Models and Applications Project

Michael Nelson

Introduction

Suppose an investor wishes to invest their entire portfolio in n risky assets and they want to determine the most optimal way to do this. In order to achieve this, the investor may want to consider this problem from an MCDM/MCDA perspective, where there are many techniques/tools that the investor can use to solve this in the best way possible (according to their personal preferences). In this project, we will discuss the paper "Fuzzy Numbers and MCDM Methods for Portfolio Optimization" by Thi T. Nguyen and Lee N. Gordon-Brown, which provides an interesting example of this.

Notation and Terminology

Let $U(\omega)$ respresent an investor's utility given that their terminal wealth is ω. In particular, $U\colon \mathbb{R}_{>0} \to \mathbb{R}$ is an analytic function which is strictly increasing (a typtical preferred utility function is $U(\omega) = \ln \omega$). Let $\boldsymbol{R} = (R_1, \dots, R_n)^\top$ be a vector in $\mathbb{R}^n$ which represents the rate of return of n risky assets and let $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{R}) = (\mu_1, \dots, \mu_n)^\top$. Let $\boldsymbol{w} = (w_1, \dots, w_n)^\top$ be a weight vector representing the proportion of wealth allocated to various assets. In particular, $\boldsymbol{w}$ should satisfy the following constraints:

$$\sum_i w_i = 1 \text{ and } \boldsymbol{w} \geq 0. \tag{1}$$

We regard $\boldsymbol{\mu}$ as being known whereas $\boldsymbol{w}$ being unknown: there are n risky assets that the investor wants their entire portfolio invested in. They can obtain $\boldsymbol{\mu}$ using historical/simulation data however $\boldsymbol{w}$ remains unknown at this point. We shall assume that ω is normalized in the sense that $\omega = 1 + r_{\mathrm{p}}$ where r_{p} is the rate of return of their entire portfolio (with respect to $\boldsymbol{w}$):

$$r_{\mathrm{p}} = \boldsymbol{w}^\top \boldsymbol{R} = \sum_i w_i R_i.$$

In paticular, if r_{p} increases, then their utility $U(\omega)$ increase as well. Since $U(\omega)$ is analytic at $\mathrm{E}(\omega)$, we can express it locally at $\mathrm{E}(\omega)$ in terms of the infinite Taylor series as:

$$U(\omega) = \sum_{k=0}^{\infty} \frac{U^{(k)}(\mathrm{E}(\omega))}{k!} (\omega - \mathrm{E}(\omega))^k \tag{2}$$

Applying the expectation operator to both sides of (2) gives us

$$\mathrm{E}(U(\omega)) = \sum_{k=0}^{\infty} \frac{U^{(k)}(\mathrm{E}(\omega))}{k!} \mathrm{E}((\omega - \mathrm{E}(\omega))^k). \tag{3}$$

The expected utility from an investment in risky assets depends on all central moments, however for numerical purposes, we only consider the first four moments. Thus we approximate $\mathrm{E}(U(\omega))$ using the fourth-order Taylor expansion:

$$\mathrm{E}(U(\omega)) \approx U(\mathrm{E}(\omega)) + \frac{1}{2!} U''(\mathrm{E}(\omega)) \sigma_{\mathrm{p}}^2 + \frac{1}{3!} U'''(\mathrm{E}(\omega)) s_{\mathrm{p}}^3 + \frac{1}{4!} U''''(\mathrm{E}(\omega)) \kappa_{\mathrm{p}}^4.$$

where we set

$$\begin{aligned} \mu_{\mathrm{p}} &= \mathrm{E}(r_{\mathrm{p}}) = \boldsymbol{w}^\top \boldsymbol{\mu} \\ \sigma_{\mathrm{p}}^2 &= \mathrm{E}((r_{\mathrm{p}} - \mu_{\mathrm{p}})^2) = \mathrm{E}((\omega - \mathrm{E}(\omega))^2) \\ s_{\mathrm{p}}^3 &= \mathrm{E}((r_{\mathrm{p}} - \mu_{\mathrm{p}})^3) = \mathrm{E}((\omega - \mathrm{E}(\omega))^3) \\ \kappa_{\mathrm{p}}^4 &= \mathrm{E}((r_{\mathrm{p}} - \mu_{\mathrm{p}})^4) = \mathrm{E}((\omega - \mathrm{E}(\omega))^4) \end{aligned}$$

