All About Risk...

A Senior Project submitted to
The Division of Science, Mathematics, and Computing
of
Bard College

by
Dexin Zhou

Annandale-on-Hudson, New York
April, 2009
Dedicated to all my family members who raised me and always supported me to pursue my dreams.
Acknowledgments

I would like to thank my senior project advisors, Greg Landweber and Cliona Golden, for their dedicated discussion and instructions on my project. I also appreciate Professor Tsu-Yu Tsao for sitting on my board and offering his suggestions from an economics perspective. Mark Shi, a director of Citigroup’s Counterparty Risk Management, gave me the initial research idea when I mentioned to him that I was interested in the application of mathematics in the financial industry in a Chinatown restaurant. Without his help, it would have been very hard to start this project. Again, I appreciate my family members for their unreserved support. Last but not the least, I want to thank all my friends at Bard, who treated me like a family member and who gave me courage to move on in the hardest time of my college life.
Abstract

This project is a survey of different aspects of financial risk. My research focuses on the nature of stock price and stock return time series, since it is one of the key factors in the simulation of returns and losses over a period of time.

The main techniques used in this project are from the field of linear algebra and the Fast Fourier Transform (FFT). Time series in nature can be viewed as vectors that live in multi-dimensions. Therefore, linear algebra naturally arises when dealing with the existing data. The main techniques used are Singular Value Decomposition (SVD) and linear projection. The former technique examines the orthogonal composition of the data and the latter examines the relationship of a stock’s projection and the stock return to the time series itself. The Fast Fourier transform treats stock returns as samples from a signal and identifies the possible patterns out of the stock return time series.

There are two main results. First, the signal processing techniques tell us that there are no obvious regular cycles in stock return time series. Secondly, the stock sector information reflects in the singular value decomposition. Finally, there are many attempts in this research that are yet to be concluded.
# Contents

Dedication 1

Acknowledgments 2

Abstract 3

1 Introduction 6

2 Background Research 9

2.1 Steps in Constructing Value at Risk (VaR) 9
2.2 Variables 10
2.3 Other Expressions of VaR 11
2.4 Normal Distribution 11
2.5 Student’s t-distribution 12
2.6 VaR for Parametric Distribution 12
2.7 Desirable Properties of VaR 13
2.8 Covariance 13
2.9 Analytic Variance-Covariance, or Delta-Normal Approach 14
2.10 Delta-Normal VaR for Other Instruments 15
2.11 Simulation Method 15
2.12 The Bootstrap 16
2.13 Computing VAR 17
2.14 Simulations with Multiple Variables 17
2.15 The Cholesky Factorization 18
2.16 Number of Independent Factors 19
2.17 Singular Value Decomposition 20
2.18 Eigenvalues and Singular Values 20
1 Introduction

Risk is commonly regarded as a concept of the precise probability of specific eventualities. It is more precisely defined as a state of uncertainty where some of the possibilities involve a loss, catastrophe, or other undesirable outcome. In fact, we face all kinds of risks throughout our lives. The risks we face in our daily lives include car accidents, sickness, and investment portfolio depreciation. Therefore, evaluating risks is very important.

To a financial institution, managing risk is always considered one of the top priorities. Insufficient or badly managed risk can cause significant consequences to the firm. For instance, AIG’s risk management practices caused the debacle of the insurance giant. The Credit Default Swap (CDS) contracts that they sold severely damaged the firm’s financial conditions in the financial crisis. As a result, the AIG stock price fell to below one dollar from over $60. Another example comes from SocGen of France, where the rogue trader Jerome Kerviel caused a loss of 4 billion Euros or 7 billion dollars.

People in the financial market face different risks. Usually the risks are categorized into three categories: market risk, liquidity risk, and credit risk. Value-at-Risk (VAR) was invented as a measurement of a portfolio’s exposure to all categories’ risks, but it is
1. INTRODUCTION

Figure 1.0.1. AIG Stock Prices

most commonly used to measure the market risk. Value at Risk (VaR) is a mathematical approach for estimating the maximum potential loss of a given portfolio within some period of time with some likelihood of occurrence. It is sometimes regarded as a barometer for the trader.

The computation of VaR can be highly dependent on the nature of financial time series. The historical simulation and Monte-Carlo simulation methods rely heavily on the actual historical time series and the computer generated time series to predict the future price movement of certain securities. Therefore, this project focuses on the study of the nature and characteristics of financial time series. More specifically, we mainly consider equity and equity index time series, because the equity market data in the study is more readily attainable than data from other markets.

Chapter 1 discusses some of the standard computation techniques and theoretical definitions of Value at Risk. The chapter is followed by an introduction to the research set-up where the goals, methodologies, data collection and data analysis techniques are introduced. Chapter 4 discusses the analysis that treats stock price and stock return series as signals and uses Discrete Fourier Transform (DFT) to analyze the data. The following chapter, What do Singular and eigenvalues tell us?, discusses the results from using sin-
1. **INTRODUCTION**

Regular value decomposition and eigen-decomposition to analyze the stock return data. The linear projection chapter studies the relationship between the linear projection of a stock onto different set of stocks and the stock time series itself. Finally, we discuss the possible topics for future research.

Since our question is rather broad, the project is open-ended and the structure of the project is rather loose. This also reflects the process of the research. We decide to pursue whatever idea that seems interesting at the time and get partial results. In other words, this project deals with multiple ideas to try to address the same big question.
2
Background Research

The following discussions and mathematical calculations are standard in academic research. Some of these are direct citations from various sources [1–9].

