

Mean-shortfall optimization problem with perturbation methods

Hayeon Won^a, Seyoung Park^{1,a}

^aDepartment of Statistics, Sungkyunkwan University

Abstract

Many researches have been done on portfolio optimization since Markowitz (1952) published a diversified investment model. Markowitz's mean-variance portfolio optimization problem is established under the assumption that the distribution of returns follows a normal distribution. However, in real life, the distribution of returns does not follow a normal distribution, and variance is not a robust statistic as it is heavily influenced by outliers. To overcome these potential issues, mean-shortfall portfolio model was proposed that utilized downside risk, shortfall, as a risk index. In this paper, we propose a perturbation method that uses the shortfall as a risk index of the portfolio. The proposed portfolio utilizes an adaptive Lasso to obtain a sparse and stable asset selection because it can reduce management and transaction costs. The proposed optimization is easily applicable as it can be computed using an efficient linear programming. In our real data analysis, we show the validity of the proposed perturbation method.

Keywords: portfolio optimization, shortfall, perturbation, linear programming

1. 서론

금융 분야에서 투자의 손실을 줄이고 높은 수익을 얻고자 하는 노력은 꾸준히 시도되어 왔고, 이를 수학적으로 해결하고자 하는 방법론이 포트폴리오 최적화(portfolio optimization)이다. 포트폴리오 최적화를 통해 투자자는 높은 기대수익률을 갖는 동시에 최소의 위험도를 갖는 최적의 포트폴리오 조합을 선택하고자 한다. 여기서 위험을 측정하는 지표로는 대표적으로 분산(variance)이 사용되어 진다. Markowitz (1952)는 분산을 위험 지표로 활용하는 평균-분산 포트폴리오(mean-variance portfolio) 최적화 문제를 도입하였다. Markowitz (1952)의 평균-분산 포트폴리오 최적화 모형은 현대 포트폴리오 이론(modern portfolio theory; MPT)의 근간이 되어, 이를 기반으로 한 Sharpe (1964)와 Lintner (1965)의 자본자산 가격결정 모형(capital asset pricing model; CAPM)이 제안되었다. 그리고 제로- β 포트폴리오로 무위험 자산(riskless asset)을 대체하는 제로- β CAPM (Black, 1972)으로 발전하였다. 이처럼 Markowitz (1952)의 평균-분산 포트폴리오 최적화 모형은 현대 금융학의 초석을 다지는 이론이 되었다.

그러나 포트폴리오 최적화 모형에서 위험 지표로 분산을 사용하는 것에 대한 여러 문제점들이 제기되었다. Markowitz (1952)의 평균-분산 포트폴리오 최적화 이론은 수익률 분포가 평균을 중심으로 대칭인 타원 분포(elliptical distribution)을 따른다는 것과 투자자의 효용 함수가 이차 효용(quadratic utility)이라는 가정을

This work was supported by a National Research Foundation of Korea grant funded by the Korea government (MSIP) (No. NRF-2019R1C1C1003805).

¹ Corresponding author: Department of Statistics, Sungkyunkwan University, 25-2, Seonggyunwan-ro, Jongno-Gu, Seoul 03063, Korea. E-mail: ishspsy@skku.edu.

기반으로 한다. 하지만 실생활에서 수익률 분포가 정확한 대칭을 이루는 경우가 드물며 (Bertsimas 등, 2004), 이차 효용 함수는 위험 회피(risk aversion) 투자자의 경우 부가 증가할수록 한계 효용이 감소하는 절대 위험 회피(absolute risk aversion)가 증가하는 특성을 갖는다 (Pratt, 1964; Arrow, 1971; Huang과 Litzenberger, 1988). 따라서 수익률 분포가 타원 분포를 따르지 않거나 위험 중립(risk neutral) 또는 위험 선호(risk seeking) 투자자의 경우에는 모형을 적용하기 어렵다는 한계가 있다. 그리고 분산을 줄이려는 시도는 하방위험(downside risk)뿐만 아니라 상방위험(upside risk)도 줄이게 되는 문제가 발생한다 (Grootveld와 Hallerback, 1998). 또한 분산은 이상치의 영향을 많이 받는 민감한 지표이므로 추정 오차(estimation error)가 커질 수도 있다는 문제점이 있다. 이러한 문제를 보완하기 위해 하방위험만을 다루는 반분산(semi-variance)과 lower partial moment (LPM), Value at Risk (VaR), conditional VaR (CVaR), 그리고 숏폴(Shortfall) 등 다양한 지표가 사용되고 있다. 하지만 LPM과 반분산 등은 계산하기 어렵다는 문제점이 있어 VaR가 주로 사용되어 왔으나, VaR는 비볼록성을 가지므로 최적화 문제에는 사용하기 어려운 한계가 있다 (Bertsimas 등, 2004). 숏폴은 VaR와 달리 볼록성을 가진 지표로 최적화에 사용하기 용이하다. 따라서 본 연구에서는 여러 하방위험 지표 중 숏폴을 사용한 평균-숏폴 포트폴리오 최적화 모형을 사용하였다.

보다 효율적인 포트폴리오 구성을 목적으로 Park 등 (2019)은 벤치마크 포트폴리오를 활용하여 포트폴리오 가중치를 제한함으로써 최적의 포트폴리오를 구하는 퍼터베이션(perturbation) 모형을 제안했다. Park 등 (2019)은 벤치마크 포트폴리오로 종합주가지수(market index)를 추종하는 추종펀드(index tracking fund) 또는 동일가중 포트폴리오(equally-weighted portfolio)를 활용하여 구성하는 포트폴리오의 일정 부분을 벤치마크에 투자하고 나머지 가중치에 편차를 주는 퍼터베이션 기법을 사용함으로써 벤치마크보다 위험률이 낮거나 수익률이 높은 포트폴리오를 구성한다. Park 등 (2019)의 퍼터베이션 기법은 West와 North 두 가지로 나뉘는데, 이는 효율적 투자선(efficient frontier) 평면에서 벤치마크를 서쪽(west) 또는 북쪽(north)로 보낸다는 의미에서 비롯되었다. 효율적 투자선 평면에서 벤치마크가 서쪽에 가까워지면 동일한 기대 수익률 하에서 위험도가 감소한다. 반대로 벤치마크가 북쪽에 가까워지면 동일한 위험도 하에서 수익률이 증가한다. 따라서 West는 위험을 낮추는 데 목적을 두는 반면 North는 높은 수익률을 얻고자 하는 목적으로 시행된다.

금융 투자에서는 낮은 위험도와 높은 수익률의 포트폴리오를 운용하는 것 못지않게 비용을 줄이는 것도 또한 중요하다. 예를 들어 고려하는 포트폴리오가 수익률이 높더라도 포트폴리오 운용 또는 거래 비용이 크다면 투자자의 실제 수익률은 감소될 수 있다. 포트폴리오 거래 비용은 포트폴리오를 조정하는 리밸런싱(rebalancing)시에 발생하는 비용으로 리밸런싱을 최소화하고 안정적인 포트폴리오(stable portfolio)를 얻음으로써 거래 비용을 줄일 수 있다. 또한 포트폴리오를 구성하는 자산의 수가 많아질수록 관리 비용이 늘어나므로 적은 수의 자산으로 구성된 포트폴리오(sparse portfolio)의 운용이 중요하다. Park 등 (2019)의 퍼터베이션 방법론은 adaptive Lasso 기법을 적용함으로써 적은 수의 자산을 사용한 안정적인 포트폴리오를 구성하였다. 그리고 보다 빠르고 효율적으로 최적화 문제를 풀기 위해 선형계획법(linear programming; LP)을 사용하였다. 본 연구에서는 Bertsimas (2004)의 평균-숏폴 포트폴리오 최적화 문제에 Park 등 (2019)의 퍼터베이션 기법을 적용한 모형을 제안하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서 숏폴과 평균-숏폴 포트폴리오 최적화 모형, 그리고 Park 등 (2019)의 퍼터베이션 기법에 대해 자세히 설명한다. 3장에서는 Park 등 (2019)의 퍼터베이션 기법과 평균-숏폴 포트폴리오 최적화 모형을 결합한 모형에 대해 제안한다. 4장에서는 뉴욕증권거래소의 다우존스산업평균지수(DOW30), 스탠다드앤드푸어스지수(S&P100; S&P500)와 런던증권거래소의 풋지수(FTSE100; FTSE250), 독일 프랑크푸르트 증권거래소의 닥스지수(DAX30) 등의 다양한 실증 데이터 활용하여 3장에서 제시한 모형의 타당성을 입증한다. 5장에서는 4장에서 제시한 결과를 바탕으로 결론을 내린다.

