1. Introduction

Financial activities involve risk. One well-known risk measure is Value-at-Risk (VaR), which is a measure related to the loss distribution and represents the predicted maximum loss with a specified probability level (such as 95%, 99%) over a certain period of time. VaR is widely used by people in the financial industry. However, VaR lacks subadditivity and may be difficult to compute and optimize. In addition, VaR pays no attention to the magnitude of losses beyond the VaR value. One well-known modification of VaR is the Conditional Value-at-Risk (CVaR). Here, for a given confidence level $\beta$, the VaR$_\beta$ associated with portfolio $x$ is given by 1.1.

\[
\text{VaR}_\beta = \alpha_\beta = \min \{ \alpha \in R \mid \int_{f(x,y)\leq \alpha} p(y)dy \geq \beta \}
\]

We define the $\beta$-CVaR associated with portfolio $x$ as:

\[
\text{CVaR}_\beta = \frac{1}{1 - \beta} \int_{f(x,y)\leq \alpha} f(x,y)p(y)dy
\]

Note that $\text{CVaR}_\beta(x) \geq \text{VaR}_\beta(x)$, i.e., the CVaR of a portfolio is always at least big as its VaR. Figure (1) shows an example of VaR and CVaR.

Modern portfolio theory (MPT) attempts to maximize investment return while minimizing risk by strategically selecting different assets to build an investment portfolio. More technically, MPT models a portfolio as a weighted combination of assets so that the return of a portfolio
is the weighted combination of the assets’ returns. By combining different assets MPT seeks to reduce the total risk of the portfolio. In many cases, choosing a particular portfolio optimization strategy results in making a priori assumptions regarding underlying asset dynamics. Due to features such as nonstationarity, model misspecification, and a lack of robustness, mean variance optimization as originally formulated is mainly an academic tool.

For a fixed period investment (disregarding mark-to-market or leverage concerns), underlying dynamics are secondary to the final joint density function of the assets involved. Using conditional value-at-risk (CVaR) as our objective function, we construct CVaR optimal portfolios under various assumptions for the terminal joint distribution.

2. Problem Description

Let \( f(w,y) \) denote a loss function as:

\[
(2.1) \quad f(w, y) = - \sum_{i=1}^{n} w_i y_{is},
\]

the weighted total loss of the portfolio under a certain scenario \( s \). By definition, \( \beta - VaR \) represents the predicted maximum loss with a specified probability level over a certain period of time (1.1). And again, the conditional value at risk, \( \beta - CVaR \), is given by (1.2). Since the definition of CVaR involves the VaR function explicitly, it is difficult to work with and optimize this function as originally stated.
Therefore, we introduce another function 2.2,

\[ F_\beta(w, \alpha) = \alpha + \frac{1}{1 - \beta} \int (f(w, y) - \alpha)_+ p(y) dy. \]

Rockafeller shows that \( \min (F) = \beta - CVaR \).

Sampling the probability distribution of \( y \) according to its density \( p(y) \), we get the approximation of \( F \) as:

\[ \tilde{F}_\beta(w, \alpha) = \alpha + \frac{1}{S(1 - \beta)} \sum_{s=1}^{S} (f(w, y_s) - \alpha)_+ \]

The objective of our portfolio optimization problem is to minimize 2.3, subject to constraints:

\[ \sum_{i=1}^{n} w_i = 1 \]  
\[ 0 \leq w_i \leq 0.05 \]  
\[ \mu' \cdot w \geq r \]

where \( r = (1.01)^{1/21} - 1 \), and \( \mu = E[y] \). That is, we are requiring a daily return that yields at least a one month return of one percent.

We investigate the problem using two approaches. One is to change the objective function to a linear function with two additional constraints, the other is to use a Minorization-Maximization Algorithm.