2.1 Steps in Constructing Value at Risk (VaR)

The steps in constructing the VaR are:

1. Mark to market the current portfolio (the current value of the portfolio on the market).

2. Measure the variability of risk factors.

3. Set the time horizon.

4. Set the confidence level (the threshold probability that the value.)

5. Report the worst loss by processing all the preceding information.
Example 2.1.1. For instance a $100M portfolio with 15 percent annual volatility (standard deviation) in 10 days at 99 percent confidence level is

$$100M \times 15\% \times \sqrt{\frac{10 \text{ days}}{252 \text{ days}}} \times 2.33 = 7M,$$

where 252 is the approximate number of trading days in a year.

2.2 Variables

The following variables are commonly used in risk management and related probability and statistics texts. Many of the following variables are similar. Usually, when we are talking about the current situation, the variables are deterministic (predictable from the existing information) and variables concerning the future are stochastic (unpredictable from the existing information.)

- \(E(.)\) is the expected value of a variable.

- \(W\) is the portfolio/asset value in general. (In the context of a standard textbook, it usually stands for the value in the future and is stochastic.)

- \(R\) is the rate of return. We define \(R = (W_t - W_0)/W_0\). (In most contexts, this term is random, since we are talking about the return in the future.)

- \(R^*\) is the cut off return. It is the maximum loss under the given confidence interval. (It is a deterministic variable, since the confidence level is given.)

- \(\mu\) is the mean rate of return. We have \(\mu = E(R)\). It is deterministic.

- \(W_0\) is the current value of the portfolio or asset. It is a deterministic term, since we can mark the portfolio to market.

- \(W^* = W_0(1 + R^*)\) is the cut off value of the portfolio/asset corresponding to \(R^*\). In general \(R^*\) is negative. Thus we can write it as \(-|R^*|\). It is a deterministic term.
2. BACKGROUND RESEARCH

- **Relative VaR** is defined as \( \text{VAR(mean)} = E(W) - W^* = -W_0(R^* - \mu) \).

- **Absolute VaR** is defined as \( \text{VAR(zero)} = W_0 - W^* = -W_0R^* \).

- \( c \) is the probability of the portfolio value above specified \( W^* \). We define
  \[
  c = \int_{W^*}^{\infty} f(w) \, dw.
  \]

- \( p \) is the chance that the portfolio value falls below \( W^* \). We have
  \[
  1 - c = \int_{-\infty}^{W^*} f(w) \, dw = P(w \leq W^*) = p.
  \]

2.3 Other Expressions of VaR

We can also express VaR as \( \text{VAR}(H, c) \), where \( H \) is the time horizon and \( c \) is the confidence interval. Then we have

\[
\text{VAR}(H, c) = \sqrt{H} \, \text{VAR}(1, c).
\]

For example, we have \( \text{VAR}(10, c) = \sqrt{10} \, \text{VAR}(1, c) \).

2.4 Normal Distribution

The normal distribution is a very commonly used distribution. People generally assume the distribution of stock returns is normal. The probability density function of the normal distribution is

\[
\Phi(x) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left( -\frac{(x - \mu)^2}{2\sigma^2} \right),
\]

where \( \mu \) is the mean and \( \sigma \) is the standard deviation. The cumulative distribution function (the integral of the probability density function) is

\[
N(x) = \frac{1}{2} \left( 1 + \text{erf} \left( \frac{x - \mu}{\sigma \sqrt{2}} \right) \right).
\]

The function \( \exp(x) \) is the exponential function and \( \text{erf}(x) \) is the error function.
2. **BACKGROUND RESEARCH**

2.5 Student’s t-distribution

The student’s t-distribution is sometimes used to deal with the “fat tail” of stock return distribution. The probability density function of the student’s t-distribution is

\[
p(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi} \Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)}
\]

where

- \( \text{mean} = 0 \) for \( \nu > 1 \), otherwise undefined.
- \( \text{variance} = \nu/(\nu - 2) \) for \( \nu > 2 \), otherwise undefined.

2.6 VaR for Parametric Distribution

The VaR computation can be simplified under the assumption that it belongs to a parametric family, such as the normal distribution. We translate the general distribution \( f(w) \) to a standard normal distribution \( \Phi(\epsilon) \).

- \( R^* \) with a standard normal deviate of \( \alpha > 0 \) by setting

\[
-\alpha = \frac{-|R^*| - \mu}{\sigma}
\]

is equivalent to setting \( 1 - c = \int_{-\infty}^{W^*} f(w) \, dw = \int_{-\infty}^{-|R^*|} f(r) \, dr = \int_{-\infty}^{-\alpha} \Phi(\epsilon) \, d\epsilon \).

- Standard Normal Cumulative Distribution is \( N(d) = \int_{-\infty}^{d} \Phi(\epsilon) \, d\epsilon \).
- The cut-off return is \( R^* = -\alpha\sigma + \mu \).
- \( \text{VAR}(\text{mean}) = -W_0(R^* - \mu) = W_0\alpha\sigma\sqrt{\Delta t} \).
- \( \text{VAR}(\text{zero}) = -W_0R^* = W_0(\alpha\sigma\sqrt{\Delta t} - \mu\Delta t) \).
2.7 Desirable Properties of VaR

A risk measure can be viewed as a function of the distribution of portfolio value \( W \). Value at Risk is summarized into a single number \( \rho(W) \) (it can be measured in the dollar value), with the following properties:

- Monotonicity: If \( W_1 \leq W_2 \), \( \rho(W_1) \geq \rho(W_2) \), or if one portfolio has systemically lower returns than another for all states of the world, its risk must be greater.

- Translation invariance. \( \rho(W + k) = \rho(W) - k \), or adding cash \( k \) to a portfolio should reduce its risk by \( k \).