2. 샷폴과 퍼터베이션 기법

2.1. 하방위험과 샷폴

금융에서 자산의 수익률이 기대 수익률에서 벗어날 가능성을 변동성(volatility)이라 하는데, 변동성의 크기를 수치화하는 지표가 분산이다. Markowitz (1952)가 처음 제안한 포트폴리오 최적화 문제의 위험 지표로 분산이 사용되었다. 하지만 분산을 위험 지표로 사용하는 것에 대한 문제점들이 제기되었다. 제기된 문제점 중 하나는 투자자의 효용 함수가 이차 효용이라고 가정한다는 점이다. 이차 효용은 부가 증가할수록 효용은 감소하는 한계 효용 체감의 성질을 갖고, 또한 수익의 변화에 대한 위험회피도의 변화 정도인 절대적 위험회피도가 증가하는 특징을 갖는다 (Huang과 Litzenberger, 1988). 절대적 위험회피도가 증가한다는 것은 수익이 증가할수록 위험 자산에 대한 선호가 감소하는 것으로, 높은 수익을 위해 위험을 기꺼이 감수하는 위험선호 투자자와 반대되는 특성을 갖는다 (Huang과 Litzenberger, 1988). 따라서, 위험 회피자가 아닌 위험 선호자와 위험 중립 자에겐 Markowitz (1952)의 평균-분산 포트폴리오 최적화 모형을 적용하기 어렵다. 그리고 Markowitz (1952)의 평균-분산 최적화 모형은 위험 자산의 수익률 분포가 평균을 중심으로 대칭인 타원 분포라는 강한 가정을 기반으로 한다는 단점이 있다.

아래의 식은 Roy (1952)의 Safety-First criterion으로 하방위험의 개념을 도입한 첫 지표이다. 이는 위험(표준편차) 대비 기대수익률과 최소 수익률 임계값의 차이를 나타낸 값으로 수익률이 임계값보다 떨어질 확률을 고려한다.

$$\frac{E(r) - \underline{r}}{\sigma_p}, \quad (2.1)$$

여기서 $E(r)$ 은 포트폴리오 기대수익률, \underline{r} 은 최소 요구 수익률, 그리고 σ_p 은 포트폴리오의 표준편차를 나타낸다.

Markowitz (1959)는 반분산(semi-variance) 개념을 도입하였고, 좌우 대칭이 아닌 꼬리가 긴(skewed) 분포에서는 분산이 어떤 정보를 나타낼 수 없으며, 극단적으로 높거나 낮은 수익률, 즉 이상치에 대해 분산은 동일한 가중치를 적용한다는 단점이 있으므로 반분산을 사용해야 한다고 지적했다(Markowitz, 1959). 반분산의 일반화된 개념이 lower partial moment (LPM)인데, Fishburn (1977)에 의해 명시된 t 에 대한 k 차 LPM 식은 다음과 같다.

$$\text{LPM}_k(t) = \int_{-\infty}^t (t - r)^k dF(r), \quad k > 0, \quad (2.2)$$

여기서 t 는 목표 수준으로, 위 식은 목표 수준 t 이하로 수익률이 떨어질 확률을 나타낸다. 반분산은 LPM의 특수한 경우로 k 가 2이고 t 가 포트폴리오 기대수익률 $E(r)$ 로 대체될 때 성립한다. Fishburn (1977)은 투자자의 위험 선호도에 따라 k 값이 달라짐을 보였는데, $k > 1$ 이면 위험 기피(risk-aversion)이고 $k < 1$ 이면 위험 선호(risk-seeking), $k = 1$ 이면 위험 중립(risk-neutral)으로 구분하였다. 이처럼 LPM을 위험 지표로 사용한 평균-LPM 포트폴리오 최적화 모형은 위험 선호도에 따라 위험도가 달라지는 성질을 바탕으로 평균-분산 모형과 달리 위험 선호도에 따른 모형을 제시할 수 있다. 또한 LPM은 대칭이 아닌 수익률 분포에도 적용 가능하므로 평균-분산 모형의 가정을 필요로 하지 않는다는 점에서 직관적으로는 유용함을 보였지만 계산 상의 어려움 때문에 널리 사용되지 않았다 (Grootveld와 Hallerbach, 1999).

또다른 하방위험 지표로 분위수에 기반한 VaR가 있는데, VaR는 목표 기간 동안 주어진 신뢰수준에서 발생할 수 있는 최대손실크기를 의미한다 (RiskMetrics, 1996). Bertsimas 등 (2004)은 다음과 같은 가정 하에서 VaR를 아래의 식 (2.3) 과 같이 표기하였다. 확률 벡터 $R = (R_1, \dots, R_p)^T$ 가 수익률이고 $\mu = E(R)$ 를 기대수익률로 갖는 p 개의 자산이 있다고 가정해보자. 포트폴리오를 구성하는 p 개의 자산에 대한 가중치

$w = (w_1, \dots, w_p)^T$ 이면, 포트폴리오 수익률과 기대수익률은 각각 확률 벡터 $R^T w$ 와 $E(R^T w) = \mu^T w$ 가 된다. 확률변수 X 에 대한 α -분위수를 $q_\alpha(X) = \inf\{x | P(X \leq x) \geq \alpha\}$ 로 표기할 때, $(1 - \alpha)$ -신뢰 수준에서의 VaR는 다음과 같다:

$$\text{VaR}_\alpha(w) = \mu^T w - q_\alpha(R^T w), \quad \forall \alpha \in (0, 1). \quad (2.3)$$

여기서 $q_\alpha(R^T w)$ 는 포트폴리오 수익률 $R^T w$ 분포의 α 분위수다. 그러나 VaR는 준가법성과 볼록성의 부재 때문에 수학적으로 계산하기 어렵다는 문제점을 갖는다(Artzner, 1999; Rockafellar와 Uryasev, 2000; Rockafellar, 2007). 이후 VaR의 대안으로 샷폴이 제안되었다(Bertsimas 등, 2004). Bertsimas 등(2004)의 $(1 - \alpha)$ -신뢰 수준에서의 샷폴은 다음과 같이 나타낸다:

$$s_\alpha(w) = \mu^T w - E \left[R^T w | R^T w \leq q_\alpha(R^T w) \right], \quad \forall \alpha \in (0, 1). \quad (2.4)$$

샷폴과 VaR의 차이점은 식의 두 번째 항에 있다. 식 (2.3)의 두 번째 항을 $q_\alpha(R^T w)$ 에 대한 조건부 기댓값으로 대체함으로써 α -분위수 이하의 수익률을 고려한다. 즉, 샷폴은 포트폴리오 수익률이 α -분위수 이하로 발생할 평균적인 손실 크기를 나타낸다. 뿐만 아니라 조건부 기댓값을 취함으로써, 비볼록인 VaR와 달리 샷폴은 볼록성을 가지게 되어 최적화시 계산이 용이해진다는 장점이 있다.

2.2. 표본 평균-샷폴 포트폴리오 모형

Bertsimas 등 (2004)은 앞 절에서 제안한 샷폴을 위험 지표로 활용한 표본 평균-샷폴 포트폴리오 최적화 모형을 제안했다. 이 모형은 과거 데이터를 사용함으로써 수익률 분포에 대한 어떠한 가정도 하지 않는다는 점에서 현실의 제약에서 자유롭다는 장점을 갖는다. 기간 $t = 1, \dots, T$ 에 대해 T 개의 과거 수익률 표본이 r_1, \dots, r_T 라고 할 때, 그 다음 기간 $T + 1$ 에서의 수익률 추정량이 과거 T 개 수익률의 평균 $\bar{r} = \sum_{i=1}^T r_i / T$ 이라고 가정한다. Bertsimas (2004) 등은 샷폴의 비모수 추정량으로 $\hat{s}_\alpha(w) = \mu^T \bar{r} - \sum_{i=1}^K r_{(i)}(w) / K$ 을 고려하는데, 여기서 $r_{(i)}(w)$ 는 T 개의 과거 수익률의 i 번째로 작은 순서통계량이며 $K = \lfloor \alpha T \rfloor$ 로 αT 보다 크지 않은 가장 큰 정수값을 갖는다. 샷폴의 비모수 추정량을 활용한 표본 평균-샷폴 포트폴리오 최적화 모형은 다음과 같다.

$$\begin{aligned} \min_w \quad & w^T \bar{r} - \frac{1}{K} \sum_{i=1}^K r_{(i)}(w), \\ \text{subject to} \quad & w^T \bar{r} = r_p, \\ & w^T \mathbf{1} = 1, \end{aligned} \quad (2.5)$$

여기서 r_p 는 목표 포트폴리오 수익률이다. 식 (2.5)에는 순서통계량이 포함되어 있으므로 다음과 같이 식을 선형계획법으로 변환한다.

$$\begin{aligned} \min_{w, t, z} \quad & w^T \bar{r} - t + \frac{1}{K} \sum_{i=1}^T z_i, \\ \text{subject to} \quad & w^T \bar{r} = r_p, \\ & w^T \mathbf{1} = 1, \\ & z_i \geq t - w^T r_i, \\ & z_i \geq 0, \quad i = 1, \dots, T. \end{aligned} \quad (2.6)$$

식 (2.5)를 식 (2.6)으로 변환하는 과정에서 새로운 변수 z_i , $i = 1, \dots, T$ 가 생성되며 자세한 증명 과정은 Bertsimas 등 (2004)에 있으므로 생략하도록 한다. 식 (2.6)의 최적화 모형에 Park 등 (2019)의 퍼터베이션 기법을 접목한 모형을 3장에서 제안한다.