These are called the **expected return**, **variance**, **skewness**, and the **kurtosis** of the portfolio respectively. We define the (n,n) covariance matrix M_2, the (n,n^2) coskewness matrix M_3, and the (n,n^3) cokurtosis matrix by

$$\begin{aligned} M_2 &= \mathrm{E}((\boldsymbol{R}-\boldsymbol{\mu})(\boldsymbol{R}-\boldsymbol{\mu})^\top) = (\sigma_{ij}) \\ M_3 &= \mathrm{E}((\boldsymbol{R}-\boldsymbol{\mu})(\boldsymbol{R}-\boldsymbol{\mu})^\top \otimes (\boldsymbol{R}-\boldsymbol{\mu})^\top) = (s_{ijk}) \\ M_4 &= \mathrm{E}((\boldsymbol{R}-\boldsymbol{\mu})(\boldsymbol{R}-\boldsymbol{\mu})^\top \otimes (\boldsymbol{R}-\boldsymbol{\mu})^\top \otimes (\boldsymbol{R}-\boldsymbol{\mu})^\top) = (\kappa_{ijkl}), \end{aligned}$$

where $\otimes$ dentoes the Kronecker product and where

$$\begin{aligned} \sigma_{ij} &= \mathrm{E}((R_i-\mu_i)(R_j-\mu_j)) \\ s_{ijk} &= \mathrm{E}((R_i-\mu_i)(R_j-\mu_j)(R_k-\mu_k)) \\ \kappa_{ijkl} &= \mathrm{E}((R_i-\mu_i)(R_j-\mu_j)(R_k-\mu_k)(R_l-\mu_l)) \end{aligned}$$

for all $1 \leq i,j,k,l \leq n$. We calculate the higher momenta of our portfolio using the matrices M_1, M_2, and M_3 defined above as:

$$\begin{aligned} \mu_\mathrm{p} &= \boldsymbol{w}^\top \boldsymbol{\mu} = \sum_i w_i \mu_i \\ \sigma_\mathrm{p}^2 &= \boldsymbol{w}^\top M_2 \boldsymbol{w} = \sum_{i,j} w_i w_j \sigma_{ij} \\ s_\mathrm{p}^3 &= \boldsymbol{w}^\top M_3(\boldsymbol{w}^{\otimes 2}) = \sum_{i,j,k} w_i w_j w_k s_{ijk} \\ \kappa_\mathrm{p}^4 &= \boldsymbol{w}^\top M_4(\boldsymbol{w}^{\otimes 3}) = \sum_{i,j,k,l} w_i w_j w_k w_l \kappa_{ijkl}. \end{aligned}$$

The central moments of the portfolio gives us a lot of information in regards to how the portfolio is structured. For instance, a high value of σ_{ii} indicates that asset i has high volatility or has high risk, whereas a low value of σ_{ii} indicates that asset i has low volatility or has low risk. A negative value of σ_{ij} where $i \neq j$ indicates that the values of the assets i and j move in opposite directions (which is a desirable feature in a diversified portfolio), and a positive value of σ_{ij} where $i \neq j$ indicates that the values of the assets i and j move in the same direction (which often occurs with stocks in companies in the same industry). The skewness and kurtosis have their own interpretations as well. Note that the μ_i, the σ_{ij}, the s_{ijk}, and the κ_{ijkl} are all regarded as being known since they can obtain them through simulation or historical data, however the central moments of the portfolio are unknown since they are functions of the variable $\boldsymbol{w}$.

Marginal Impact

The **marginal impact** of the asset i to the porfolio's return, variance, skewness, and kurtosis is defined to be the ith component of the vectors given below:

$$\begin{aligned} \nabla_{\boldsymbol{w}} \mu_\mathrm{p} &= \boldsymbol{\mu} \\ \nabla_{\boldsymbol{w}} \sigma_\mathrm{p}^2 &= 2M_2\boldsymbol{w} \\ \nabla_{\boldsymbol{w}} s_\mathrm{p}^3 &= 3M_3\boldsymbol{w}^{\otimes 2} \\ \nabla_{\boldsymbol{w}} \kappa_\mathrm{p}^4 &= 4M_4\boldsymbol{w}^{\otimes 3}. \end{aligned}$$