- Homogeneity: \( \rho(bW) = b\rho(W) \), or increasing the size of a portfolio by \( b \) should simply scale its risk by the same factor (this rules out liquidity effects for large portfolios).

- Subadditivity: \( \rho(W_1 + W_2) \leq \rho(W_1) + \rho(W_2) \), or merging portfolios cannot increase risk.

2.8 Covariance

The covariance between two real-valued random variables \( X \) and \( Y \), with expected values \( E(X) = \mu \) and \( E(Y) = \nu \) is defined as

\[
\text{cov}(X, Y) = E((X - \mu)(Y - \nu)).
\]

Covariance can also be expressed as

\[
\text{cov}(X, Y) = \rho \sigma_x \sigma_y,
\]

where \( \rho \) is called correlation coefficient and \( \sigma_x \) and \( \sigma_y \) are the standard deviation of variable \( X \) and \( Y \).
2.9 Analytic Variance-Covariance, or Delta-Normal Approach

The analytic variance-covariance approach assumes that the distribution of the changes in portfolio value is normal.

**Example 2.9.1.** Consider a portfolio with two stocks, with $n_1$ shares of stock 1 valued at $S_1$ and $n_2$ shares of stock 2 valued at $S_2$. The value of the portfolio is $W = n_1S_1 + n_2S_2$.

We can select the stock prices $S_1$ and $S_2$ as the risk factors. We have:

$$R_W = \frac{\Delta W}{W} = \frac{n_1S_1}{W} \frac{\Delta S_1}{S_1} + \frac{n_2S_2}{W} \frac{\Delta S_2}{S_2} = \omega_1R_1 + \omega_2R_2,$$

where

- $R_W$ = the rate of return on the portfolio,
- $R_i$ = the rate of return on stock $i$, i.e., $R_i \equiv \Delta S_i / S_i$,
- $\omega_i$ = the percentage of the portfolio invested in stock $i$. $\sum \omega_i = 1$,
- $\Delta S$ is the difference between the security prices at time $t$ and time 0.

Prices are expected to be log-normally distributed, so that the log-returns during the period $(t - 1, t)$, i.e.,

$$R_t = \ln \frac{S_t}{S_{t-1}} = \ln \left( 1 - \frac{S_t - S_{t-1}}{S_{t-1}} \right) \sim \frac{\Delta S_t}{S_{t-1}},$$

are normally distributed.

We have $R_W \sim N(\mu_W, \sigma_W)$ with
2. BACKGROUND RESEARCH

\[ \mu_W = \sum_{i=1}^{2} \omega_i \mu_i, \quad \text{and} \]
\[ \sigma^2_W = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \text{cov}(R_1, R_2) \]
\[ = \omega_1^2 \sigma_1^2 + \omega_2^2 \sigma_2^2 + 2\omega_1 \omega_2 \rho \sigma_1 \sigma_2 \]
\[ = (\omega_1 \omega_2) \begin{pmatrix} \sigma_1^2 & \rho \sigma_1 \sigma_2 \\ \rho \sigma_1 \sigma_2 & \sigma_2^2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \]
\[ = (\omega_1 \omega_2) \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \]
\[ = w \Omega w^T = w \sigma C \sigma w^T, \]

where \( w = (\omega_1 \omega_2) \), \( \Omega \) is the variance-covariance matrix, \( \sigma \) is the diagonal standard deviation matrix, and \( C \) is the correlation matrix.

2.10 Delta-Normal VaR for Other Instruments

We can use a first-order Taylor expansion of the pricing equation to generalize our previous analysis, giving us

\[ dW = \sum_{i=1}^{n} \frac{\partial W}{\partial f_i} df_i = \sum_{i=1}^{n} \Delta_i df_i, \]

where \( \Delta_i \) denotes the “delta” of the position in the instrument. It follows that

\[ \sigma(dW) = \sqrt{\sum_{i=1}^{n} \Delta_i^2 \sigma^2(df_i) + \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \Delta_i \Delta_j \text{cov}(df_i, df_j)}. \]

It assumes that the risk factors are normally distributed and the change in asset value \( dW \) is also normally distributed.

2.11 Simulation Method

The first step in the simulation consists of a particular stochastic model for the behavior of prices. A commonly used model is geometric brownian motion (GBM). The model assumes
that innovations in the asset price are uncorrelated over time and that small movements in prices can be described by

\[ dS_t = \mu_t S_t dt + \sigma_t S_t dz, \]

where \( dz \) is a random variable distributed normally with mean zero and variance \( dt \). It is assumed that the price of the assets does not depend on random variables. It is Brownian in the sense that its variance continuously decreases with the time interval, \( V(dz) = dt \). This rules out processes with sudden jumps.

The variables \( \mu_t \) and \( \sigma_t \) represent the instantaneous drift and volatility at time \( t \). They evolve over time. The variables \( \mu_t \) and \( \sigma_t \) can be functions of past variables, in which case it would be easy to simulate time variation in a GARCH (General Autoregressive Conditional Heteroscedasticity) process.

In practice, the process with infinitesimally small increment \( dt \) can be approximated by discrete moves of size \( \Delta t \). Define \( t \) as the present time, \( T \) as the target time, and \( \tau = T - t \) as the (VAR) time horizon. We can first chop \( \tau \) into \( n \) increments, with \( \Delta t = \tau/n \). (The choice of number of steps should depend on the VAR horizon and the required accuracy. A smaller number of steps will be faster to implement but may not provide a good approximation to the stochastic process.)

Integrating \( ds/s \) over a finite interval, we have approximately

\[ \Delta S_t = S_{t-1}(\mu \Delta t + \sigma \epsilon \sqrt{\Delta t}), \]

where \( \epsilon \) is a standard normal random variable (with mean zero and unit variance). We have \( E(\Delta S/S) = \mu \Delta t \) and \( V(\Delta S/S) = \sigma^2 \Delta t \).