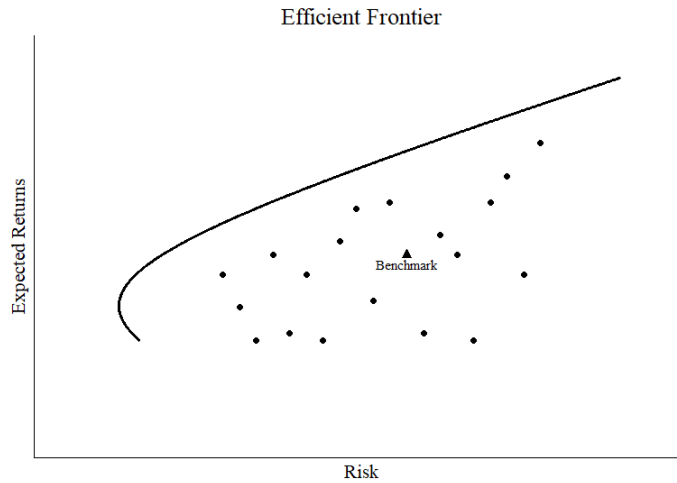


Figure 1: Efficient frontier.

2.3. 퍼터베이션 기법

금융 분야에서 벤치마크를 활용하여 더 높은 수익률을 가지는 포트폴리오를 운용하는 역사는 오래되었다. 하지만 벤치마크보다 높은 초과 수익률(excess return)을 얻으면서 벤치마크와 포트폴리오 간의 수익률 차이에서 비롯되는 추적 오차(tracking error)만을 줄이는 전략에는 전체 포트폴리오 위험도가 오히려 커져버리는 딜레마가 존재했다 (Roll, 1992). Jorion (2003)은 추적 오차와 전체 포트폴리오 위험도 모두 줄이면서 초과 수익을 얻는 최적의 포트폴리오를 얻는 모형을 제안했다.

이에 착안하여 Park 등 (2019)이 제안한 퍼터베이션(perturbation) 기법은 인덱스펀드 등과 같은 벤치마크를 활용하여 수익률을 벤치마크화하며 위험도를 낮추는 동시에 포트폴리오 비용까지 조정하는 방법론으로 벤치마크의 가중치에 약간의 편차(perturbation)를 줌으로써 포트폴리오의 성능을 조정한다. 포트폴리오의 위험도와 수익률 중 어느 측도에 초점을 두느냐에 따라 West와 North, 두 가지 방법으로 나뉘는데 West는 벤치마크를 활용하여 포트폴리오의 위험도를 줄이는 전략이고, North는 위험도를 일정 수준 유지하되 수익률을 높이는 전략이다.

Figure 1는 포트폴리오의 효율적 투자선(efficient frontier)을 나타낸다. 그래프에서 수평축은 위험도(risk)를, 수직축은 기대 수익률(expected return)을 의미한다. 동일한 위험도에서 최대의 기대 수익률을 갖는 효율적인(efficient) 포트폴리오들을 연결한 곡선이 효율적 투자선이며, 효율적 투자선 아래의 각각의 점들은 비효율적인(inefficient) 포트폴리오를 나타낸다. 따라서 효율적 투자선에 가까운 포트폴리오일수록 동일한 위험도에서 더 높은 기대 수익률을 갖는 더 나은 포트폴리오임을 의미한다.

퍼터베이션 기법의 West와 North는 이 효율적 투자선 평면 상에서 움직이고자 하는 목표 방향과 관련이 있다. Figure 1에서 점 Benchmark는 벤치마크 포트폴리오를 의미하는데, 벤치마크보다 효율적 투자선의 서쪽(west)에 가까워지면 동일한 기대 수익률에서 위험도가 감소한다. 반면, 벤치마크보다 효율적 투자선의 북쪽(North)으로 가까워지면 동일한 위험도에서 기대 수익률은 증가하게 된다. 즉, West와 North는 벤치마크를 활용하여 벤치마크보다 더 낮은 위험도를 갖거나, 또는 더 높은 기대 수익률을 갖는 포트폴리오를 구성하는 기법이다. 따라서, 벤치마크보다 낮은 위험도 또는 높은 기대 수익률 어느 전략에 초점을 두느냐에 따라서 투자자는 West와 North 중에 하나를 선택하여 포트폴리오 최적화 문제에 적용할 수 있다. 벤치마크보다 위험도를 낮추는 데 중점을 두는 투자자는 West 기법을, 반대로 벤치마크보다 높은 기대수익률을 얻는 데 중점을

두는 North 기법을 적용한 포트폴리오를 운용할 수 있다. Park 등 (2019)의 퍼터베이션 기법은 벤치마크를 활용하여 어느 정도 안정적인 수익률을 얻으면서 퍼터베이션 기법을 통해 포트폴리오의 성능을 조절하는 동시에, 포트폴리오 비용을 줄일 수 있는 현실적인 제안도 하였다는 점에서도 의의가 있다.

포트폴리오 비용은 포트폴리오를 구성하는 자산의 수에 따라 결정되는 운용 비용(management cost)과 포트폴리오 가중치를 조정하는 리밸런싱(rebalancing)에서 발생하는 세금과 수수료 등의 거래 비용(transaction cost) 등 두 가지로 나뉜다. 아무리 수익률이 높고 위험도가 낮은 포트폴리오라도, 포트폴리오를 구성하는 자산의 수가 많거나 리밸런싱의 빈도가 높을수록 비용이 증가하므로 투자자에게 돌아가는 수익이 줄어드는 바람직하지 않은 결과가 발생한다. 따라서 Park 등 (2019)은 adaptive Lasso를 적용하여 자산의 수가 적고 안정적인 포트폴리오를 얻어 비용을 줄일 수 있는 모형을 제안했다. 예를 들어 하나의 포트폴리오를 구성하는 자산의 수가 p 개라고 가정할 때 자산의 수익률 벡터는 $r = (r_1, \dots, r_p)^T$ 로 표기한다. p 개의 자산에 대한 기대수익률 벡터와 공분산 행렬은 각각 $\mu = (\mu_1, \dots, \mu_p)^T$ 와 Σ 로 표기되며, 그 추정치는 $\hat{\mu} = (\hat{\mu}_1, \dots, \hat{\mu}_p)^T$ 와 $\hat{\Sigma}$ 로 나타낸다. p 개의 자산으로 이뤄진 포트폴리오 가중치는 $w = (w_1, \dots, w_p)^T$ 으로 표기하고 합을 1로 가정한다. 즉, $\sum_i w_i = 1$ 이다.

전체 포트폴리오에서 β 만큼은 벤치마크인 v 에 투자하고, 나머지 $1 - \beta$ 는 금융 시장의 자산들에 투자할 때, 포트폴리오 가중치는 $w = \beta v + \delta$ 으로 표기할 수 있으며, 여기서 $\delta = (\delta_1, \dots, \delta_p)^T$ 는 벤치마크에 대한 편차(퍼터베이션)을 나타낸다. $\beta \neq 1, 0 < c_1 \leq 1$ 일 때, West 퍼터베이션 모형은 다음과 같다.

$$\begin{aligned} \min_{\delta \in \mathbb{R}^p} \quad & \sum_{i=1}^p \frac{|\delta_i|}{|\hat{\delta}_i|}, \\ \text{subject to} \quad & \delta^T \mathbf{1} = 1 - \beta, \\ & (\beta v + \delta)^T \hat{\Sigma} (\beta v + \delta) \leq c_1 \hat{\sigma}_v^2, \end{aligned} \quad (2.7)$$

여기서 $\hat{\sigma}_v^2$ 은 벤치마크 위험(분산)의 추정치이고, $\hat{\delta}_i$ 는 이전 시간에서의 i 번째 자산에 대한 가중치이다. $\hat{\sigma}_v^2$ 은 추정 시점으로부터 과거 12개월의 표본을 활용하여 추정되며, 업데이트 되는 가중치 δ_i 는 과거 12개월의 자료를 활용하여 추정된다. 자산의 수가 적고 안정적인 포트폴리오를 구성하기 위해 adaptive Lasso 기법을 적용하여 식 (2.7)의 목적함수를 구성하는데, ℓ_1 패널티인 $\sum_{i=1}^p |\delta_i|$ 를 최소화함으로써 자산의 수가 적은 포트폴리오를 얻을 수 있고, $1/|\hat{\delta}_i|$ 만큼의 가중치를 부여함으로써 안정적인 포트폴리오를 얻을 수 있다. 예를 들어 이전 시간에서의 가중치 값 $\hat{\delta}_i$ 가 0에 가까운 작은 값을 갖는다면 현 시점에서의 가중치 추정값 δ_i 도 0에 가까운 값을 갖게 될 개연성이 크다. 반면에 이전 시간에서의 가중치의 절대값이 크다면 현 시점에서의 추정값도 상대적으로 커질 개연성이 있다.