Assets with higher marginal impact will make large changes to the moments of the portfolio relative to small changes of the weight of that asset; thus they will have *more* influence to the overall portfolio compared to the other assets. Note that the marginal impact of asset i to the whole portfolio return is given by the expected return μ_i which does not depend on the other assets j where $j \neq i$, however the marginal impact of asset i to the higher moments will typically take into account the other assets j where $j \neq i$. For instance, since the σ_{ij} are symmetric in i and j, the portfolio's variance can be expressed as

$$\sigma_\mathrm{p}^2 = \sum_{1\leq i\leq n} w_i^2 \sigma_i^2 + 2 \sum_{1\leq i<j\leq n} w_i w_j \sigma_{ij}.$$

The second term on the right represents the diversification effect of the overall portfolio's variance. The variance marginal impact of asset i to the whole portfolio's variance is given by

$$\partial_{w_i} \sigma_\mathrm{p}^2 = 2 \sum_{1\leq j\leq n} w_j \sigma_{ij},$$

which clearly depends on all assets in the portfolio. For each criteria (return, variance, skewness, and kurtosis) we calculate two weight vectors related to minimum and maximum circumstances. We then use them to calculate the *exact* marginal impacts of the assets to the portfolio momenta using the partial derivative formulas above. Thus we can calculate the marginal impacts of each asset in the extreme cases, however the exact contribution of an asset on the portfolio's higher moments is uncertain before choosing a weight vector $\boldsymbol{w}$. We handle this problem using fuzzy numbers.

Optimal Solutions in Extreme Cases

For each $j = \{2, 3, 4\}$ corresponding to variance, skewness, and kurtosis, we set $\boldsymbol{w}^j_{\min}$ and $\boldsymbol{w}^j_{\max}$ to be an optimal objective solutions to the following mathematical programs:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{w}^\top M_j(\boldsymbol{w}^{\otimes j}) \\ \text{subject to} & \sum_i w_i = 1 \text{ and } \boldsymbol{w} \geq 0, \end{array} \qquad \begin{array}{ll} \text{maximize} & \boldsymbol{w}^\top M_j(\boldsymbol{w}^{\otimes j}) \\ \text{subject to} & \sum_i w_i = 1 \text{ and } \boldsymbol{w} \geq 0. \end{array}$$

Moreover, for each $i \in \{1, \ldots, n\}$ we set

$$\varepsilon_{ij} = jM_j(\boldsymbol{w}^j_{\min})^{\otimes(j-1)} \quad \text{and} \quad \delta_{ij} = jM_j(\boldsymbol{w}^j_{\max})^{\otimes(j-1)}.$$

In particular, $\boldsymbol{w}^2_{\min}$ is the weight vector which gives the smallest portfolio variance, and $\varepsilon_{i,2}$ is the marginal impact of asset i to the portfolio variance at the extremal point $\boldsymbol{w}^2_{\min}$ where portfolio variance is minimized.

Fuzzy Numbers

A **fuzzy set** A is a pair (X, f) where X is a subset of $\mathbb{R}$ and f is a function from X to $[0, 1]$. We denote $f = f_A$ and call it the **membership function** of A. For each $x \in X$, the value $f_A(x)$ is called the **grade** of membership of x in A. Intuitively, if $f_A(x) = 1$, then we think of x as *fully* belonging to the fuzzy set A, and if $f_A(x) = \varepsilon$ where $0 < \varepsilon < 1$ then we think of x as *partially* belonging to A. We set

$$\begin{aligned} A^{\geq\alpha} &:= \{f_A \geq \alpha\} \\ A^{>\alpha} &:= \{f_A > \alpha\} \\ \operatorname{Supp}(A) &:= \{f_A > 0\} \\ \operatorname{Core}(A) &:= \{f_A = 1\} \end{aligned}$$

We call these the **α-cut** of A, the **strict α-cut** of A, the **support** of A, and the **core** of A respectively. We say A is **normal** if $\operatorname{Core}(A) \neq \emptyset$. In this paper, the authors only considered fuzzy sets of the form $A = (\mathbb{R}, f)$ where the membership function is given by