### 2.12 The Bootstrap

An alternative to generating random numbers from a hypothetical distribution is to sample from historical data. Thus we are agnostic about the distribution. For example, suppose
2. BACKGROUND RESEARCH

we observe a series of $M$ returns $R = \Delta S/S$, then $\{R\} = (R_1, \ldots, R_M)$, which can be assumed to be independent identically distributed random variables drawn from an unknown distribution. The historical simulation methods consist of using this series once to generate pseudo returns.

2.13 Computing VAR

The steps for simulating VAR are as follows:

1. Choose a stochastic process and parameters.

2. Generate a pseudosequence of variables $\epsilon_1, \epsilon_2, \ldots, \epsilon_n$, from which prices are computed as $S_{t+1}, S_{t+2}, \ldots, S_{t+n}$.

3. Calculate the value of the asset (or portfolio) $F_{t+n} = F_T$ under this particular sequence of prices at the target horizon.

4. Repeat steps 2 and 3 as many times as necessary, i.e. $n$ times.

This process creates a distribution of values $F^1_T, \ldots, F^n_T$. We can sort the observations and tabulate the expected value $E(F_T)$ and the quantile $Q(F_T, c)$. Then

$$\text{VAR}(c, T) = E(F_T) - Q(F_T, c)$$

2.14 Simulations with Multiple Variables

If the variables are uncorrelated, the randomization can be performed independently for each variable

$$\Delta S_{j,t} = S_{j,t-1}(\mu_j \Delta t + \sigma_j \epsilon_{j,t} \sqrt{\Delta t}),$$

where the $\epsilon$ values are independent across the time or periods and series $j = 1, \ldots, N$. 
Generally, however, variables are correlated. To account for this correlation, we start with a set of independent variables $\eta$, which are then transformed into $\epsilon$ variables setting,

\[\epsilon_1 = \eta_1,\]
\[\epsilon_2 = \eta_1 + (1 - \rho^2)^{1/2}\eta_2,\]

where $\rho$ is the correlation coefficient between the variable $\epsilon_i$. The variable $\epsilon_2$ is unity, since $V(\epsilon_2) = \rho^2 V(\eta_1) + [(1 - \rho^2)^{1/2}]^2 V(\eta_2) = 1$. Then the covariance of the $\epsilon_i$ is

\[\text{cov}(\epsilon_1, \epsilon_2) = \text{cov}[\eta_1, \rho \eta_1 + (1 - \rho^2)^{1/2}\eta_2] = \rho \text{cov}(\eta_1, \eta_2) = \rho.\]

**2.15 The Cholesky Factorization**

We suppose that we have a vector of $N$ values of $\epsilon$, which we would like to display some correlation structure $V(\epsilon) = E(\epsilon\epsilon') = R$. Since the matrix $R$ is a symmetric real matrix, it can be decomposed into its Cholesky factors

\[R = TT',\]

where $T$ is a lower triangular matrix with zeros in the upper right corner.

Then start from an $N$-vector $\eta$, which is composed of independent variables all with unit variances. In other words, $V(\eta) = I$, where $I$ is the identity matrix. Next, construct the variable $\epsilon = T\eta$. Its variance matrix is

\[V(\epsilon) = E(\epsilon\epsilon') = E(T\eta\eta'T') = TE(\eta\eta')T' = TIT = TT' = R.\]

Thus we have confirmed that the values of $\epsilon$ have the desired correlations.

For example, in a two-variable case, the matrix can be decomposed into

\[
\begin{bmatrix}
1 & \rho \\
\rho & 1
\end{bmatrix} = \begin{bmatrix}
\alpha_{11} & 0 \\
\alpha_{12} & \alpha_{22}
\end{bmatrix} \begin{bmatrix}
\alpha_{11} & \alpha_{12} \\
0 & \alpha_{22}
\end{bmatrix} = \begin{bmatrix}
\alpha_{11}^2 + \alpha_{12}^2 & \alpha_{11}\alpha_{12} \\
\alpha_{11}\alpha_{12} & \alpha_{22}^2 + \alpha_{22}^2
\end{bmatrix}.
\]
Because the Cholesky matrix is triangular, the factors can be found by successive substitution by setting

$$\alpha_{11}^2 = 1,$$
$$\alpha_{11}\alpha_{12} = \rho,$$
$$\alpha_{12}^2 + \alpha_{22}^2 = 1,$$

which yields

$$\begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \rho & (1 - \rho^2)^{1/2} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}.$$ 

If $A$ has real entries and is symmetric and positive definite, then $A$ can be decomposed as

$$A = LL^T,$$

where $L$ is a lower triangular matrix with strictly positive diagonal entries. If $A$ is a symmetric positive-definite matrix with real entries, $L$ can be assumed to have real entries as well. Cholesky factorization is mainly used for the numerical solution of linear equations $Ax = b$. If $A$ is positive definite, then we have $A = LL^T$. Then solve $L\eta = b$ for $\eta$ and then find $L^Tx = \eta$ for $x$. Cholesky factorization, in our case, helps generate correlated random variables through a variance-covariance matrix (variance-covariance matrix is positive symmetric).

### 2.16 Number of Independent Factors

The variable $R$ must be a positive-definite matrix for the decomposition to work. Otherwise, there is no way to transform $N$ independent sources of risks into $N$ correlated variables of $\epsilon$. If the matrix is not positive-definite, the Cholesky factorization will not work.