2.1. Linear Programming Approach.

In order to solve this problem, we introduce an auxiliary variable \( z = (z_1, z_2, \ldots, z_s) \) to replace \( (f(w, y_s) - \alpha)_+ \) in (2.3), this is achieved by imposing the constraints (2.7).

\[ z_s \geq f(w, y) - \alpha \]  
\[ z_s \geq 0 \]

In this way, we modify the optimization problem to (2.9), subject to constrains (2.4) and (2.7). We use the Matlab function linprog to solve the above problem.

\[ \min_{(w, z, y)} \alpha + \frac{1}{S(1 - \beta)} \sum_{s=1}^{S} z_s \]

2.2. MM Approach.

2.2.1. The problem.
As an alternative to the Linear Programming method, we consider a derivative-based numerical approach.

\[
\text{CVaR} = \min \alpha + \frac{1}{(1 - \beta)S} \sum_{s=1}^{S} \left[ - \sum_{j=1}^{N} y_{js} w_j - \alpha \right]^+
\]

subject to \( w \geq 0 \) and \( \sum_{i=1}^{n} w_i = 1 \)

2.2.2. MM algorithm approach.

MM algorithm, which is short for Minorization-Maximization or Majorization-Minimization, is a generalization of the well-known EM algorithm method. The algorithm consists of the following steps:

1. \( f(\theta) \) is the target function we want to maximize, and \( \theta^{(m)} \) a fixed value of \( \theta \).
2. \( g(\theta|\theta^{(m)}) \) is a real-valued function which minorizes \( f(\theta) \) at point \( \theta^{(m)} \), i.e.
   \[
g(\theta|\theta^{(m)}) \leq f(\theta) \quad \text{for all} \quad \theta
\]
   \[
g(\theta^{(m)}|\theta^{(m)}) = f(\theta^{(m)})
\]
3. Denote \( (\theta^{(m+1)}) \) to be the maximizer of \( g(\theta|\theta^{(m)}) \),
   \[
   \theta^{(m+1)} = \arg \max_{\theta} g(\theta|\theta^{(m)})
   \]

The g-function for CVaR computation is:

\[
g(w, \alpha|w^{(m)}, \alpha^{(m)}) = \alpha + \frac{1}{(1 - \beta)S} \sum_{s=1}^{S} h \left( - \sum_{j=1}^{N} y_{js} w_j - \alpha \right) - \sum_{j=1}^{N} y_{js} w_j^{(m)} - \alpha^{(m)}
\]

where \( h(x|x^{(m)}) = \frac{1}{4|x^{(m)}|} x^2 + \frac{1}{2} x + \frac{1}{4} |x^{(m)}| \)

We note the following drawback: the factor \( \frac{1}{|x^{(m)}|} \) leads to instability when \( |x^{(m)}| \) approaches 0.

2.2.3. Revised MM algorithm.

In Alexander’s paper, they proposed an quadratic function \( \rho \) to approximate the piecewise linear function.

\[
\tilde{\text{CVaR}} = \alpha + \frac{1}{(1 - \beta)S} \sum_{s=1}^{S} \rho_t \left( - \sum_{j=1}^{N} y_{js} w_j - \alpha \right)
\]
Given a resolution parameter $\epsilon > 0$,

$$
\rho_\epsilon(z) = \begin{cases} 
  z & \text{if } z \geq \epsilon; \\
  \frac{1}{4\epsilon} z^2 + \frac{1}{2} \frac{1}{4} \epsilon & \text{if } z \leq 0.
\end{cases}
$$

We have a continuous approximation with continuous derivative. However, the Hessian matrix is not continuous, so only a gradient-based first-order method be implemented.

![Figure 2. Quadratic Piecewise Continuous approximate function](image)

2.2.4. B-spline function approach.

We proposed a similar but second-order continuous approach to Alexander’s paper by using B-spline function to approximate the piecewise linear function,

$$
S_i(z) = \frac{1}{6} [z^3, z^2, z, 1] M [y_{i-1}, y_i, y_{i+1}, y_{i+2}] \quad \text{where } z \in [0, 1]
$$

where $M = \begin{pmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{pmatrix}$

While we were not able to implement this approach, we hope to see increased stability and speed from this smooth approximation.

3. Data and Distribution

We use the S&P 500 index as a reference to evaluate the optimized stock portfolio under study. For comparison, monthly returns of S&P 500 from January 2002 to December 2009 are collected. We focus on a
sample of 125 stocks with a market capitalization over $10 billion, using daily close prices from January 3, 2000 to December 31, 2009. Our data takes into account mergers and acquisitions, bankruptcy, dividends, and other financial operations.

