Diversification in the Managed Futures Universe

Derek G. Nokes, CUNY

Friday, May 22, 2015

Contents

Abstract

- outline thesis
- summarize findings / outline arguments

Introduction and Motivation

In portfolio allocation applications the objective is to maximize investors' future wealth by determining how to allocate capital among a set of available investments in such a way as to maximize *compound* growth subject to a set of constraints. Maximizing wealth requires that we take advantage of the powerful positive effects of compounding. When we reinvest, the magnitude of investment returns and the variability of those returns make *equal* contributions to compounded total return. Reducing the variability of returns thus has as much impact on total return as the magnitude of returns. The variability of portfolio return is a function of the co-variability of investment component returns. If the component returns tend to move together, the magnitude of fluctuations in the monthly value of the portfolio is higher than if the components move in different directions. A portfolio comprised of components with diversified returns will achieve higher compound growth than a portfolio with less diversified components holding average component returns constant. In the simplest terms, portfolio allocation is primarily about selecting sets of investments with *future* positive average returns and low co-variability.

As the size of a portfolio increases, the number of inter-relationships between components explodes. It becomes increasingly difficult to understand the drivers of portfolio return as the number of components rises because the number of independent parameters in a covariance matrix grows with the square of the number of investments. Grouping investments that tend to move together and focusing on trying to find groups that are independent is one common way to reduce portfolio variability. This can be accomplished through the use of *factor models*.

In this paper we focus on the application of a statistical factor models to the investment universe of Commodity Trading Advisor (CTA) programs. Our objective is to model the

Portfolio Return & Its Variability

In this section, we introduce definitions for portfolio return and the variability that will be used in the allocation application developed in later sections of the paper.

Portfolio Return

Portfolio return is obviously a function of the weights and the returns of investment components in the portfolio. We define the portfolio return for I component investments for the month m given the monthly returns and portfolio weights for each component investment i as:

$$r_{P,m} = \sum_{i=1}^{I} (r_{i,m} w_{i,m})$$

The holdig period return (HPR) for the portfolio is one plus the portfolio return for the month m:

$$HPR_{P,m} = 1 + \sum_{i=1}^{I} (r_{i,m} w_{i,m}) = 1 + r_{P,m}$$

The holding period return is the factor by which we mulitply the starting value of the portfolio to get the ending value of a portfolio, given the monthly returns and weights of each component investment.

Similarly, we define the terminal wealth relative (TWR) as the factor by which we multiply the starting value of the portfolio to get the ending value of the portfolio given the return streams and weights for a sequence of M months:

$$TWR_{P,M} = \prod_{1=m}^{M} \left(1 + \left(\sum_{i=1}^{I} (r_{i,m} w_{i,m}) \right) \right) = \prod_{1=m}^{M} HPR_{P,m}$$

We define the portfolio compounded return for the interval M as the portfolio terminal wealth relative minue one$:

$$r_{P,M} = \left(\prod_{1=m}^{M} \left(1 + \left(\sum_{i=1}^{I} (r_{i,m} w_{i,m}) \right) \right) \right) - 1 = \left(\prod_{1=m}^{M} (1 + r_{P,m}) \right) - 1 = \left(\prod_{1=m}^{M} HPR_{P,m} \right) - 1 = TWR_{P,M} - 1$$

Portfolio Variability

We define the standard deviation of the portfolio as:

$$\sigma_{P,M} =$$

Parametric VaR

We can estimate the value-at-risk (VaR) using a parameteric approach by assuming that portfolio returns are drawn from an indepedent and identically distributed normal random variable:

$$VaR = -\alpha_{CL} \sigma_{P,M}$$

Where

α_{CL} is the critical value at the confidence level CL

$\sigma_{P,M}$ is the standard deviation of the portfolio returns over the time interval M