We can use the Singular Value Decomposition (principal component analysis) to reduce the computation without much loss.
2. BACKGROUND RESEARCH

2.17 Singular Value Decomposition

Suppose \( M \) is an \( m \times n \) matrix with real entries. Then there exists a factorization of the form

\[
M = U \Sigma V^T,
\]

where \( U \) is an \( m \times m \) matrix, \( \Sigma \) is an \( m \times n \) diagonal matrix with nonnegative numbers on the diagonal, and \( V^T \) is an \( n \times n \) matrix. Such a factorization is called a Singular Value Decomposition. We have that:

- \( V \) contains a set of orthonormal “input” basis vectors for \( M \).
- \( U \) contains a set of orthonormal “output” basis for \( M \).
- the matrix \( \Sigma \) contains the singular values.

SVD allows us to truncate the matrix into \( \tilde{M} = U_t \Sigma_t V_t^T \), where \( \tilde{M} \) is the closest approximation of rank \( t \) to the matrix \( M \).

2.18 Eigenvalues and Singular Values

Eigenvalues and singular values are similar concepts. Suppose the matrix \( A \) has column vectors centered at zero. Then \( A^T A \) is essentially a covariance matrix, where the entry \( a_{i,j} = \langle A_i, A_j \rangle \) and the covariance between the columns related to item \( i \) and \( j \) is \( \sigma_{i,j} = \frac{a_{i,j}}{n - 1} \). In this case, the eigenvalues of the matrix \( A^T A \) are just the squares of the singular values of \( A \) multiplied by some normalization constant.

2.19 Component VaR, Marginal VaR and Incremental VaR

**Marginal VaR** measures the change in return-VaR resulting from a marginal change in the relative position in instrument \( i \). Hence we have:

\[
M - \text{VAR} \equiv \frac{\partial r_p}{\partial w_i} \text{ for } i \in p.
\]
2. BACKGROUND RESEARCH

**Incremental VaR** measures the change in return-VaR resulting from adding a small additional position in the portfolio. We have

\[ \Delta \text{VAR} \approx w_N \left( \frac{\partial r_p}{\partial w_N} \bigg|_{w_N=0} \right). \]

**Component VaR** denotes the instrument’s i component return-VaR in the context of the overall portfolio. According to the definition of CVaR, we have

\[ r_p = \sum_{i \in p} \text{CVAR}_i. \]

### 2.20 Connecting Component VaR, Marginal VaR and Incremental VaR

We can construct Component VaR as follows:

\[ \text{CVAR} = (\Delta \text{VAR}) \times w_i W = \text{VAR}_i w_i. \]

In this case, we have

\[ \text{CVAR}_1 + \ldots + \text{CVAR}_n = \text{VAR} \sum_{i=1}^n w_i \beta_i = \text{VAR}. \]

The component VaR can also be simplified into

\[ \text{CVAR}_i = \text{VAR} w_i \beta_i = (\alpha \sigma_i w_i W) \rho_i = \text{VAR}_i \rho_i. \]

### 2.21 Monte Carlo Simulation

Monte Carlo methods are a class of computational algorithms that rely on random sampling to compute results. Monte Carlo methods are most suited to calculation by a computer. Monte Carlo methods tend to be used when it is infeasible or impossible to compute an exact result with a deterministic algorithm.
2. BACKGROUND RESEARCH

2.22 R-Squared Value

R-Squared Value is a concept from linear regression. Suppose $y_i$ are observed data and $f_i$ are corresponding model-predicted values. Then we have

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2$$

is the total sum of squares and

$$SS_{\text{reg}} = \sum_i (f_i - \bar{f})^2$$

is the regression sum of squares and

$$SS_{\text{err}} = \sum_i (y_i - f_i)^2$$

is the sum of squared errors. In the previous equations, the variables $\bar{y}$ and $\bar{f}$ are the average of observed data and model predicted data. Then

$$R^2 \equiv 1 - \frac{SS_{\text{err}}}{SS_{\text{tot}}}.$$
3
Research Set-up

3.1 Research Goal

My research goal is to use mathematical tools to identify risks in an equity portfolio. The research here emphasizes the quantification of Value at Risk. In other words, we try to quantify the risk in a portfolio with the Value at Risk measurement.

3.2 Research Methodologies

Instead of using purely theoretical mathematical methodologies, I take a different route in my research. I start by processing data using different techniques and then I try to explain the theories that justify the findings. The research methodology can be summarized in the following steps.

1. Data collection and data processing.

2. Data observation.

3. Make a proposition.

4. Test the proposition.
3. RESEARCH SET-UP

5. Revise the proposition.

6. Justify the proposition.

3.3 Data Collection and Data Processing

The stock data sample is collected from Google Finance and the index data (Dow Jones Industrial Average) are collected from Yahoo Finance. Each data point is a close price of the day. Holidays and weekends, during which there are no trading activities, are excluded from our data sets.

The data range from January 23rd, 1999 to January 20th, 2009. We may test a chosen range of time in this range. This time-frame allows us to test the time-series in both the long term and the short-term.

When choosing stocks for the data sample, I intentionally spread out the sectors of the stocks so that stocks in different sectors can be tested. The stocks in the sample are listed as follows:

The components of the sample are as follows.
MATLAB 2007 is the primary tool for the data analysis. The return data matrix is generated by calculating the log-return using the following formula

$$r_t = \log(P_t) - \log(P_{t-1})$$

for each of the stocks.
4

Stock Price Time Series

“It will fluctuate.” - JP Morgan, when asked what the stock market would do.

4.1 Hypothesis

Stock prices change from time to time. Stock price movement is itself a time series. The stock prices do rise and fall from people’s buying and selling activities. Therefore, there might be some common frequencies at which stock generate abnormal returns. For instance, in some literatures it is found that the stock prices are generally lower in December and higher in January. People’s mentality may also contribute to the mean-return of stock prices in small scales. Therefore, the stock prices or stock returns are not totally random walks. By using time series analysis, some interesting facts about the stock prices may be revealed.