$$f_A(x) = \begin{cases} f^\ell_A(x) & a \leq x \leq b \\ \theta & b \leq x \leq c \\ f^r_A(x) & c \leq x \leq d \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $a \leq b \leq c \leq d$, where $\theta \in (0, 1]$, where $f^\ell_A \colon [a, b] \to [0, \omega]$ is increasing, and where $f^r_A \colon [c, d] \to [0, \omega]$ is decreasing. The authors refer to these fuzzy sets as **fuzzy numbers** (this isn't standard terminology in the literature). In the special case where $\theta = 1$, $f^\ell_A(x) = (x - a)/(b - a)$ and $f^r_A(x) = (d - x)/(d - c)$, then we call this a **normal trapezoidal fuzzy number** and we denote it by $A(a, b, c, d)$. They are called "trapezoidal" since the graph of f_A takes the shape of a trapezoid. In order to compare fuzzy numbers with other fuzzy numbers, we use their representative crisp number, which is obtained via the centroid-based defuzzification method. More specifically, let $A = A(a, b, c, d)$ be a normal trapezoidal fuzzy number. Then we set

$$\widetilde{x}(A) = \frac{1}{3}\left(a + b + c + d - \frac{cd - ab}{(c + d) - (a + b)}\right) \quad \text{and} \quad \widetilde{y}(A) = \frac{1}{3}\left(1 + \frac{c - b}{(c + d) - (a + b)}\right).$$

Centroids on the horizontal axis are used as a basis to evaluate assets. If horizontal coordinates $\widetilde{x}$ of all assets in the portfolio ar equal then we use the vertical centroid coordinates $\widetilde{y}$ instead, though this situation seldom occurs in practice where the numbers of assets is large enough. In particular, he representative location on the horizontal axis is more important than the average height in comparing fuzzy numbers.

Modeling Marginal Impacts with Fuzzy Numbers

Fix a proportion parameter $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)^\top \in [0,1]^4$ which represents bias level towards preferred extremes. For each $i \in \{1, \ldots, n\}$ (corresponding to the n assets) and for each $j \in \{1,2,3,4\}$ (coresponding to return, variance, skewness, kurtosis), let $A_{ij} = A_{ij}(a_{ij}, b_{ij}, c_{ij}, d_{ij})$ be the normal trapezoidal fuzzy number where $a_{ij} = \min\{\varepsilon_{ij}, \delta_{ij}\}$ and $d_{ij} = \max\{\varepsilon_{ij}, \delta_{ij}\}$ and where if $j \in \{2,3\}$ then we set

$$b_{ij} = \min\left\{\frac{\varepsilon_{ij}+\delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij}+\delta_{ij}}{2}\right), \varepsilon_{ij}\right\}$$
$$c_{ij} = \max\left\{\frac{\varepsilon_{ij}+\delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij}+\delta_{ij}}{2}\right), \varepsilon_{ij}\right\}$$

and if $j \in \{1,3\}$, then we set

$$b_{ij} = \min\left\{\frac{\varepsilon_{ij}+\delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij}+\delta_{ij}}{2}\right), \delta_{ij}\right\}$$
$$c_{ij} = \max\left\{\frac{\varepsilon_{ij}+\delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij}+\delta_{ij}}{2}\right), \delta_{ij}\right\}$$

We set $x_{ij} := \widetilde{x}(A_{ij})$ and $y_{ij} := \widetilde{y}(A_{ij})$.

Example 0.1. The authors give an example of what A_{i2} looks like in the case where $a_i = \varepsilon_{ij}$ and $d_i = \delta_{ij}$:

This circumstance is illustrated by Fig. 1.

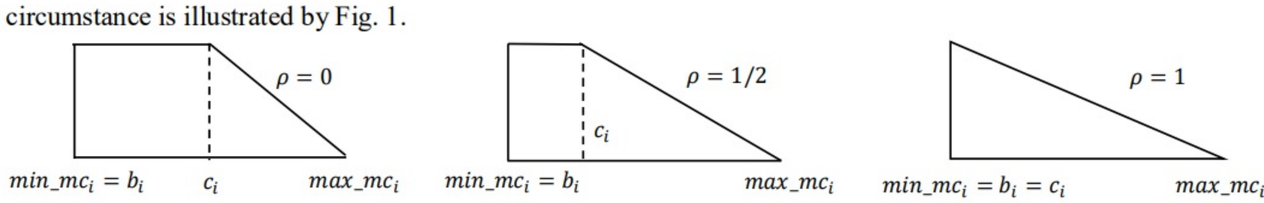

Fig. 1. Variance and kurtosis fuzzy numbers when $a_i = min\_mc_i$ and $d_i = max\_mc_i$