다음으로 $\beta \neq 1, c_2 \geq 1$ 일 때, North 퍼터베이션 모형은 다음과 같다.

$$\begin{aligned} \min_{\delta \in \mathbb{R}^p} \quad & \sum_{i=1}^p \frac{|\delta_i|}{|\hat{\delta}_i|}, \\ \text{subject to} \quad & \delta^T \mathbf{1} = 1 - \beta, \\ & (\beta v + \delta)^T \hat{\Sigma} (\beta v + \delta) \leq c_1 \hat{\sigma}_v^2, \\ & \hat{\mu}_v a + \hat{\mu}^T \delta \geq \hat{\mu}_v + c_2 |\hat{\mu}_v|. \end{aligned} \quad (2.8)$$

여기서 $\hat{\mu}_v$ 는 벤치마크 기대수익률의 추정값이다. $\hat{\mu}_v$ 와 $\hat{\sigma}_v^2$ 도 West와 마찬가지로 과거 12개월의 표본을 활용하여 표본외 방식으로 추정된다. 하지만 West와의 차이점은 세 번째 제약조건인데, 이 조건이 추가됨으로써 벤치마크의 기대수익률보다 최소한 $c_2 |\hat{\mu}_v|$ 만큼 더 높은 기대 수익률을 갖는 포트폴리오를 얻을 수 있다.

3. 퍼터베이션 기법을 활용한 평균-숏폴 포트폴리오

3.1. 표본 평균-숏폴 모형

본 절에서는 2장의 표본 평균-숏폴 최적화 모형에 퍼터베이션 기법을 접목한 모형들을 제안하고자 한다.

특정 기간 t 에서 p 개의 자산에 대한 수익률 표본 벡터를 $r_t = (r_{t1}, \dots, r_{tp})^T$ 라고 할 때, 기간 $t = 1, \dots, T$ 동안의 p 개의 자산에 대한 수익률 표본의 평균은 벡터 \bar{r} 로 표기한다. 포트폴리오 퍼터베이션 가중치를 $\delta = (\delta_1, \dots, \delta_p)^T$, 벤치마크 비율은 β 로 표기할 때, West 최적화 모형은 다음과 같다: $\beta \neq 1, \lambda > 0$, 그리고 $0 \leq c_3 \leq 1$ 에 대하여,

$$\min_{\delta, t, z} (\beta v + \delta)^T \bar{r} - t + \frac{1}{K} \sum_{i=1}^T z_i + \lambda \sum_{j=1}^p \frac{|\delta_j|}{|\bar{\delta}_j|}, \quad (3.1)$$

$$\text{subject to } \delta^T \mathbf{1} = 1 - \beta,$$

$$z_i \geq t - (\beta v + \delta)^T r_i,$$

$$z_i \geq 0, \quad i = 1, \dots, T,$$

$$\bar{r}_v \beta + \bar{r}^T \delta \geq c_3 \bar{r}_v.$$

여기서 $\bar{\delta}_j$ 는 이전 기간의 j 번째 포트폴리오 가중치이며, $\bar{r}_v = v^T \bar{r}$ 는 벤치마크의 추정 기대수익률이다. 전체 포트폴리오 가중치에서 β 만큼은 벤치마크에 투자하고 $1 - \beta$ 는 나머지 자산들에 투자하므로 그 조건이 첫 번째로 주어진다. 두번째 그리고 세번째 제약조건은 (2.6)에서 볼 수 있듯이 Bertsimas 등(2004)의 표본-평균 숏폴 최적화 모형을 선형계획법으로 치환하는 과정에서 도출된 제약 조건이다. 마지막으로 West는 벤치마크를 활용하여 포트폴리오의 위험도를 줄이는 데 목표를 두지만, 포트폴리오 수익률이 너무 낮아지는 것을 방지하기 위해 벤치마크의 수익률의 c_3 만큼 보다는 크다는 조건을 추가하였다. 모형 (3.1)에서 네 번째 제약조건에서의 c_3 값이 감소할 때 가능해 영역(feasible region)이 커지므로 제약조건을 만족하는 해가 존재하게 될 확률이 커지지만, 포트폴리오의 수익률은 낮아질 가능성이 크다. 따라서 가능해 영역이 존재하는 상황에서 모형 (3.1)을 풀기 위해 충분히 큰 c_3 값인 1부터 시작하여 0.1씩 그 값을 감소시키면서 가능해 영역이 존재하게 만들어주는 가장 큰 c_3 에 대해서 (3.1)를 풀었다.

North 모형은 다음과 같다: $\beta \neq 1, \lambda > 0$, 그리고 $0 < c_4 \leq 1$ 에 대하여,

$$\min_{\delta, t, z} (\beta v + \delta)^T \bar{r} - t + \frac{1}{K} \sum_{i=1}^T z_i + \lambda \sum_{j=1}^p \frac{|\delta_j|}{|\bar{\delta}_j|}, \quad (3.2)$$

$$\text{subject to } \delta^T \mathbf{1} = 1 - \beta,$$

$$z_i \geq t - (\beta v + \delta)^T r_i,$$

$$z_i \geq 0, \quad i = 1, \dots, T,$$

$$\bar{r}_v \beta + \bar{r}^T \delta \geq \bar{r}_v + c_4 \bar{r}_v.$$

North 모형과 West 모형의 차이점은 네번째 제약 조건에 있다. 위험도를 낮추는 데 집중한 West와 달리, North는 위험도는 일정 수준 유지하면서 수익률을 높이는 데 목적을 둔다. 따라서 모형의 네 번째 제약 조건에 벤치마크 기대수익률에 더하여 c_4 만큼 보다 더 크다는 조건을 추가하였다. 모형 (3.1)에서와 같은 방법으로 제약 조건들을 모두 만족하는 가능해 영역(feasible region)이 존재하는 상황에서의 해를 구하기 위해, c_4 값을 1부터 0.1씩 감소시켜 네 번째 제약 조건을 완화함으로써 실현가능한 해를 구한다.

3.2. 포트폴리오 가중치의 금융학적 의미

본 절에서는 3.1 절에서 제안한 표본 평균-숫폴 포트폴리오 최적화 식을 수학적으로 바라본 후 퍼터베이션 가중치의 해 $\tilde{\delta}$ 의 금융학적 성질에 대해서 설명을 하고자 한다. 구체적으로는 라그랑즈 승수법(lagrange multiplier method)을 이용하여 KKT(Karush-Kuhn-Tucker) 최적화 조건을 만족하는 해의 성질에 대해서 알아 본다.

라그랑즈 승수법에 따르면, West 모형은 다음과 같이 라그랑즈 함수 L_1 로 재구성될 수 있다.

$$\begin{aligned} L_1 = & (\beta v + \delta)^T \bar{r} - t + \frac{1}{K} \sum_{i=1}^T z_i + \lambda \sum_{j=1}^p \frac{|\delta_j|}{|\tilde{\delta}_j|} \\ & - \sum_{i=1}^T \lambda_{1i} \left\{ z_i - \left(t - (\beta v + \delta)^T r_i \right) \right\} - \sum_{i=1}^T \lambda_{2i} z_i \\ & - \lambda_3 \left(\bar{r}_v \beta + \bar{r}^T \delta - c_3 \bar{r}_v \right) + \lambda_4 \left(1 - \beta - \delta^T 1 \right) \end{aligned} \quad (3.3)$$

유클리드 공간 상의 벡터 g 를 $\sum_{j=1}^p |\delta_j|/|\tilde{\delta}_j|$ 의 sub-gradient라고 가정하면, 벡터 g 에 대한 KKT 최적화 조건들은 다음과 같다:

$$\begin{aligned} \bar{r}_j + \lambda g_j - \sum_{i=1}^T \lambda_{1i} r_{ij} - \lambda_3 \bar{r}_j - \lambda_4 &= 0, \quad x_j = 1, \dots, p, \\ \frac{1}{K} - \lambda_{1i} - \lambda_{2i} &= 0, \quad i = 1, \dots, T, \\ -1 + \sum_{i=1}^T \lambda_{1i} &= 0, \\ \lambda_{1i} \left\{ z_i - \left(t - (\beta v + \delta)^T r_i \right) \right\} &= 0, \quad \lambda_{2i} z_i = 0, \quad \lambda_3 (\bar{r}_v \beta + \bar{r}^T \delta - c_3 \bar{r}_v) = 0, \\ z_i \geq t - (\beta v + \delta)^T r_i, \quad z_i &\geq 0, \\ \bar{r}_v \beta + \bar{r}^T \delta \geq c_3 \bar{r}_v, \quad \delta^T 1 &= 1 - \beta. \end{aligned}$$

여기서 처음 세 조건(complementary slackness)은 다음을 나타낸다.

$$\begin{aligned} \lambda g_j &= \sum_{i=1}^T \lambda_{1i} r_{ij} - \bar{r}_j + \lambda_3 \bar{r}_j + \lambda_4 \\ &= \sum_{i=1}^T \lambda_{1i} (r_{ij} - \bar{r}_j) + \lambda_3 \bar{r}_j + \lambda_4, \quad j = 1, \dots, p. \end{aligned}$$