4.2 Model

The model is built upon a sunspot model (please refer to Appendix B for the code). The mathematical tools reveal the relative strength or energy on each of the frequency nodes.
4. STOCK PRICE TIME SERIES

That is to say, we will be able to observe the possible abnormal return/stock price change frequencies.

4.3 Discrete Fourier Transform

The Discrete Fourier Transform (DFT) transforms one function into another. The DFT is called the frequency domain representation of the original function (which is often a function defined on the time domain).

The mathematical definition of DFT is as follows [8].

The sequence of \( N \) complex numbers \( x_0, \ldots, x_{N-1} \) is transformed into the sequence of \( N \) complex numbers \( X_0, \ldots, X_{N-1} \) by the DFT according to the following formula:

\[
X_k = \sum_{n=0}^{N-1} x_n e^{-\frac{2\pi i}{N} kn} \quad k = 0, \ldots, N - 1,
\]

where \( e^{-\frac{2\pi i}{N} kn} \) is a primitive \( N \)’th root of unity.

4.4 Some Experiments

We use MATLAB to examine some of the stock and index time series that we obtained from various data sources. The JP Morgan time series from January 23rd, 1999 to January 20th, 2009 is:
After performing a DFT, we get

\[ \text{JP Morgan Time Series} \]

\[ \text{JP Morgan Stock Price DFT Analysis} \]
4. STOCK PRICE TIME SERIES

The period with the highest frequency is the longest period. The cycle of the whole sequence has more power than any of the the shorter periods. Since the time series is not centered at zero, the DFT analysis is not adequate. A more hopeful approach, however, is instead to run the DFT on the return of the stocks. The return time series for JP Morgan is shown in the following figure:

![JP Morgan Return Time Series](image)

The findings of the DFT are shown in the following figure:

![JP Morgan Return DFT](image)
4. **STOCK PRICE TIME SERIES**

Besides seeing that the energies are concentrated in the lower period, there is no obvious conclusion. The experiment is repeated on a few other time series.

We applied a similar analysis to Dow Jones Industrial Average Index:
ConocoPhillips is an integrated energy company based in US. We again apply DFT to analyze its stock return time series.
Again, all the time series above display very similar properties under DFT. In the following section, we will try to explain the behaviors of the stock return time series.
4. STOCK PRICE TIME SERIES

4.5 Some Observations

Despite the unexciting graphs we get from applying the Discrete Fourier Transform, we can still observe some trends of particular stocks from the graphs. We see 500 days as a period that covers about two years (there are about 250 trading days in a year). JP Morgan has a peak at a period close to 3-year cycle. In the shorter term, a scaled image reveals some facts. With the help of MATLAB’s data cursor, some additional peaks can be located and tentatively explained.

<table>
<thead>
<tr>
<th>Days/Cycle</th>
<th>Energy</th>
<th>Possible Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.96</td>
<td>11.86</td>
<td>Weekly Fluctuation</td>
</tr>
<tr>
<td>22.06</td>
<td>8.709</td>
<td>Monthly Fluctuation</td>
</tr>
<tr>
<td>38.11</td>
<td>6.258</td>
<td>N/A</td>
</tr>
<tr>
<td>76.21</td>
<td>4.384</td>
<td>Quarterly Fluctuation</td>
</tr>
<tr>
<td>114.3</td>
<td>4.747</td>
<td>N/A</td>
</tr>
</tbody>
</table>

The Dow Jones Industrial shows a peak on approximately a two-year and one-year cycle. The 83.77 days/cycle period has relatively strong energy perhaps related to the quarterly changes. The 228.5 days/cycle period perhaps corresponds to the yearly fluctuation.

ConocoPhillips also has some interesting peaks.

<table>
<thead>
<tr>
<th>Days/Cycle</th>
<th>Energy</th>
<th>Possible Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.03</td>
<td>8.953</td>
<td>Weekly Fluctuation</td>
</tr>
<tr>
<td>24.9</td>
<td>4.434</td>
<td>Monthly Fluctuation</td>
</tr>
<tr>
<td>73.97</td>
<td>2.877</td>
<td>Quarterly Fluctuation</td>
</tr>
<tr>
<td>86.72</td>
<td>3.33</td>
<td>N/A</td>
</tr>
<tr>
<td>251.5</td>
<td>1.829</td>
<td>Semi-Annual Fluctuation</td>
</tr>
<tr>
<td>838.3</td>
<td>0.8711</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Before getting too optimistic with our findings, let’s examine a random sequence. The sequence is generated by some simple MatLab codes (please refer to Appendix A). The sequence is a random walk with fixed variance and zero mean (here 0.01 is specified).
This pseudo return time series is clearly distinct from any stock return time series that was shown previously. The returns are more bounded and compact in the pseudo return series and volatility looks uniform throughout the series. The DFT Analysis gives rise to the following graph:

The pseudo return time series is not that different from the actual stock return! Also, if we want to, we can assign peaks meanings.
4. **STOCK PRICE TIME SERIES**

The only observable difference between this and the previous DFT analysis charts is that the tail of this DFT chart is much higher than the three analyses on the real historical data, suggesting more energy on the longer periods. Whether this observation is true to most of the stocks is yet to be concluded.

However, we can see that the cyclical behavior on DFT analysis is mainly a result of random walk behavior. At least, it is safe to conclude that there is no strong indication of stocks’ cyclic with fixed frequency.
5

What Do Singular and Eigenvalues Tell Us?