This simple parameteric model can be

Too Many Moving Parts

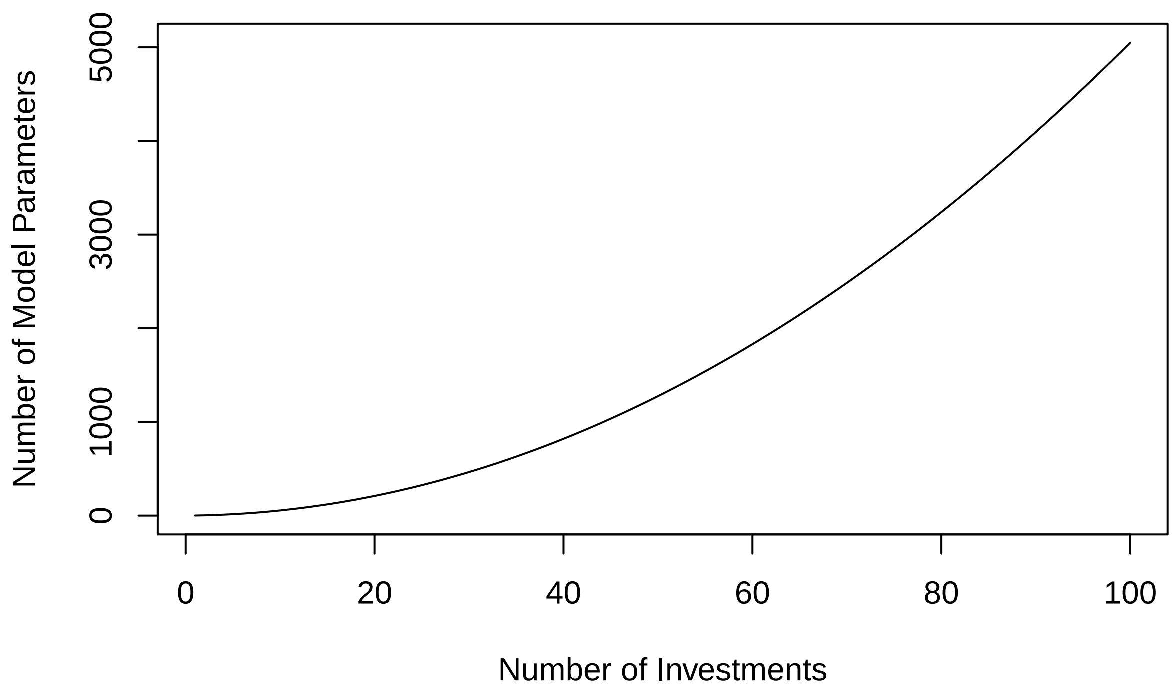
The number of independent parameters in a covariance matrix grows with the number of investments according to the following function:

$$P = \frac{I (I + 1)}{2}$$

Where

P is the number of independent paramters in the covariance matrix

I is the number of distinct investments



Using any

The number of independent parameters to be estimated in the covariance matrix grows with the square of the number of investments, while the number of data points available to estimate a covariance matrix grows only linearly with the number of investments. In other words, the larger the portfolio, the more historical data we need typically need to estimate the covariance matrix reliably. This is particularly problematic when our interest is in the temporal evolution of the relationships between investments in the portfolio.

- difficulty in understanding a portfolio as the number of components increase
- research problem / question
- data science workflow / approach to answer the question

– show the picture of the workflow

- outline each section [Introduction & Motivation, Theory, Data, Applications in R]

– outline each subsection

Data

- quick overview of the data collected
- raw data

– data has not been normalized; not transaction-oriented

– storage has been set-up for one time analysis

- preprocessed data

– although return data quality appears high, the quality of data pertaining to manager and program information is much lower.

Raw Data Extraction, Transformation, and Loading (ETL)

-outline the data (see data dictionary in Appendix A)

– steps taken

– code for the data extraction, exploration, and cleaning etc

Data Exploration

- explore the manager and program information
- explore the monthly return data

– outlier identification

– mean returns vs standard deviation of returns for each period

Data Cleaning

Theory

Modeling Context

In the previous section, we provided an overview of the process used to obtain the monthly returns for all distinct CTA programs available in the Altegris managed futures database.

In this section, we provide a brief overview the theoretical underpinnings of the modeling approach employed in our application [outlined in section blaw blaw].

Standardized Returns

Standardization rescales a variable while preserving its order.