위 식은 λ_{2i} 의 조건에 따라 나누어 볼 수 있다. 만약 $\lambda_{2i} = 0$, 즉 $z_i \geq 0$ 이면, λg_j 는 다음과 같다:

$$\begin{aligned} \lambda g_j &= \sum_{i=1}^T \frac{1}{K} (r_{ij} - \bar{r}_j) + \lambda_3 \bar{r}_j + \lambda_4 \\ &= \lambda_3 \bar{r}_j + \lambda_4 \end{aligned}$$

반면에 $\lambda_{2i} \neq 0$, 즉 $z_i = 0$ 이면, λg_j 는 다음과 같다:

$$\begin{aligned}\lambda g_j &= \sum_{i=1}^T \left(\frac{1}{K} - \lambda_{2i} \right) (r_{ij} - \bar{r}_j) + \lambda_3 \bar{r}_j + \lambda_4 \\ &= \sum_{i=1}^T \left(1 - \frac{T}{K} \right) (r_{ij} - \bar{r}_j) + \lambda_3 \bar{r}_j + \lambda_4 \\ &= \lambda_3 \bar{r}_j + \lambda_4.\end{aligned}$$

따라서, 모든 λ_{2i} 값에 대하여 $g_j = (\lambda_3/\lambda)\bar{r}_j + \lambda_4/\lambda$ 이다. 여기서 만약 $\lambda_3 \neq 0$ 이면, $\hat{\delta}_j$ 는 다음과 같다:

$$\begin{aligned}g_j \geq 0, \quad \hat{\delta}_j \geq 0, \quad \text{for } \bar{r}_j \geq -\frac{\lambda_4}{\lambda_3}, \\ g_j \leq 0, \quad \hat{\delta}_j \leq 0, \quad \text{for } \bar{r}_j \leq -\frac{\lambda_4}{\lambda_3}.\end{aligned}$$

$\hat{\delta}_j$ 의 부호는 j 번째 자산의 공매도(short selling)의 여부를 나타낸다. 구체적으로 j 번째 자산의 가중치 $\hat{\delta}_j$ 가 음수이면 공매를 하는 경우(short position)을 의미하고, $\hat{\delta}_j$ 가 양수이면 증권을 매입하는 것(long position)을 의미한다. 공매도는 주가 하락이 예상되는 종목의 증권을 빌려서 매도하여 차익을 얻는 투자 방법으로, 그 종목의 평균적인 과거 수익률의 특정 값 이하면 현재 값도 그 값 이하일 것으로 예상하여 공매도가 증가할 수 있다. 따라서 위 식들은 j 번째 자산의 과거 평균 수익률인 \bar{r}_j 가 특정 값보다 크면 공매를 하고, \bar{r}_j 가 특정 값보다 크면 증권을 매입한다는 금융학적인 의미를 갖는다.

반면에, $\lambda_3 = 0$ 이면, $\hat{\delta}_j$ 는 다음과 같다:

$$g_j = \frac{\lambda_4}{\lambda} \geq 0, \quad \hat{\delta}_j \geq 0.$$

만약 포트폴리오 가중치 $\hat{\delta}_j$ 가 모두 0보다 크면 공매도를 하지 않은 롱온리(long-only) 포트폴리오이며, 그렇지 않을 경우 $\hat{\delta}$ 공매도가 포함된 롱-숏(long-short) 포트폴리오를 의미한다. 공매도는 주식 시장의 유동성과 효율성을 높일 수는 있으나, 차입 공매도(cover short)에서 수수료와 이자 비용이 발생하며 개인 투자자의 경우 기관 투자자 대비 정보의 비대칭성으로 인해 공매도를 통해 차익을 실현하기가 쉽지 않다. 또한 공매도는 시장의 안정성을 해친다는 점에서 시장이 불안정할 땐 금지되곤 한다. 이러한 점에서 롱온리 포트폴리오를 구성할 수 있다는 점은 실용성이 있다. West 모형은 위와 같은 특정 조건하에서 롱온리 포트폴리오 또는 롱-숏 포트폴리오로 운용될 수 있다는 의미로 해석할 수 있다.

마찬가지로 North 모형 또한 라그랑주 함수 L_2 로 재구성될 수 있다.

$$\begin{aligned}L_2 &= (\beta v + \delta)^T \bar{r} - t + \frac{1}{K} \sum_{i=1}^T z_i + \lambda \sum_{j=1}^p \frac{|\delta_j|}{|\tilde{\delta}_j|} \\ &\quad - \sum_{i=1}^T \lambda_{1i} \left\{ z_i - (t - (\beta v + \delta)^T r_i) \right\} - \sum_{i=1}^T \lambda_{2i} z_i \\ &\quad - \lambda_3 (\bar{r}_v \beta + \bar{r}^T \delta - \bar{r}_v - c_4 |\bar{r}_v|) + \lambda_4 (1 - \beta - \delta^T 1)\end{aligned}$$

유클리드 공간 상의 벡터 g 를 $\sum_{j=1}^p |\delta_j|/|\tilde{\delta}_j|$ 의 sub-gradient라고 가정하면, 벡터 g 에 대한 KKT 최적화 조건

들은 다음과 같다:

$$\begin{aligned} \bar{r}_j + \lambda g_j - \sum_{i=1}^T \lambda_{1i} r_{ij} - \lambda_3 \bar{r}_j - \lambda_4 &= 0, \quad j = 1, \dots, p \\ \frac{1}{K} - \lambda_{1i} - \lambda_{2i} &= 0, \quad i = 1, \dots, T \\ -1 + \sum_{i=1}^T \lambda_{1i} &= 0, \\ \lambda_{1i} \{z_i - (t - (\beta v + \delta)^T r_i)\} &= 0, \quad \lambda_{2i} z_i = 0, \quad \lambda_3 (\bar{r}_v \beta + \bar{r}^T \delta - \bar{r}_v - c_4 |\bar{r}_v|) = 0, \\ z_i &\geq t - (\beta v + \delta)^T r_i, \quad z_i \geq 0, \\ \bar{r}_v \beta + \bar{r}^T \delta &\geq \bar{r}_v + c_4 |\bar{r}_v|, \quad \delta^T \mathbf{1} = 1 - \beta. \end{aligned}$$

West와 마찬가지로, KKT 최적화 조건을 통해 다음과 같은 해를 얻을 수 있다. 자세한 증명 과정은 생략하도록 한다.

$\lambda_3 \neq 0$ 일때, $\hat{\delta}$ 은 다음과 같다:

$$\begin{aligned} g_j &\geq 0, \quad \hat{\delta}_j \geq 0, \quad \text{for } \bar{r}_j \geq -\frac{\lambda_4}{\lambda_3}. \\ g_j &\leq 0, \quad \hat{\delta}_j \leq 0, \quad \text{for } \bar{r}_j \leq -\frac{\lambda_4}{\lambda_3}. \end{aligned}$$

$\lambda_3 = 0$ 일때, $\hat{\delta}$ 은 다음과 같다:

$$g_j = \frac{\lambda_4}{\lambda} \geq 0, \quad \delta_j \geq 0.$$

North도 West와 마찬가지로 $\lambda \neq 0$ 이면 j 번째 자산의 과거 평균 수익률이 특정 값보다 크면 공매를 하고 반대의 경우엔 증권을 매입하나, $\lambda = 0$ 이면 공매도가 포함되지 않은 롱온리 포트폴리오를 운용하는 금융학적 의미를 갖는다.

4. 실증 자료 분석

4.1. 자료 설명

본 절에서는 실제 금융 자료를 활용하여 모형의 성능을 측정하고자 한다. 자료들은 2004년 1월부터 2015년 12월까지 기록된 것으로, 뉴욕증권거래소의 다우존스산업평균지수(DOW30)와 스탠다드앤드푸어스지수(S&P100; S&P500), 런던증권거래소의 풋시지수(FTSE100; FTSE250), 그리고 독일 프랑크푸르트 증권거래소의 닥스지수(DAX30) 등 총 6개의 데이터셋을 사용하였다.

먼저, DOW30 지수와 S&P100 지수, S&P500 지수의 데이터셋에는 각각 29개, 89개, 445개의 자산이 포함되어 있다. 이는 12년의 기간동안 발생한 결측을 제거하는 과정에서 비롯된 결과이다. 그리고 FTSE100 지수와 FTSE250 지수는 81개, 170개의 자산이 있으며, DAX30 지수에는 30개의 자산이 있다.

4.2. 알고리즘

West 모형인 최적화 식 (3.1)에서의 시간의 흐름에 따른 알고리즘은 다음과 같다. 먼저 입력값을 지정하고 아래의 3단계의 과정을 T 번 반복한 후, 식의 해를 출력한다.

입력: 시간과 포트폴리오 가중치의 초기값; $t = T_0$ (월), $\tilde{\delta}^{(T_0)} = (1/p, \dots, 1/p)^T$

출력: 포트폴리오 가중치 편차 벡터; $\hat{\delta}^{(t)} \in \mathbb{R}^p$, $t = T_0, \dots, T_0 + T$.

다음 1-3단계를 $t = T_0 + T$ 가 될 때까지 T 번 반복한다.

