Portfolio Optimization using Rank Correlation

Chanaka Edirisinghe
University of Tennessee, USA

Wenjun Zhou
University of Tennessee, USA

INTRODUCTION

Since the introduction of modern portfolio theory (Markowitz, 1952), quantitative analysis of financial data has contributed tremendously to informed decision making in finance, such as portfolio selection, risk management, and asset pricing, to name a few. As we witness the tremendous growth in the financial markets both in terms of the quantum of data generated and innovative financial products created, effective risk management has become an even more pressing need for the financial industry.

Apart from the so-called *risk-free* assets, such as treasury securities, most investments offer returns with some form of risk. While individual assets may carry varying degrees of return-risk characteristics, collectively, when incorporated within a portfolio of investments, the resulting portfolio is expected to yield a certain diversification of the risks so as to offer the investor with an acceptable return-risk profile. The success of this process is largely dependent not only on the specific exposure (or allocation) in each asset, and thus, its own marginal distributions of return, but also on the correlations among returns of the underlying assets. A diversification is expected to lessen risk exposure since each asset class has a different correlation to the others; when stocks rise, for example, bonds often fall. At a time when the stock market begins to fall, real estate may begin generating above average returns. Therefore, a specific investment allocation is influenced sig-nificantly by the underlying correlations among different assets.

Given the current price of an asset, its return (over a specific period of time in the future) is the price difference (positive/negative if the future price is higher/lower than the current) divided by the current price. Since the future price is unknown, asset return is uncertain and it may be modeled as a random variable. Considering a specific historical period of (observed) asset prices, the future asset return distribution may be hypothesized, and its key statistical parameters estimated, such as the mean and variance. Suppose there is a set of assets under consideration and their asset returns are shown to be uncorrelated. It is well known that by increasing the number of such assets in a portfolio, risk in portfolio return is diversified, and in the limit when the number of such assets becomes infinite, the resulting portfolio risk asymptotically disappears. However, the existence of such a large group of perfectly uncorrelated instruments defies the common-logic in the market place. Consequently, the investor is confronted with the question of how to diversify the portfolio risk in the face of existing correlated assets. The seminal work by Harry Markowitz (1952) is an attempt in this direction, whereby an optimal set of weights is determined for each asset simultaneously to achieve a prescribed level of expected return in the future whilst associating such a portfolio with a minimum level of risk, with risk being described by portfolio variance (Elton, Gruber, Brown, & Goetzmann, 2006). In this approach, the basic rule is simple: low

correlation makes for good diversification and highly correlated assets or asset classes are to be avoided. Thus, correlations play a fundamental and important role in the investment selection process.

Throughout the decades since, the above quantitative framework has inspired a large body of further work in financial data analysis and modeling, resulting in a wide range of software programs and tools that serve as decision support or help execute automated trading. There are also several subtle issues associated with estimating correlations, such as appropriate data transformation to ensure the assumed distribution, choice of statistical models, and predictions based on historical and simulated data.

An important aspect of correlations in aiding portfolio selection is the specific notion of the co-varying nature of asset returns. Modern risk management calls for an understanding of stochastic dependence that goes well-beyond simple linear (i.e., Pearson's) correlation (Rodgers & Nicewander, 1988). However, the dependency between risks, say, in the case of non-linear derivative assets, invalidates many of the distributional assumptions underlying the use of linear correlation. Stock returns, in empirical testing, violate these assumptions due to their inherent skewness and heavy-tailedness, i.e., multivariate stock returns distributions are non-elliptical (Chicheportiche & Bouchaud, 2012). In such a situation, the elegance of using linear correlations within Markowitz risk management framework breaks down. We conjecture that rank correlations among asset returns are preferable in this case in modeling the dependency structure.

In this work, we focus on Kendall's τ correlation coefficient (Kendall, 1970) with financial data. When the historical data set is sizeable (involving a large volume of transactions as needed for short term portfolio optimization), rank correlations become computationally more-intensive. In the case of intra-day high frequency trading using, say, hourly or 15 second-pricing information, computational complexities further compound barring the analyst from engaging in trading action with finer granularity.

While the Pearson's correlation measures the degree of linear dependence between two random variables, Kendall's rank correlation measures the degree of monotonic dependence (Embrechts, Mcneil, & Straumann, 2012). In this chapter, we illustrate how rank correlations can be efficiently computed based on an iterative scheme we develop. Moreover, by way of application to equity portfolio optimization, we demonstrate the superiority of the rank correlation-based portfolio selection using out-of-sample testing of the optimal portfolios so-devised.

BACKGROUND

Stock returns are quite well-known to have non-symmetric distributions. More importantly, stock return distributions are shown to have 'fatter' tails than the normal distribution would imply (Ziemba, 2003). Consequently, normal distribution assumption (or many other theoretical distributions) often leads to portfolios that perform poorly in practice (Edirisinghe, 2007, 2010). The main reason is that extreme events do occur much more frequently than most theoretical distributions correspond to. A linear correlation estimator such as Pearson's correlation coefficient is known to be susceptible to heavy-tailed noise and outliers, and thus, the use of it to determine mean-variance optimized investment portfolios can lead to poor risk control of the investments. Stock correlations should not be viewed as static; in fact, they can vary dynamically, and sometimes to a great extent (Engle, 2009).

A well-known deficiency of the linear correlation is that it is not invariant under nonlinear strictly increasing transformations, say $T : \Re \to \Re$ of two random variables X and Y. That is, in general, $\rho\left(T\left(X\right), T\left(Y\right)\right) \neq \rho\left(X, Y\right)$, and in particular, $\left|\rho\left(T\left(X\right), T\left(Y\right)\right)\right| \leq \left|\rho\left(X, Y\right)\right|$ where X and Y are bivariate normal, see for instance, Kendall

12 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/portfolio-optimization-using-rank-correlation/107375

Related Content

A Tree-Based Approach for Detecting Redundant Business Rules in Very Large Financial Datasets
Nhien-An Le-Khac, Sammer Markosand Tahar Kechadi (2012). *International Journal of Business Intelligence Research (pp. 1-13).*
www.irma-international.org/article/tree-based-approach-detecting-redundant/74732

Improving Cache Energy Efficiency for Green Computing
Sparsh Mittal (2014). *Encyclopedia of Business Analytics and Optimization (pp. 1155-1161).*
www.irma-international.org/chapter/improving-cache-energy-efficiency-for-green-computing/107314

Modeling-Centered Data Warehousing Learning: Methods, Concepts and Resources
Nenad Jukicand Boris Jukic (2012). *International Journal of Business Intelligence Research (pp. 74-95).*
www.irma-international.org/article/modeling-centered-data-warehousing-learning/74735

Artificial Intelligence in FinTech: Understanding Stakeholders Perception on Innovation, Disruption, and Transformation in Finance
Farida S. Rasiwalaand Bindya Kohli (2021). *International Journal of Business Intelligence Research (pp. 48-65).*
www.irma-international.org/article/artificial-intelligence-in-fintech/269446

Building Adaptive E-Business Infrastructure for Intelligent Enterprises
Liang-Jie Zhangand Jen-Yao Chung (2004). *Intelligent Enterprises of the 21st Century (pp. 92-105).*
www.irma-international.org/chapter/building-adaptive-business-infrastructure-intelligent/24243