3.1. Missing data.
Very often some data is missing from the time series of observations. Simply ignoring missing data will generally lead to biased estimates. Imputation strategies are the most widely used methods. On the one hand, one may interpolate missing data. However, a more reliable imputation is to assume a statistical model for data set and estimate parameters of that model. Missing data can be randomly drawn from the model. We carried out an approach named Amelia II developed by professor Gary King from Harvard University (see http://gking.harvard.edu/amelia/ for detail). The imputation model in Amelia II assumes that the complete data (that is, both observed and unobserved) is multivariate normal.

Choosing an appropriate distribution for stock returns is an important issue of this project. The two distributions most commonly used in the analysis of financial asset returns and prices are normal distribution and its cousin, log normal distribution. As we know, the normal distribution is often not a good approximation for returns. We use a student-t distribution; taking advantage of fatter tails.

As we know, the normal distribution has a bell-shaped curve described by two parameters $\mu$ for mean and $\sigma$ for variance. To investigate returns...
of multiple stocks, we begin by using a multivariate normal distribution. The multivariate normal distribution is a generalization of the univariate normal distribution to higher dimensions. As in the univariate normal distribution, two parameters describe the MVN: they are the mean vector $\mu$ and covariance matrix $\Sigma$. Both $\mu$ and $\Sigma$ can be easily estimated from our dataset once missing data is imputed.

![Graph showing comparison between standard normal distribution and Student-t distribution](image)

**Figure 4.** Compare Standard Normal Distribution with Student-t Distribution

3.2.1. **Drawbacks of Normal Distribution.**

Figure (3.2.1) shows how a normal distribution has a thinner tail and a low degree of peakedness. We note that stock market returns are interspersed with outliers that reflect news, events or information released by firms, thereby increasing the kurtosis of the return distribution. We randomly select 4 stocks, fitted both normal distribution and Student-t distribution. Figure (3.2.1) shows fitted distributions for these stock returns. We can easily see that the normal distribution does not fit the stock returns well. However, the Student-t distribution is able to capture the fatter tails observed.
3.3. Multivariate Student-t Distribution.

The overall shape of the probability density function of the t-distribution resembles the bell shape of a normally distributed variable, except that it is a bit lower and wider. As the number of degrees of freedom grows, the t-distribution approaches the normal distribution.

3.3.1. Choice of Degree of Freedom.

We used statistical software R to fit multivariate student t distribution for a trailing window of two years stock returns, and obtained a vector of degrees of freedom for fitted multivariate Student-t distribution. We observed the fitted degrees of freedom varied from 5 to 19 over our historical period.

4. Procedures

Step 1: Two multivariate distributions (Normal and Student’s t) are chosen for portfolio returns. The parameters (such as location and dispersion) are estimated from 2-year long historical returns of the portfolio under study and 2000 scenarios are generated to optimize portfolio.

Step 2: The Solvers (LP and MM) take the scenarios from the Step 1 and search for the optimal weight for each stock in the portfolio. We compare these methods in term of computing time.

Step 3: Optimal weights are computed at the end of every calendar month. We evaluate actual performance ex-post, and rebalance the full portfolio at each month end.
Table 1. Performance statistics for the CVaR optimal portfolios compared with S&P 500 index.

<table>
<thead>
<tr>
<th></th>
<th>S&amp;P500</th>
<th>LPNormal</th>
<th>LPStudentt</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>-23.37 %</td>
<td>-14.75 %</td>
<td>-13.38 %</td>
</tr>
<tr>
<td>2003</td>
<td>26.38 %</td>
<td>22.12 %</td>
<td>21.56 %</td>
</tr>
<tr>
<td>2004</td>
<td>8.99 %</td>
<td>10.65 %</td>
<td>10.92 %</td>
</tr>
<tr>
<td>2005</td>
<td>3.00 %</td>
<td>0.33 %</td>
<td>1.03 %</td>
</tr>
<tr>
<td>2006</td>
<td>13.62 %</td>
<td>20.86 %</td>
<td>22.30 %</td>
</tr>
<tr>
<td>2007</td>
<td>3.53 %</td>
<td>4.64 %</td>
<td>4.78 %</td>
</tr>
<tr>
<td>2008</td>
<td>-38.49 %</td>
<td>-16.85 %</td>
<td>-17.77 %</td>
</tr>
<tr>
<td>2009</td>
<td>23.45 %</td>
<td>29.51 %</td>
<td>28.78 %</td>
</tr>
<tr>
<td>Total Return</td>
<td>-2.87 %</td>
<td>57.38 %</td>
<td>60.12 %</td>
</tr>
<tr>
<td>Annualized Return</td>
<td>-0.36 %</td>
<td>5.83 %</td>
<td>6.06 %</td>
</tr>
<tr>
<td>Average Monthly Return</td>
<td>0.07 %</td>
<td>0.53 %</td>
<td>0.55 %</td>
</tr>
<tr>
<td>Max Drawdown</td>
<td>-16.94 %</td>
<td>-11.30 %</td>
<td>-11.57 %</td>
</tr>
<tr>
<td>Max Gain</td>
<td>9.39 %</td>
<td>7.50 %</td>
<td>7.25 %</td>
</tr>
<tr>
<td>Annualized Return Volatility</td>
<td>15.60 %</td>
<td>11.84 %</td>
<td>11.86 %</td>
</tr>
<tr>
<td>Annualized Sharpe ratio</td>
<td>-0.02</td>
<td>0.49</td>
<td>0.51</td>
</tr>
<tr>
<td>β</td>
<td>1.00</td>
<td>0.24</td>
<td>0.66</td>
</tr>
<tr>
<td>α</td>
<td>0.00</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