We denote the monthly return of the i^{th} investment for the m^{th} month as $r_{i,m}$ and define the standardized return as:

$$\hat{r}_{i,m} = \frac{(r_{i,m} - \bar{r}_{i,M})}{\sigma(r_{i,M})}$$

Where

$\hat{r}_{i,m}$ is the standardized return of the i^{th} investment for the m^{th} month using data over the time interval M

$r_{i,m}$ is the observed return of the i^{th} investment for the m^{th} month

$\bar{r}_{i,M} = \frac{1}{M}\sum_{m=1}^{M}(\hat{r}_m)$ is the mean of the return stream of the i^{th} investment over the time interval M

$\sigma(r_{i,M}) =$ is the standard deviation of the returns for the i^{th} investment over the time interval M

Using a little linear algebra we can standardize the return with the following operation:

Correlations

We represent the standardized returns as an I x M matrix $\hat{R}$ with an empirical correlation matrix C defined as:

$$C = \frac{1}{M}\hat{R}\hat{R}^T$$

Where

T denotes the matrix transform

The correlation matrix (C) of returns $(\hat{R})$ and the covariance matrix $(\Sigma_{\hat{R}})$ of standardized returns $(\hat{R})$ are identical.

$\bar{r}_i = \frac{1^T \hat{r}_i}{I}$

$$\sigma_{i,j} = \frac{1}{M}\hat{r}_i^T\hat{r}_j - \bar{r}_i\bar{r}_j$$

$$\Sigma_{\hat{R}} = \frac{\hat{R}^T\hat{R}}{M} - \left(\bar{R}_i\bar{R}_j\right)$$

Principal Component Analysis (PCA)

The objective of principal component analysis (PCA) is to find a linear transformation Ω that maps a set of observed variables $\hat{R}$ into a set of uncorrelated variables F. We define the I x M statistical factor matrix as

$$F = \Omega\hat{R}$$

Where each row f_k $(k = 1, \dot{;} N)$ corresponds to a factor F of $\hat{R}$ and the transformation matrix Ω has elements $\omega_{k,i}$. The first row of ω_1 (which contains the first set of factor coefficients) is chosen such that the first factor (f_1) is aligned with the direction of maximal variance in the I-dimensional space defined by $\hat{R}$. Each subsequent factor (f_k) accounts for as much of the remaining variance of the standardized returns $\hat{R}$ as possible, subject to the constraint that the ω_k are mutually orthogonal. We further constrain the vectors ω_k by requiring that $\omega_k\omega_k^T = 1$ for all k.

The correlation matrix C is an I x I diagonalizable symmetric matrix that can be written in the form

$$C = \frac{1}{M}EDE^T$$

Where D is a diagonal matrix of eigenvalues d and E is an orthogonal matrix of the corresponding eigenvectors.

The eigenvectors of the correlation matrix C correspond to the directions of maximal variance such that $\Omega = E^T$, and one finds the statistical factors / principal components F using the diagonalization in .

If the sign of every coefficient in a statistical factor f_k is reversed, neither the variance of f_k nor the orthogonality of ω with respect to each of the other eigenvectors changes. For this reason, the signs of factors (PCs) are arbitrary. This feature of PCA can be problematic when we are interested in the temporal evolution of factors.

Proportion of Variance

The covariance matrix Σ_F for the statistical factor matrix F can be written as:

$$\Sigma_F = \frac{1}{M} F F^T = \frac{1}{M} \Omega \hat{R} \hat{R}^T \Omega^T = D$$

Where D is the diagonal matrix of eigenvalues d.

The total variance of the standardized returns $\hat{R}$ for the I investments is then

$$\sum_{i=1}^{I} \sigma^2(\hat{r}_i) = tr(\Sigma_{\hat{R}}) = \sum_{i=1}^{I} d_i = \sum_{i=1}^{N} \sigma^2(f_i) = tr(D) = I$$

Where $\Sigma_{\hat{R}}$ is the covariance matrix for $\hat{R}$

$\sigma^2(\hat{r}_i) = 1$ is the variance of the vector $\hat{r}_i$ of standardized returns for investment i.

