier to accept the current VaR model.

Another observation is that the interval for N usually is large. For example, if c is 99 percent and T is equal to 255, we accept the model as long as $N < 7$ at the test confidence level 95 percent. However, there is a high probability that $N < 7$ and the model is incorrect. How accurate our decision is? To measure the decision error, classically type I and type II error are involved. Type I error is the probability that the model is correct, but we reject it, and type II error is the probability that the model is

incorrect, but we accept it. A higher test confidence level leads to a smaller type I error but a larger type II error.

The Basel Committee Rules for Back Testing

The following table is the Basel Committee rules for back testing with the test confidence level to be 95 percent. It provides a simple way to accept or reject the current VaR model.

VaR Confidence Level	Nonrejection region for number of exceptions N		
	T = 255days	T = 510 days	T = 1000days
99%	$N < 7$	$1 < N < 11$	$4 < N < 17$
97.5%	$2 < N < 12$	$6 < N < 21$	$15 < N < 36$
95%	$6 < N < 21$	$16 < N < 36$	$37 < N < 65$
92.5%	$11 < N < 28$	$27 < N < 51$	$59 < N < 92$
90%	$16 < N < 36$	$38 < N < 65$	$81 < N < 120$

Table 5.1 Model back testing, 95% nonrejection test confidence regions

Source: Jorion (page 136)

Suppose there are two years of data ($T = 510$ days), then the expected number of exceptions is $\mu = T(1 - c) = 510 \times (1 - 95\%) \approx 26$. However, by the above table, VaR users will not reject the model as long as $N \in (16, 36)$. Values of N greater than or equal to 36 indicate that the VaR number is too small, or the current model understates the probability of large losses. Values of N less than or equal to 16 indicate that the VaR model is over conservative.

The table also shows that, as the sample size increases, the interval N/T shrinks, which means we can reject a model more easily if the sample size is large. For example, consider the row with a 95 percent confidence level. When $T = 255$, the N/T interval is approximately (0.024, 0.082); when $T = 510$, the interval becomes (0.031, 0.071); when $T = 1000$ days, the interval is (0.037, 0.065).

5.2 Model Back Testing with Conditional Coverage

In the last section, we have shown that with a VaR confidence level c , the number of exceptions should be in an interval around $T(1-c)$. Now we will focus on the distribution of the exceptions. Ideally they should be identically and independently distributed.

For instance, the expected number of exceptions with the VaR confidence level $c = 95\%$ is about $250 \times (1 - 95\%) \approx 13$ every year. If we actually observed 15 exceptions, with 10 of them occurred in the last two weeks, certainly this should invalidate the VaR model. However, back testing with exceptions is not able to reject the model. Therefore, back testing with conditional coverage is introduced to test the distribution of exceptions.

Setup for Conditional Coverage Model

Suppose we have the data of portfolio returns of T days. Firstly, give an indicator to each day, and set the indicator value as follows (Chatfield 2001):

Indicator $i = 1$ if it does not exceed VaR;

Indicator $i = 0$ otherwise.

Thus we have a sequence $\{I_t\}_{t=1}^T$ of 0s and 1s. For any two consecutive days, there are only four cases: 00, 01, 10 and 11.

Next, define $T_{i,j}$ ($i = 0, 1; j = 0, 1$) as the number of days that i, j occurs, T_0 as the number of days that the previous day's indicator is 0, and T_1 as the number of days that the previous day's indicator is 1.

Note that T_0 and T_1 equal to the number of 0 and 1 indicators respectively, and the sum of T_0 and T_1 is T .

Also define π as the ratio of the number of 1s T_1 to the total number of observations T , i.e. the probability of non-exception occurs in the sample. Let π_0 be

the conditional probability of 01 occurs if the previous day is 0 and π_1 be the conditional probability of 11 occurs if the previous day is 1. Then $\pi_0 = \frac{T_{01}}{T_0}, \pi_1 = \frac{T_{11}}{T_1}$, and the sum of π_0 and π_1 is π .

Assuming above definitions, the following table can be constructed.

	Conditional		Unconditional
	Day Before		
	Exception	No Exception	
Current day:			
Exception	$T_{00} = T_0 (1 - \pi_0)$	$T_{10} = T_1 (1 - \pi_1)$	$T (1 - \pi)$
No Exception	$T_{01} = T_0 (\pi_0)$	$T_{11} = T_1 (\pi_1)$	$T (\pi)$
Total	$T_0 = T_{00} + T_{01}$	$T_1 = T_{10} + T_{11}$	$T = T_0 + T_1$

Table 5.2 Expected number of exceptions table.

Now using this setup, we can do the likelihood ratio (LR) test for the conditional coverage model. The conditional coverage test consisting of two parts: the LR test of unconditional coverage (LR_{uc}) and the LR test of independence (LR_{ind}). Let's discuss the LR_{uc} test first.