5.1 Motivation

Principal Component Analysis (PCA) can help people indentify the factor that causes the most impact [12]. PCA plays a very important role in risk management techniques. The technique can be used to reduce the dimension of a covariance matrix. Besides its application in Monte Carlo simulation, Principle Component Analysis can also reduce the computation time of the analytic method. However, the modified matrix using PCA has a downward bias on the VaR. Whereas people hope to facilitate the computation, underestimating risks can cause big problems to risk management. Therefore, it is important to study the nature of singular and eigenvalues of a matrix, which are central concepts in PCA.

5.2 Eigenvalue vs Singular Value

Eigenvalues and singular values are the same concept when we treat a covariance matrix, since the eigen-decomposition already yields a orthogonal basis. In fact, the singular values are the square roots of the eigenvalues of the covariance times a certain constant. However,
when we study the data matrix, singular value decomposition provides a more convenient
way to look at the matrix, since the data matrices are usually not square matrices.

5.3 Hypothesis

The singular value indicates the magnitude of variation in a certain direction. The first hypothesis is that the more concentrated the portfolio, the greater the size of the corresponding eigenvalues, since more risks will be concentrated on the principal direction.

The second hypothesis is that the relative importance (percentage of the sum of eigenvalues) of the dominant eigenvalue will likely decrease as the portfolio diversifies. Conversely, the relative importance of the least eigenvalue is likely to increase.

The relative magnitude of a singular value is defined as

\[ M_{\text{relative}} = \frac{\sigma_i^2}{\sum \sigma_i^2} \]

5.4 PCA Experiments

Several portfolio configurations have been carried out to test the properties.

The first portfolio stock selection includes five financial stocks. They are JP Morgan, Bank of America, Citigroup, Morgan Stanley and Wells Fargo. The first diversification comes from replacing Wells Fargo by News Corp. The second diversified portfolio consists of JP Morgan, Bank of America, Morgan Stanley, News Corp and Microsoft. The third diversified portfolio consists of Coca-Cola, Bank of America, Citigroup, Morgan Stanley and News Corp. The fourth consists of JP Morgan, WalMart, Morgan Stanley, News Corp and Microsoft. Finally, the last set of stocks has JP Morgan, Bank of America, ConocoPhillips, Coca-Cola and GlaxoSmithKline. The start date of our testing is January
5. WHAT DO SINGULAR AND EIGENVALUES TELL US?

23rd, 1999. The length of the data that we use is 251 trading days, or approximately one year.

This experiment shows that our first hypothesis is questionable, however, the second hypothesis may be valid in some sense. Therefore, additional experiments are carried out.

This time, a different set of portfolios are being tested and, to show the consistency of the theory, the set of stock selections are tested under different dates.

<table>
<thead>
<tr>
<th>Stock Pool No.</th>
<th>Company Tickers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>COP XOM CHK XTO BHP</td>
</tr>
<tr>
<td>2</td>
<td>COP XOM CHK XTO PFE</td>
</tr>
<tr>
<td>3</td>
<td>COP BAC CHK XTO PFE</td>
</tr>
<tr>
<td>4</td>
<td>COP BAC JPM XTO PFE</td>
</tr>
</tbody>
</table>
5. **WHAT DO SINGULAR AND EIGENVALUES TELL US?**

The start dates of the time series are 300 trading days, 600 trading days and 1200 trading days after Jan 20, 1999 in the three separate trials. The undiversified portfolio includes 4 oil production/refinery firms and one basic material (BHP Bilton), which has a lot of exposure to oil price. We gradually diversify the portfolio with financial stocks and Pfizer, a major drug production company.

Experiment 1:

![The Absolute Magnitude of Singular Values](image)

![The Relative Magnitude of Singular Values](image)

Experiment 2:
5. WHAT DO SINGULAR AND EIGENVALUES TELL US?

### Experiment 3:

**The Absolute Magnitude of Singular Values**

- **COP XOM CHK XTO BHP**
- **COP XOM CHK XTO PFE**
- **COP BAC CHK XTO PFE**
- **COP BAC JPM XTO PFE**

**The Relative Magnitude of Singular Values**

- **COP XOM CHK XTO BHP**
- **COP XOM CHK XTO PFE**
- **COP BAC CHK XTO PFE**
- **COP BAC JPM XTO PFE**
5. WHAT DO SINGULAR AND EIGENVALUES TELL US?

These experiments in some sense validated our hypothesis, as we observe from the tables, that the leading singular values decline as we diversify the portfolio compositions. However, some anomalies also appear. For example, the diversified portfolio has more relative magnitude on the first singular value than the second and third singular value in the second experiment. That is to say, after diversifying the portfolio, the variation on the principal direction does not decrease versus the other directions.

5.5 Stock Clustering by Singular Vectors

Considering the vectoral nature of our time series presentation, I reflected on my early work on Data Clustering. The question now can be viewed as: “can we use the singular vectors to cluster stocks?”, which means, stocks with similar movement are set apart from the rest in the group. For the specific technique on clustering with SVD, please refer to [11].

Our experiment starts with oil/energy sector, since oil/energy stocks have very strong comovements due to the significant influence of oil prices. The five stocks are XOM, XTO, CHK, COP and BHP.
5. WHAT DO SINGULAR AND EIGENVALUES TELL US?