1단계: $\tilde{\delta} = \tilde{\delta}^{(t)}$ 와 $\alpha, \beta, \lambda, c_3$ 값을 넣어 최적화 식 (3.1)의 해 $\hat{\delta}^{(t)}$ 을 구한다.

2단계: $\tilde{\delta}^{(t+1)}$ 를 $\hat{\delta}^{(t)}$ 으로 업데이트한다.

3단계: $t \leftarrow t + 1$

North 모형인 최적화 식 (3.2)에서의 시간의 흐름에 따른 알고리즘도 West와 마찬가지로, 먼저 입력값을 지정하고 3단계의 과정을 T 번 반복한 후, 식의 해를 출력한다.

입력: 시간과 포트폴리오 가중치의 초기값 $t = T_0$ (월), $\tilde{\delta}^{(T_0)} = (1/p, \dots, 1/p)^T$

출력: 포트폴리오 가중치 편차 벡터 $\hat{\delta}^{(t)} \in \mathbb{R}^p$, $t = T_0, \dots, T_0 + T$.

다음 1-3단계를 $t = T_0 + T$ 가 될 때까지 T 번 반복한다.

1단계: $\tilde{\delta} = \tilde{\delta}^{(t)}$ 와 $\alpha, \beta, \lambda, c_4$ 값을 넣어 최적화 식 (3.2)의 해 $\hat{\delta}^{(t)}$ 을 구한다.

2단계: $\tilde{\delta}^{(t+1)}$ 를 $\hat{\delta}^{(t)}$ 으로 업데이트한다.

3단계: $t \leftarrow t + 1$

4.3. 모수 선택

평균-숏폴 최적화 모형인 식 (3.1)과 식 (3.2)를 푸는 데 있어서 adaptive Lasso의 모수 λ 의 값으로 $\{1, 10, 50, 100, 500, 1000, 5000, 10000\}$ 를 고려하였고, 10-fold 교차 검증법(cross-validation)을 통해 가장 성능이 좋은 λ 를 최종 모수로 선택하였다. 교차 검증 성능을 평가하는 지표로서 포트폴리오 평균수익률(Mean return) 또는 샤프지수(Sharpe ratio), 이 두 가지 지표를 사용하였다. 좀더 구체적으로는 최근 1년의 데이터를 랜덤하게 10 등분하여 9개의 그룹을 훈련 집합(training set)으로 사용하여 각각의 λ 값에 대해 식 (3.1)과 식 (3.2)를 풀고, 남은 1개의 그룹을 테스트 집합(test set)으로 사용하여 수익률과 샤프지수를 각각 계산한다. 10개의 그룹을 각각 테스트 데이터로 이용하여 수익률과 샤프지수의 값들의 평균을 구한 후에, 가장 큰 평균값을 갖는 λ 를 최종 모수로 선택한다. 이때 사용하는 지표인 기대수익률과 샤프지수에 대해 각각의 최적의 모수가 결정된다.

4.4. 성능 평가 지표

모형의 성능을 평가하기 위해 6가지 지표가 사용되었다. $w_t = \beta v_t + \delta_t$ 와 $r_t = (r_{t1}, \dots, r_{tp})^T$ 는 각각 시간 t 에서의 포트폴리오 가중치와 자산의 수익률 벡터를 의미한다. $A_t = \{\delta_{t,j} \mid \delta_{t,j} \neq 0, 1 \leq j \leq p\}$ 는 퍼터베이션 가중치 δ_t 에서 0이 아닌 원소들의 집합을 의미한다. 포트폴리오 성능은 시간 $\tau + 1$ 부터 시간 T 까지 $T - \tau - 1$ 의 기간의 데이터를 기반으로 다음 지표를 활용하여 측정하였다.

(1) 기대수익률: $\mu_w = (1/(T - \tau)) \sum_{t=\tau+1}^T w_t^T r_t$ (Park 등, 2019)

(2) 위험률: $s_w = w_t^T \tilde{r}_t - (1/K) \sum_{t=\tau+1}^K w_{(t)}^T r_{(t)}$; 숏폴의 표본 비모수 추정량 (Bertsimas, 2004).

(3) 샤프지수: μ_w/s_w ; 위험률 대비 기대수익률의 크기 (Sharpe, 1966). 투자의 성과를 판단하는 지수로, 샤프 지수가 커질수록 더 나은 포트폴리오를 의미한다.

(4) 시간 t 에서의 희소성: $|A_t|/p$; 시간 t 에서 포트폴리오를 구성하는 자산 수의 비율 (Park 등, 2019). 포트폴리오를 구성하는 자산이 적을수록 포트폴리오 운용 비용도 감소한다.

Table 1: Expected return using 95% confidence level

Data	Bench	West		North	
	Index	MR	SR	MR	SR
DOW30	0.0953	0.0968 (0.0015)	0.0954 (0.0001)	0.1057 (0.0103)	0.1079 (0.0126)
S&P100	0.1045	0.1041 (-0.0004)	0.1023 (-0.0022)	0.1114 (0.0069)	0.1071 (0.0026)
S&P500	0.1105	0.1082 (-0.0023)	0.1092 (-0.0014)	0.1047 (-0.0058)	0.1115 (0.0010)
FTSE100	0.0331	0.0335 (0.0004)	0.0353 (0.0022)	0.0479 (0.0148)	0.0477 (0.0146)
FTSE250	0.1168	0.1277 (0.0110)	0.1283 (0.0115)	0.1312 (0.0145)	0.1336 (0.0168)
DAX30	0.1164	0.1224 (0.0060)	0.1246 (0.0082)	0.1293 (0.0129)	0.1332 (0.0168)

Data	Bench	West		North	
	Equally	MR	SR	MR	SR
DOW30	0.1449	0.1400 (-0.0048)	0.1400 (-0.0048)	0.1549 (0.0100)	0.1555 (0.0106)
S&P100	0.1398	0.1365 (-0.0033)	0.1377 (-0.0021)	0.1486 (0.0089)	0.1435 (0.0038)
S&P500	0.1401	0.1357 (-0.0044)	0.1376 (-0.0025)	0.1433 (0.0031)	0.1402 (0.0000)
FTSE100	0.1544	0.1490 (-0.0054)	0.1424 (-0.0120)	0.1840 (0.0296)	0.1798 (0.0254)
FTSE250	0.1787	0.1840 (0.0053)	0.1800 (0.0013)	0.1889 (0.0102)	0.1894 (0.0107)
DAX30	0.1242	0.1273 (0.0031)	0.1287 (0.0045)	0.1448 (0.0206)	0.1441 (0.0199)

(5) 시간 t 에서의 안정성: $|A_t \Delta A_{t-1}|/p$; 시간 $t-1$ 와 시간 t 에서의 포트폴리오 자산 집합의 상대적인 차이를 나타내는 지표 (Park 등, 2019). 안정적인 포트폴리오일수록 포트폴리오 거래 비용도 감소한다.

(6) 회전율(turnover): $(1/(T-\tau-1)) \sum_{t=\tau}^{T-1} \sum_{j=1}^p |\delta_i^{t+1} - \delta_i^t|$; $T-\tau-1$ 일의 거래 기간 동안 리밸런싱된 가중치의 평균 크기 (Gerhard and Hess, 2003). 포트폴리오의 가중치 변화가 작을수록 거래 비용 또한 감소한다.

모수 λ 선택과 퍼터베이션 가중치 $\hat{\delta}$ 를 구하는 데 Jegadeesh과 Titman (2001) 및 Bruni 등 (2017), Park 등 (2019)과 마찬가지로 금융 분야의 시계열 예측에서 자주 사용되는 롤링 윈도우(rolling time window) 방법을 사용하였다. 모수 λ 값을 결정하기 위한 교차 검증에서는 표본 내 성능검증(in-sample testing), $\hat{\delta}$ 값을 추정하는데 표본 외(out-of-sample) 예측 방식을 사용하였다. 윈도우 크기는 250일(12개월)로 지정하였으며, 이전 250개 관측치를 활용하여 그 다음 21개(1개월)의 값을 업데이트하였다.