The proportion of the total variance in $\hat{R}$ explained by the k^{th} factor is then

$$\frac{\sigma^2(f_k)}{\Sigma_{i=1}^{I} \sigma^2(\hat{r}_i)} = \frac{d_k}{\Sigma_{i=1}^{I} d_k} = \frac{d_k}{I}$$

The proportion of the variance from the k^{th} factor is equal to the ratio of the k^{th} largest eigenvalue d_k to the number of investments I.

The large variance in investment returns explained by a single factor implies that there is a large amount of common variation in the investment universe.

Random Matrix Theory (RMT)

Number of Significant Components

determine how many statistical factors are needed to describe the correlations between investments PCA is widely used to produce lower-dimensional representations of multivariate data by retaining a few "significant" components and discarding all other components. Many heuristic methods have been proposed for determining the number of significant factors, but there is no widespread agreement on an optimal approach

Apply two techniques to find the number of significant components. The assumes that a factor is significant if its eigenvalue $d > 1/N$. Any component that satisfies this criterion accounts for more than a fraction $(1/N)$ of the variance of the system. It is considered significant because it is assumed to summarize more information than any single original variable. The second approach is to compare the observed eigenvalues to the eigenvalues for random data and can be understood by considering the scree plot (figure ???). A scree plot shows the magnitudes of the eigenvalues as a function of the eigenvalue index, where the eigenvalues are sorted such that $\beta 1 \beta 2 \beta N$. The leftmost data point in the scree plot indicates the magnitude of the largest eigenvalue, and the rightmost data point indicates the magnitude of the smallest eigenvalue. The number of significant PCs is given by the number of eigenvalues in the scree plot for which the eigenvalue for the observed data is larger than the corresponding eigenvalue for random data

We explore

- topic intro
- outline how we determine # of significant components

- outline factor sensitivities [time permitting]
- outline VaR [time permitting]

Modeling: Results

- data driven approach

-use statistical methods to select and weight factors

- approach uses returns as the independent variables and factors as the dependent variables
- variety of estimation procedures, including classification trees, k-means, and principal components - that can be used to estimate these models.

-statistic is established to determine the criteria for a successful model

- algorithm of the statistical method evaluates the data and compares the results against the criteria.

Data Preprocessing

- de-trend and scale the returns

Statistical Factor Analysis

In this section was

- do PCA
- number of significant components
- universe diversification over time

Significant Statistical Factor Coefficients

An increase in the variance for which a factor accounts might be the result of increases in the correlations among only a few assets (which then have large factor coefficients) or a effect in which many investments begin to make significant contributions to the factor, This is an important distinction, because the two types of changes have very different implications for portfolio management. It becomes much more difficult to reduce risk by diversifying across different investment when correlations between all investments increase. In contrast, increases in correlations within an investment type that are not accompanied by increases in correlations between investment types have a less significant impact on diversification.

Inverse Participation Ratio (IPR) The inverse participation ratio I_k of the k^{th} factor ω_k is defined as:

$$IPR_k = \sum_{i=1}^{I} (\omega_{k,i})^4$$

The IPR quantifies the reciprocal of the number of elements that make a significant contribution to each eigenvector.

The behavior of the IPR is bounded by two cases:

[1] An eigenvector with identical contributions $\omega_{k,i} = frac1\sqrt{I}$ from all I investments has $IPR_k = \frac{1}{I}$

[2] An eigenvector with a single factor $\omega_{k,i} = 1$ and remaining factors equal to zero has $IPR = 1$

The inverse of the IPR - the so-called participation ratio - provides a more intuitive measure of the significance of a given factor as a large PR indicates that many investments contribute to the factor, while a small PR signals that few investments contribute to the factor:

$PR = \frac{1}{IPR_k}$

Temporal Evolution

We show as a function of time the fraction of the variance $fracd_kN$ due to the first 5 statistical factors $f_k(k = 1, ..., 5)$.

We also investigate temporal changes in the number of investments that make significant contributions to each statistical factor.