The LR test of unconditional coverage

Consider the indicator sequence $\{I_n\}_{t=1}^T$. The expected probability of exceptions is the confidence level c , thus in the unconditional coverage test, the hypothesis $E[I_t] = c$ should be tested against the alternative $E[I_t] \neq c$.

For a sample data, the likelihood function is the probability of obtaining a particular set of data. With the notation defined in the setup, define the null likelihood function $L_0 = L(c; I_1, I_2, \dots, I_T) = (1-c)^{T_0} c^{T_1}$.

If the real probability is not c , we can write it as x and $L(x; I_1, I_2, \dots, I_T) = (1-x)^{T_0} x^{T_1}$.

Those values of parameters that maximize the sample likelihood are called maximum likelihood estimate (MLE). And the MLE for this sample of T days should be $\pi = \frac{T_1}{T}$. Therefore, the maximum likelihood function can be written as

$$L_1 = L(\pi; I_1, I_2, \dots, I_T) = (1-\pi)^{T_0} \pi^{T_1}.$$

The ratio $\lambda = \frac{L_0}{L_1}$ is always between 0 and 1. Now let $\text{LR}_{\text{uc}} = \chi^2 = -2 \ln \lambda$, the LR_{uc} is asymptotically distributed as chi-square with degree of freedom 1. By substitution, we have

$$\text{LR}_{\text{uc}} = -2T_0 \ln(1-c) - 2T_1 \ln c + 2T_0 \ln(1-\pi) + 2T_1 \ln \pi \quad (5.3)$$

The less likely the assumption is, the smaller λ will be, and the larger χ^2 will be. The critical value for chi-square at the 95 percent confidence interval with degree of freedom 1 is 3.84, therefore we will reject the model if $\text{LR}_{\text{uc}} > 3.84$.

The LR test of independence

Now we want to test the independence part of our conditional coverage hypothesis. Independence will be tested against the first order Markov alternative.

Consider a binary first-order Markov chain $\{I_n\}$, with transition probability matrix

$$\Pi_1 = \begin{bmatrix} 1-\pi_0 & \pi_0 \\ 1-\pi_1 & \pi_1 \end{bmatrix}.$$

If the sequence is independent, we should have the null transition matrix

$$\Pi_0 = \begin{bmatrix} 1-c & c \\ 1-c & c \end{bmatrix}$$

Therefore, the likelihood function for the null is $L_0 = L(\Pi_0, I_1, I_2, \dots, I_T) = (1-c)^{T_0} c^{T_1}$, and for the alternative is $L_1 = L(\Pi_1; I_1, I_2, \dots, I_T) = (1-\pi_0)^{T_{00}} \pi_0^{T_{01}} (1-\pi_1)^{T_{10}} \pi_1^{T_{11}}$.

Again, take the ratio $\lambda = \frac{L_0}{L_1}$ and let $\text{LR}_{\text{ind}} = \chi^2 = -2 \ln \lambda$, therefore, we have

$$\text{LR}_{\text{ind}} = -2T_0 \ln(1-c) - 2T_1 \ln c + 2T_{00} \ln(1-\pi_0) + 2T_{01} \ln \pi_0 + 2T_{10} \ln(1-\pi_1) + 2T_{11} \ln \pi_1.$$

Similarly, LR_{ind} is asymptotically distributed as $\chi^2(1)$. At the 95 percent confidence interval, we will reject the model if $\text{LR}_{\text{ind}} > 3.84$.

Chapter 6

Conclusions

This report contains a systematic study of Value at Risk.

In the introduction part, we introduce VaR as a risk management method developed in response to those famous financial disasters such as Barings's fall.

VaR measures the worst loss over some time horizon within some confidence level, and there are various methods of VaR computation. Since the essence of VaR modeling is the prediction of the highest expected loss for a given portfolio, these VaR techniques estimate losses by approximating the lower quantile in the portfolio return distribution.

VaR tools are extremely useful to risk management, which include marginal VaR, incremental VaR and component VaR. These tools can be used to diversify the portfolio and hence reduce the risk encountered.

Since the risk factors always change over time, the next topic is to measure the time varying risk. From the simplest moving average method to more complicated GARCH process and RiskMetrics approach, we can actually see the changes of risk factors.

Finally, the back testing methods are discussed. This method is used to validate the VaR model. We will accept the current VaR model only if the number of exceptions is within some range and these exceptions are independent.

References

[1] Chatfield Christopher, "Time Series Forecasting", Chapman & Hall / CRC, 2001, 181-214.

[2] Irina N. K & Svetlozar T. R, "Value at Risk: Recent Advances", Chapman & Hall / CRC, 2000, 801-858.

[3] Philippe Jorion, "Value at Risk", 2nd edition, McGraw-Hill, 2001.