5. Result and Evaluation

The performance statistics of the CVaR optimal portfolio from 2002/1 to 2009/12 are shown in Table 1. We use the S&P 500 index as a benchmark for comparison. In general, the CVaR optimal portfolios perform better than S&P 500 index with higher annualized return, higher average monthly return, lower max drawdown, smaller annualized volatility and higher Sharpe ratio. Moreover, the CVaR optimal portfolios have a small positive $\alpha$ and low $\beta$, indicating that they have small independent-of-market risk. It should be noted that the multivariate normal distribution and the multivariate student t distribution used to model the portfolio returns yield very similar results. Most of absolute differences are within 0.1%. This can probably be explained by the fact that the missing data imputation strategy currently used assumes a normal distribution for the complete data set. We expect that using Student-t distribution to fill missing data may make the difference more pronounced.

Figure 6 shows the performance of our optimized portfolio comparing S&P 500 index, giving the simulation of CVaR. While CVaR gives an accurate description of our market risk, the optimized portfolio with
minimized CVaR successfully hedged some of the risk and shows a better total return than S&P 500.

We also compared the computing efficiency of the LP and MM method used in the present work. Most of time, the MM method is faster than the LP method. However, a careful analysis (e.g. fixed accuracy) may be needed to make a conclusion.

Since we are using numerical method to solve the CVaR minimization problem, it is necessary to check the convergence of the LP and MM methods as the number of scenarios increases. For the LP method, CVaR increases initially with increase of scenarios and levels off when a large number of scenarios are generated. For the MM method, however, no clear convergence is shown. A further study should be carried out to clarify this issue.

6. DISCUSSION OF FUTURE WORK

It is not unusual to see that the correlations of stocks are close to zero for large positive shocks and close to one for large negative shocks.
The financial market behaves differently according to the shocks acting on it. In other words, the financial market has different states. In this sense, a mixture of distributions (multivariate normal or Student-t), one for each state, may be appropriate, and may be expected to describe the market more accurately. Another improvement we may make is on missing data imputation. The Amelia II method used in the present work assumes that the complete data are multivariate normal. As we discussed above, however, this is likely not appropriate for stock returns. Other multivariate distributions such as multivariate Student-t should be used to impute these data.

Acknowledgement 1. We would like to express our gratitude to all those who gave us the possibility to complete this project. We want to thank the Institute of Mathematics and its Application. We are also deeply indebted to our Mentor, Christopher Bemis, whose help, stimulating suggestions and encouragement helped us in all the time of research for this project.
LIST OF FIGURES

1 VaR and CVaR 2
2 Quadratic Piecewise Continuous approximate function 5
3 Cubic B-spline approximation 6
4 Compare Standard Normal Distribution with Student-t Distribution 7
5 Compare Distributions for Single Stocks 8
6 S&P 500 Monthly Return, Expected Return from Calculated Weight, CVaR from 2002.1 to 2009.12 10

LIST OF TABLES

1 Performance statistics for the CVaR optimal portfolios compared with S&P 500 index. 9
2 Computing efficiency of the LP and MM methods in term of time (in seconds) taken to optimize the portfolio 10