4.5. 실증 데이터 분석을 통한 모형 성능 비교

West와 North 모형의 성능을 평가하기 위해 벤치마크인 인덱스 펀드(market index)와 동일가중 포트폴리오(equally-weighted portfolio)를 기준으로 모형의 성능을 비교하였다. 벤치마크와 비교했을 때 성능이 향상되면 진한 글씨로 표시하였다. Table 1부터 Table 6은 $\alpha = 0.05$ 인 95% 신뢰수준에서 각 모형의 기대수익률,

Table 2: Risk using 95% confidence level

Data	Bench	West		North	
	Index	MR	SR	MR	SR
DOW30	1.0358	1.0059 (0.0299)	1.0033 (0.0325)	1.0623 (-0.0265)	1.0685 (-0.0327)
S&P100	1.0161	0.9888 (0.0272)	0.9690 (0.0471)	1.0105 (0.0055)	0.9876 (0.0284)
S&P500	1.0599	1.0452 (0.0147)	1.0246 (0.0353)	1.0722 (-0.0123)	1.0688 (-0.0089)
FTSE100	1.3511	1.3542 (-0.0031)	1.3441 (0.0070)	1.3702 (-0.0191)	1.3709 (-0.0198)
FTSE250	1.4877	1.4321 (0.0557)	1.4326 (0.0551)	1.4500 (0.0377)	1.4524 (0.0354)
DAX30	1.7171	1.6736 (0.0434)	1.6759 (0.0412)	1.6933 (0.0237)	1.6902 (0.0268)

Data	Bench	West		North	
	Equally	MR	SR	MR	SR
DOW30	1.0105	0.9892 (0.0213)	0.9852 (0.0253)	1.0473 (-0.0367)	1.0479 (-0.0373)
S&P100	1.1120	1.0770 (0.0350)	1.0637 (0.0483)	1.1204 (-0.0084)	1.1166 (-0.0046)
S&P500	1.1441	1.1076 (0.0365)	1.0913 (0.0528)	1.1490 (-0.0049)	1.1454 (-0.0014)
FTSE100	1.4156	1.4127 (0.0029)	1.4112 (0.0044)	1.3912 (0.0244)	1.4589 (-0.0433)
FTSE250	1.3644	1.3017 (0.0626)	1.2978 (0.0666)	1.4492 (-0.0848)	1.4091 (-0.0447)
DAX30	1.7227	1.6873 (0.0354)	1.6887 (0.0340)	1.6900 (0.0327)	1.6892 (0.0335)

위험률, 샤프지수, 안정성, 그리고 회전율을 나타낸다. 또한 모수 λ 선택을 위한 교차 검증에서 사용한 지표를 MR과 SR로 표기하였는데, MR은 기대수익률(mean return)이고 SR은 샤프지수(Sharpe ratio)를 의미한다.

먼저 Table 1에서 인덱스 펀드(Index)를 벤치마크로 두었을 때, DOW30과 FTSE100, FTSE250, DAX30은 West와 North 모두 기대수익률이 벤치마크 대비 증가하였다. 반면 S&P100과 S&P500은 North에서만 기대수익률이 증가하였다. 동일가중 포트폴리오(Equally) 벤치마크에 대해선 FTSE250과 DAX30은 West와 North 모두 기대수익률이 증가한 반면, DOW30과 S&P100, S&P500, FTSE100은 North만 기대수익률이 증가하였다. 즉, West보다 North의 기대수익률이 증가한 빈도가 잦으며, 증가량 또한 더 많다.

Table 2는 인덱스 펀드 벤치마크에 대해서, S&P100과 FTSE250, DAX30은 West와 North 모두 위험률이 벤치마크 대비 감소한 반면, DOW30과 S&P500, FTSE100은 West만 위험률이 감소하였다. 동일가중 포트폴리오를 벤치마크로 활용하였을 때에는, West는 모든 지수의 위험률이 감소하였으나, North는 FTSE100과 DAX30만 위험률이 감소하였다. 전반적으로 모든 데이터의 North보다 West의 위험률이 더 많이 감소하였다.

Table 3에선 벤치마크가 인덱스 펀드일 때, S&P500의 West-MR과 North-MR을 제외한 모든 모형들은 샤프지수가 벤치마크보다 상승하였다. 벤치마크가 동일가중 포트폴리오일 때 West 모형에서는 DOW30과 FTSE100 샤프지수는 감소한 반면, North 모형은 모든 데이터의 샤프지수가 상승하였다. 대부분의 데이터들의 샤프지수가 벤치마크보다 상승하였으므로, West와 North 모형들이 벤치마크보다 더 나은 포트폴리오를

Table 3: Sharpe ratio using 95% confidence level

Data	Bench	West		North	
	Index	MR	SR	MR	SR
DOW30	0.0077	0.0080 (0.0003)	0.0079 (0.0003)	0.0083 (0.0006)	0.0084 (0.0007)
S&P100	0.0086	0.0088 (0.0002)	0.0088 (0.0002)	0.0092 (0.0006)	0.0090 (0.0005)
S&P500	0.0087	0.0086 (-0.0001)	0.0089 (0.0002)	0.0081 (-0.0005)	0.0087 (0.0000)
FTSE100	0.0020	0.0021 (0.0000)	0.0022 (0.0001)	0.0029 (0.0009)	0.0029 (0.0009)
FTSE250	0.0065	0.0074 (0.0009)	0.0075 (0.0009)	0.0075 (0.0010)	0.0077 (0.0011)
DAX30	0.0056	0.0061 (0.0004)	0.0062 (0.0005)	0.0064 (0.0007)	0.0066 (0.0009)

Data	Bench	West		North	
	Equally	MR	SR	MR	SR
DOW30	0.0119	0.0118 (-0.0001)	0.0118 (-0.0001)	0.0123 (0.0004)	0.0124 (0.0004)
S&P100	0.0105	0.0106 (0.0001)	0.0108 (0.0003)	0.0111 (0.0006)	0.0107 (0.0002)
S&P500	0.0102	0.0102 (0.0000)	0.0105 (0.0003)	0.0104 (0.0002)	0.0102 (-0.0000)
FTSE100	0.0091	0.0088 (-0.0003)	0.0084 (-0.0007)	0.0110 (0.0019)	0.0103 (0.0012)
FTSE250	0.0109	0.0118 (0.0009)	0.0116 (0.0006)	0.0109 (-0.0000)	0.0112 (0.0003)
DAX30	0.0060	0.0063 (0.0003)	0.0064 (0.0003)	0.0071 (0.0011)	0.0071 (0.0011)

구성한 것으로 볼 수 있다.

Table 4-6은 West와 North 모형의 회소성과 안정성, 회전율을 벤치마크와 비교하여 표로 나타낸 것이다. Table 4에서 West보다 North의 회소성 값이 더 크지만, 전반적으로 0.09 이하의 값을 갖는다. 이는 전체 자산 중에서 9% 이하의 자산으로만 포트폴리오가 운용되며, 그만큼 포트폴리오 운용 비용도 감소한다는 의미이다.

Table 5 또한 West보다 North의 안정성 값이 더 크지만, 그 값이 0.08을 넘어가지 않는다. 예를 들어 FTSE250의 170개의 자산으로 구성된 포트폴리오에서 North-MR 모형의 안정성이 0.0147일 때, 이전 시간과 비교한 포트폴리오를 구성하는 자산 집합의 차이가 2,499, 즉 3개 이하라는 것이다. 포트폴리오를 구성하는 자산의 변동이 작으므로 포트폴리오 거래 비용 또한 절감할 수 있다.

Table 6도 회소성 및 안정성과 마찬가지로 West보다 North의 회전율 값이 더 크지만, 가장 큰 값이 0.21 이하이다. 이는 전체 거래 기간 동안 조정된 가중치의 평균 크기가 0.21 이하임을 의미한다.

4.6. 신뢰수준의 변화에 따른 모형 성능 비교

신뢰수준의 변화에 따른 모형의 성능을 비교하기 위해 α 값에 따라 7개의 신뢰수준별로 모형의 기대수익률, 위험률, 샤프지수, 회소성, 안정성, 그리고 회전율을 구하였고, 그 값들은 Supplementary materials의 Supplemen-

Table 4: Sparsity using 95% confidence level

Bench	Market index			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0544	0.0631	0.0758	0.0801
S&P100	0.0174	0.0201	0.0269	0.0280
S&P500	0.0191	0.0209	0.0224	0.0236
FTSE100	0.0343	0.0382	0.0358	0.0367
FTSE250	0.0227	0.0256	0.0193	0.0201
DAX30	0.0465	0.0504	0.0724	0.0739

Bench	Equally-weighted portfolio			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0495	0.0520	0.0724	0.0748
S&P100	0.0187	0.0218	0.0242	0.0251
S&P500	0.0156	0.0184	0.0197	0.0204
FTSE100	0.0218	0.0261	0.0264	0.0283
FTSE250	0.0220	0.0259	0.0168	0.0174
DAX30	0.0445	0.0470	0.0709	0.0695

Table 5: Stability using 95% confidence level

Bench	Market index			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0325	0.0325	0.0709	0.0714
S&P100	0.0110	0.0130	0.0277	0.0311
S&P500	0.0116	0.0137	0.0270	0.0270
FTSE100	0.0209	0.0216	0.0362	0.0356
FTSE250	0.0132	0.0148	0.0177	0.0164
DAX30	0.0159	0.0188	0.0685	0.0605

Bench	Equally-weighted portfolio			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0232	0.0222	0.0616	0.0596
S&P100	0.0103	0.0121	0.0256	0.0266
S&P500	0.0095	0.0109	0.0207	0.0227
FTSE100	0.0143	0.0144	0.0261	0.0278
FTSE250	0.0130	0.0132	0.0147	0.0142
DAX30	0.0139	0.0139	0.0615	0.0585

tary Tables A.1–A.24에서 확인할 수 있다. 신뢰수준은 $\alpha = 0.5, 0.3, 0.2, 0.1, 0.05, 0.025$ 에 대해 50%, 70%, 80%, 90%, 95%, 97.5%로 설정하였다. Supplementary table들을 살펴보면 전반적으로 신뢰수준이 변화하더라도 모형의 성능도 그에 따라 달라지는 양상이 보이진 않는다. 그리고 Supplementary Tables A.1–A.24까지의 희소성, 안정성, 회전율은 4.5절의 경우와 마찬가지로 각각 0.09, 0.08, 0.21 이하의 값을 갖는 것을 확인할 수 있다.