Conclusions

- state conclusions
- state how conclusions help direct future work

– facilitate sensitivity and scenarios analysis / stress testing

- state limitations of linear correlation

– correlation vs. causal

- discuss potential for future work

– Probabilistic Graphical Models (PGM): Bayesian Networks

References

[1]

[2] Hadley Wickham Refs

Fabozzi, Frank J.; Focardi, Sergio M.; Kolm, Peter N. (2010-01-29). Quantitative Equity Investing: Techniques and Strategies (Frank J. Fabozzi Series) (Kindle Location 4602). Wiley. Kindle Edition.

Acknowledgements

Appendix A: GitHub Repository

All of the R code used to produce this paper can be found in the following github repository:

https://github.com/dgn2/IS607_Final_Project

The R code required to:

- extract CTA manager, program, and monthly return data from the Altegris managed futures website
- create a MySQL database with tables to store extracted CTA manager, program, and monthly return data
- load CTA manager, program, and monthly return data to the MySQL database
- conduct limited exploratory analysis of the data
- conduct limited cleaning of the data used in subsequent statistical modeling
- estimate statistical factors based on the monthly returns of a select set of CTA programs

The github repository also includes the .Rmd file used to generate the .pdf working paper file.

Appendix B: Data Dictionary

The data dictionary for the data extracted from the Altegris managed futures website can be found in the github repository:

https://github.com/dgn2/IS607_Final_Project

Apprendix C: Fundamental Laws of Investing

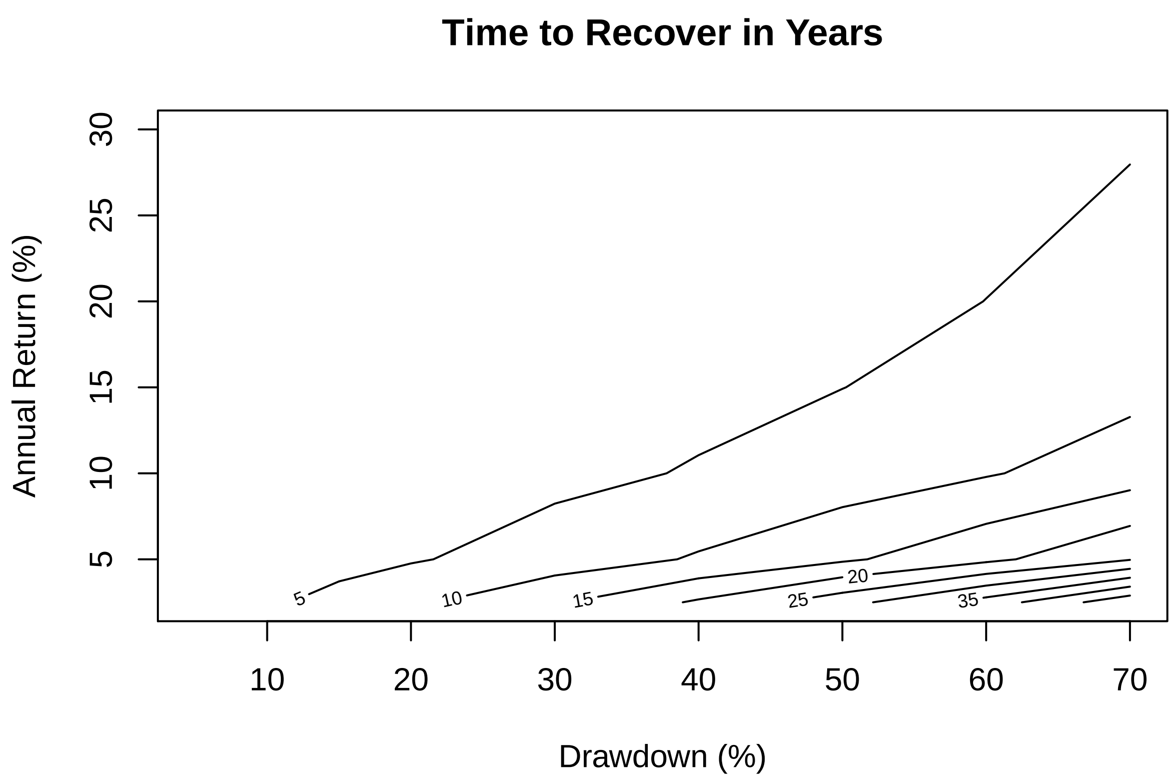
Compounding: Typical Return and Return Variability