The test results are as follows, indicating each stock’s corresponding value on the first singular vector:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>COP</td>
<td>-0.3243</td>
</tr>
<tr>
<td>XOM</td>
<td>-0.2607</td>
</tr>
<tr>
<td>CHK</td>
<td>-0.6752</td>
</tr>
<tr>
<td>XTO</td>
<td>-0.5108</td>
</tr>
<tr>
<td>BHP</td>
<td>-0.3318</td>
</tr>
</tbody>
</table>

As we can observe from the table, all singular values are negative. Then some of the stocks in the set are replaced by stocks from other sectors. For instance, we when we replace COP by a financial stock, JPM, the result becomes

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>COP</td>
<td>0.3290</td>
</tr>
<tr>
<td>XOM</td>
<td>0.2621</td>
</tr>
<tr>
<td>CHK</td>
<td>0.7260</td>
</tr>
<tr>
<td>XTO</td>
<td>0.5431</td>
</tr>
<tr>
<td>GSK</td>
<td>-0.0317</td>
</tr>
</tbody>
</table>

The value corresponding GSK becomes negative while the other values are positive, distinguishing GSK from the rest of the stocks in the set. Replacing one more stock in the portfolio, the result is presented as follows:

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>COP</td>
<td>0.3623</td>
</tr>
<tr>
<td>XOM</td>
<td>0.2870</td>
</tr>
<tr>
<td>CHK</td>
<td>0.8842</td>
</tr>
<tr>
<td>XTO</td>
<td>0.5431</td>
</tr>
<tr>
<td>PFE</td>
<td>-0.0531</td>
</tr>
<tr>
<td>GSK</td>
<td>-0.0415</td>
</tr>
</tbody>
</table>

Now we look into a different sector. The financial sector stocks usually have a lot of comovements. Therefore, we take a set of five financial stocks and apply SVD.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPM</td>
<td>0.4464</td>
</tr>
<tr>
<td>BAC</td>
<td>0.0017</td>
</tr>
<tr>
<td>C</td>
<td>0.5294</td>
</tr>
<tr>
<td>MS</td>
<td>0.6312</td>
</tr>
<tr>
<td>WFC</td>
<td>0.3495</td>
</tr>
</tbody>
</table>

As we expected, since all of the stocks are from the same sector, the signs of the corresponding values on the first singular vector are the same. Meanwhile, we notice that the
second value, corresponding to BAC is comparatively small. Therefore, the second singular value is examined.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPM</td>
<td>-0.0348</td>
</tr>
<tr>
<td>BAC</td>
<td>0.9977</td>
</tr>
<tr>
<td>C</td>
<td>-0.0299</td>
</tr>
<tr>
<td>MS</td>
<td>0.0491</td>
</tr>
<tr>
<td>WFC</td>
<td>-0.0038</td>
</tr>
</tbody>
</table>

This signifies that BAC’s volatility is concentrated more on a much different direction.
6
Linear Projection of Stocks

6.1 Projection Test

Two sets of four-stock combinations are picked. Set one consists of four stocks from financial sectors, namely, JP Morgan (JPM), Bank of America (BAC), Citigroup (C) and Morgan Stanley (MS). The other set consists of stocks from mixed sectors MasterCard (M), News Corp (NWS), Coca-Cola (KO) and Microsoft (MSFT). Then we project Wells Fargo (WFC) return time series onto the two sets of stocks separately. Finally, we run a linear regression of Wells Fargo return time series against its projections onto the two sets of stocks. The regression equation is \( f(x) = p_1 x + p_2 \). The results are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Test 1</th>
<th>Test 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>(0.9653, 1.035)</td>
<td>(0.9252, 1.075)</td>
</tr>
<tr>
<td>( p_2 )</td>
<td>0.0002603(-0.0003428, 0.0008634)</td>
<td>0.0001612(-0.0006439, 0.0009663)</td>
</tr>
<tr>
<td>R-square</td>
<td>0.5592</td>
<td>0.2146</td>
</tr>
</tbody>
</table>

The numbers in the parenthesis are the 95% confidence intervals.

6.2 How to Interpret?

In both cases, the linear regression yields a line that has slope of 1 and is very close to going through the origin (0, 0). Thus, there is no fundamental change in the time series. However,
the R-square, a variable that explains the total variation explained by the model, dropped from the first test to the second test. Thus, the projection projected onto the financial sector is closer to the original series than the one projected onto the mixed sector. Of course, this result is more or less expected.
The nature of stock return time series is very intriguing. One direction that particularly interests me and is, I think, deserving of further research is using signal processing to analyze the stock return time series. The inconclusiveness of the analysis in this project could well be a result of the basic nature of the signal processing tool implemented. Some more sophisticated techniques may tell more interesting stories about the data.

Singular Value Decomposition may provide some previously unidentified relationship among the stocks in a portfolio. Also, by identifying the greatest direction of variation, we may be able to hedge the portfolio better by adding a component that counteracts the risk on that direction.
Appendix A

Pseudo Stock Prices Generator

%This function generates random stock price time series.
%mu specifies the average return over the time
%sigma specifies the standard deviation of the time series
%dt specifies the time interval of each step of simulation
%h gives the time horizon (number of days) of the simulation
%s0 is the start price (price of day 1)

function sn=gen_s_i(mu,sigma,dt,h,s0)

sn=zeros(1,h); %create a vector that stores the time series.
sn(1)=s0; %set the day 1 price to s0
epsilon=randn(h-1); %create the drift of each day after day 1 from normal distribution
for i=1:h-1
    %calculate the prices of each day in the time series
    sn(i+1)=sn(i).*(1+mu*dt+sigma.*dt.^((1/2)).*epsilon(i));
end
%Please refer to MatLab Fourier Analysis Demo

Y = fft(timeseries);
N = length(Y);
Y(1) = [];
power = abs(Y(1:N/2)).^2;
nyquist = 1/2;
freq = (1:N/2)/(N/2)*nyquist;
plot(freq,power), grid on %plot power against frequency
period = 1./freq;
plot(period,power), grid on %plot power against period
Bibliography