Table 6: Turnover using 95% confidence level

Bench	Market index			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0726	0.0929	0.1593	0.1559
S&P100	0.0543	0.0884	0.1612	0.1786
S&P500	0.0613	0.0904	0.1727	0.1712
FTSE100	0.1120	0.1155	0.1544	0.1595
FTSE250	0.1034	0.1025	0.1322	0.1246
DAX30	0.0816	0.0698	0.1921	0.1851
Bench	Equally-weighted portfolio			
	West		North	
Data	MR	SR	MR	SR
DOW30	0.0555	0.0734	0.1996	0.1768
S&P100	0.0610	0.0902	0.1825	0.1910
S&P500	0.0511	0.0836	0.1925	0.2059
FTSE100	0.0553	0.0706	0.1507	0.1731
FTSE250	0.0847	0.0956	0.1476	0.1323
DAX30	0.0733	0.0745	0.1852	0.1942

5. 결론

본 논문에서는 다양한 현실 상황에 적용 가능한 표본 평균-숏폴 포트폴리오 최적화 문제에 perturbation 방법을 접목한 모형을 제안하였고, 다양한 실증 데이터 분석을 통해 그 성능을 검증하였다. 기존의 Markowitz (1952)의 평균-분산 포트폴리오 최적화 모형은 다음과 같은 두 가지 강한 가정 하에서 성립한다. 첫째, 수익률 분포가 타원 분포를 따른다. 둘째, 투자자의 효용 함수는 이차 효용이다. 이러한 가정들은 현실에서 발생 가능한 타원 분포가 아닌 다양한 분포, 그리고 절대 위험 회피가 증가하는 특성을 갖는 위험 회피 투자자의 경우에 모형을 적용하기 어렵다는 한계가 있다. Bertsimas (2004)의 표본 평균-숏폴 최적화 문제는 이러한 가정들로부터 자유로우며, 이상치의 영향을 많이 받는 분산을 하방위험을 대체하였다는 점에서 의의가 있다. 또한 하방위험 중에서 볼록성을 가짐으로써 최적화 문제에서 계산이 용이한 숏폴을 위험 지표로 사용하였다. 이러한 표본 평균-숏폴 최적화 문제에 Park 등 (2019)의 퍼터베이션 방법을 적용함으로써 보다 효율적인 포트폴리오를 구성하는 모형을 제안했다. 퍼터베이션 방법은 West와 North 두 가지로 나뉘는데, West는 위험도를 낮추는 데 목적을 두는 위험도 기반 모형이고 North는 위험도를 어느 정도 유지하면서 기대수익률을 높이는 데 목적을 두는 수익률-위험도 기반 모형이다. 또한 퍼터베이션 방법은 adaptive Lasso 기법을 활용함으로써 포트폴리오를 구성하는 자산의 수와 가중치 변동을 줄이는 안정적인 포트폴리오를 운용한다. 이는 포트폴리오 운용 비용 및 거래 비용과 연관이 있어, 자산의 수가 적고 변동성이 적은 포트폴리오일수록 비용이 감소한다는 의의가 있다. 실증 자료 분석에서는 α 값에 따른 여러 신뢰수준 하에서 모형의 성능을 기대수익률, 위험도, 샤프지수, 회소성, 안정성, 회전율로 나누어 평가하였다. 95% 신뢰수준에 대해선 West보다 North의 기대수익률이 더 많이 증가하는 한편, 위험도는 West가 더 많이 감소하는 것을 확인하였다. 회소성과 안정성, 회전율 또한 West보다 North의 값들이 조금 더 높았지만, 특정 값 이하로 나타남으로써 포트폴리오 비용을 감소시킬 수 있음을 확인하였다. 회전율은 안정성과 회소성에 비해 다소 높은 값을 가지는데, 회전율을 낮추는 것을 향후 과제로 제시한다. 제안된 모형은 여러 가정들의 제약을 받지 않는 모형으로서 현실의 다양한 상황에 적용 가능하다. 또한 최적의 포트폴리오를 얻을 뿐만 아니라, 포트폴리오 운용 및 거래 비용에 대해서도 고려함으로써 보다 현실적인 포트폴리오를 얻을 수 있다는 점에서 의의가 있다.

References

- Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing*, North-Holland, Amsterdam.
- Artzner, P., Delbaen, F., Eber, J., and Heath, D. (1999). Coherent Measures of Risk, *Mathematical Finance*, **9**, 203–228.
- Bertsimas, D., Lauprete, G. J., and Samarov, A. (2004). Shortfall as a risk measure: properties, optimization and applications, *Journal of Economic Dynamics & Control*, **28**, 1353–1381.
- Black, F. (1972). Capital market equilibrium with restricted borrowing, *The Journal of Business*, **45**, 444–455.
- Bruni, R., Cesarone, F., Scozzari, A., and Tardella, F. (2017). On exact and approximate stochastic dominance strategies for portfolio selection, *European Journal of Operational Research*, **259**, 322–329.
- Fishburn, P. C. (1977). Mean-Risk Analysis with Risk Associated with Below-Target Returns, *The American Economic Review*, **67**, 116–126.
- Gerhard, F. and Hess, D. (2003). Multivariate Market Risk Estimators: Reliability and Transaction Costs in the Context of Portfolio Selection, *European Journal of Finance*, **9**, 1–18.
- Grootveld, H., Hallerbach, W. (1998). Variance vs downside risk: Is there really that much difference?, *European Journal of Operational Research*, **114**, 304–319.
- Huang, C. and Litzenberger, R. H. (1988). *Foundations for Financial Economics*. Prentice Hall, New Jersey, 59–62
- Jegadeesh, N. and Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations, *The Journal of Finance*, **56**, 699–720.
- Jorion, P. (2003). Portfolio optimization with tracking-error constraints, *Financial Analysts Journal*, **59**, 70–82.
- Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *The Review of Economics and Statistics*, **47**, 13–37.
- Markowitz, H. (1952). Portfolio selection, *Journal of Finance*, **7**, 77–91.
- Park, S., Song, H., and Lee, S. (2019). Linear programming models for portfolio optimization using a benchmark, *The European Journal of Finance*, **25**, 435–457.
- Park, S., Lee, E., Lee, S., and Kim, K. (2019). Dantzig type optimization method with applications to portfolio selection, *Sustainability*, **11**, 1–32.
- Pratt, J. W. (1964). Risk aversion in the small and in the large, *Econometrica*, **32(1/2)**, 122–136.
- RiskMetrics (1996). Technical Document. New York: Morgan Guarantee Trust Company, Global Research.
- Rockafeller, R. T. and Uryasev, S. (2000). Optimization of conditional value-at-risk, *Journal of Risk*, **2**, 21–42.
- Rockafeller, R. T. (2007). Coherent Approaches to Risk in Optimization Under Uncertainty, *Informa*.
- Roy, A. D. (1952). Safety first and the holding of assets, *Econometrica*, **20**, 431–449.
- Roll, R. (1992). A mean–variance analysis of tracking error, *Journal of Portfolio Management*, **18**, 13–22.
- Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk, *The Journal of Financial*, **19**, 425–442.
- Sharpe, W. F. (1966). Mutual fund performance, *The Journal of Business*, **39**, 119–138.

Received November 3, 2020; Revised December 5, 2020; Accepted December 21, 2020

퍼터베이션 방법을 활용한 평균-숏폴 포트폴리오 최적화

원하연^a, 박세영^{1,a}

^a성균관대학교 통계학과

요약

Markowitz (1952)의 분산투자 모형 발표 이후 포트폴리오 최적화에 대한 많은 연구가 이루어졌다. 마코 위츠의 평균-분산 포트폴리오 최적화 모형은 수익 분포가 정규분포를 따른다는 가정하에서 성립한다. 그러나 실생활에서는 수익 분포가 정규분포를 따르지 않는 경우가 존재한다. 또한 분산은 이상치의 영향을 많이 받는 민감한 지표이다. 이런 분산의 단점을 보완할 수 있는 하방위험인 숏폴(Shortfall)을 위험 지표로 적용함으로써 수익 분포에 대해 최적화가 가능한 평균-숏폴 포트폴리오 모형이 제안되었다. 또한 Jorion (2003)과 Park (2019)은 포트폴리오의 위험도를 최소화하는 동시에 적은 수의 자산으로 구성(sparse)되고 안정적(stable)인 포트폴리오를 얻는 퍼터베이션 방법을 제안하였다. 본 논문에서는 평균-숏폴 포트폴리오 모형에 퍼터베이션 방법과 adaptive Lasso를 적용하여 사용되는 자산의 수가 적으면서 안정적이고 쉽게 적용 가능한 포트폴리오 모형을 제안한다. 그리고 실증 데이터 분석을 통하여 모형의 타당성을 입증한다.

주요용어: 포트폴리오 최적화, 숏폴, 퍼터베이션, 선형 계획법