- outline performance

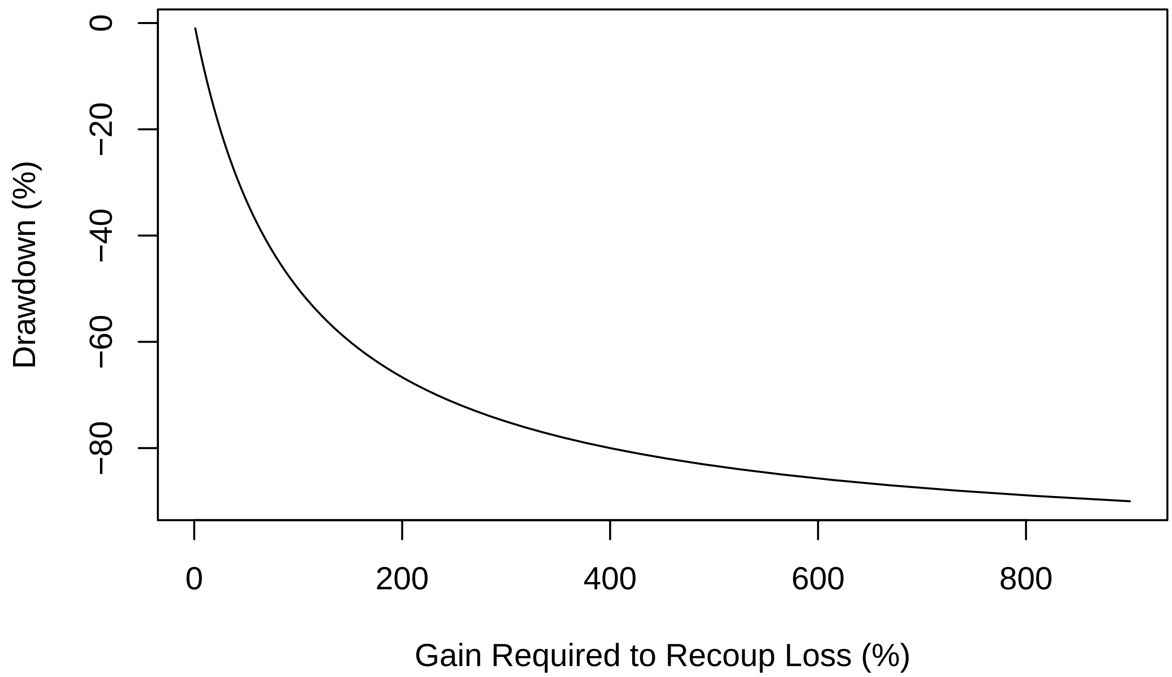
Importance of Capital Preservation

The amount to recover from a loss increases geometrically with the magnitude of the loss.

$$G = \left(\frac{1}{1-L}\right) - 1$$



	2.5	5	10	15	20	30
5	2.08	1.05	0.54	0.37	0.28	0.20
10	4.27	2.16	1.11	0.75	0.58	0.40
15	6.58	3.33	1.71	1.16	0.89	0.62
20	9.04	4.57	2.34	1.60	1.22	0.85
30	14.44	7.31	3.74	2.55	1.96	1.36
40	20.69	10.47	5.36	3.65	2.80	1.95
50	28.07	14.21	7.27	4.96	3.80	2.64
60	37.11	18.78	9.61	6.56	5.03	3.49
70	48.76	24.68	12.63	8.61	6.60	4.59



A loss of 20% requires a gain of 25% to recoup the loss.

A loss of 30% requires a gain of 43% to recoup the loss.

A loss of 40% requires a gain of 67% to recoup the loss.

A loss of 50% requires a gain of 100% to recoup the loss.

A loss of 60% requires a gain of 150% to recoup the loss.

A loss of 70% requires a gain of 233% to recoup the loss.

Importance of Diversification