Modeling Marginal Impacts by Fuzzy Numbers

In general, a rational investor would typically prefer to maximize return and skewness and minimize variance and kurtosis There are many ways to do this depending in the investor's preference utility. To this end, we consider a weighting scheme

$$\mathrm{s} = (\mathrm{s}_r : \mathrm{s}_v : \mathrm{s}_s : s_\kappa)$$

which measures the investors preference. For instance, the scheme $(1:1:0:0)$ indicates that the investor is only concerned about return and variance whereas they are not concerned about skewness and kurtosis. The scheme $(4:3:2:1)$ indicates that the investor is concerned about all four moementa, however they favor return the most, variance the second-most, skewness the third-most, and favors kurtosis the least. For experimental purposes, the authors considered the following weighting schemes:

TABLE I
INVESTOR PREFERENCE WEIGHTING SCHEMES USED FOR EXPERIMENTS

Weighting schemes $\boldsymbol{s = (s_r : s_v : s_s : s_k)}$	(2:1:2:1)	(1:2:1:2)	(4:3:2:1)	(1:2:3:4)	(1:1:0:0)	(0:0:1:1)
Normalized ratios $\boldsymbol{\overline{s} = (\overline{s_r}, \overline{s_v}, \overline{s_s}, \overline{s_k})}$	(1/3, 1/6, 1/3, 1/6)	(1/6, 1/3, 1/6, 1/3)	(2/5, 3/10, 1/5, 1/10)	(1/10, 1/5, 3/10, 2/5)	(1/2, 1/2, 0, 0)	(0, 0, 1/2, 1/2)
$\boldsymbol{\rho}$ **in fuzzy number designs**	(1, 1/2, 1, 1/2)	(1/2, 1, 1/2, 1)	(1, 3/4, 1/2, 1/4)	(1/4, 1/2, 3/4, 1)	(1, 1, 0, 0)	(0, 0, 1, 1)

Here, $\overline{\mathrm{s}} = \mathrm{s}/\|\mathrm{s}\|$ and $\boldsymbol{\rho} = \mathrm{s}/\mathrm{s}_{\max}$ where $\|\mathrm{s}\|^2 = \mathrm{s}_r^2 + \mathrm{s}_v^2 + \mathrm{s}_s^2 + \mathrm{s}_\kappa^2$ and $\mathrm{s}_{\max} = \max\{\mathrm{s}_r, \mathrm{s}_v, \mathrm{s}_s, \mathrm{s}_\kappa\}$. In particular, if $\rho_v = 1$ means the investor is most concerned about variance whereas $\rho_v = 0$ means the investor is least concerned about variance at all.

The SAW Method Approach to Portfolio Selection

We are now ready to describe the SAW method approach to portfolio selection.

The Decision Matrix

Fix a proportion parameter $\boldsymbol{\rho} = (\rho_1, \rho_2, \rho_3, \rho_4)^\top$ which represents bias level towards preferred extremes. For each $i \in \{1, \ldots, n\}$ (corresponding to the n assets) and for each $j \in \{1, 2, 3, 4\}$ (coresponding to return, variance, skewness, kurtosis), let $A_{ij} = A_{ij}(a_{ij}, b_{ij}, c_{ij}, d_{ij})$ be the normal trapezoidal fuzzy number where $a_{ij} = \min\{\varepsilon_{ij}, \delta_{ij}\}$ and $d_{ij} = \max\{\varepsilon_{ij}, \delta_{ij}\}$ and where if $j \in \{2, 4\}$ then we set

$$
\begin{aligned}
b_{ij} &= \min\left\{\frac{\varepsilon_{ij} + \delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij} + \delta_{ij}}{2}\right), \varepsilon_{ij}\right\} \\
c_{ij} &= \max\left\{\frac{\varepsilon_{ij} + \delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij} + \delta_{ij}}{2}\right), \varepsilon_{ij}\right\}
\end{aligned}
$$

and if $j \in \{1, 3\}$, then we set

$$
\begin{aligned}
b_{ij} &= \min\left\{\frac{\varepsilon_{ij} + \delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij} + \delta_{ij}}{2}\right), \delta_{ij}\right\} \\
c_{ij} &= \max\left\{\frac{\varepsilon_{ij} + \delta_{ij}}{2} + \rho_j\left(\varepsilon_{ij} - \frac{\varepsilon_{ij} + \delta_{ij}}{2}\right), \delta_{ij}\right\}
\end{aligned}
$$

Example 0.2. The authors give an example of what A_{i2} looks like in the case where $a_i = \varepsilon_{ij}$ and $d_i = \delta_{ij}$:

This circumstance is illustrated by Fig. 1.

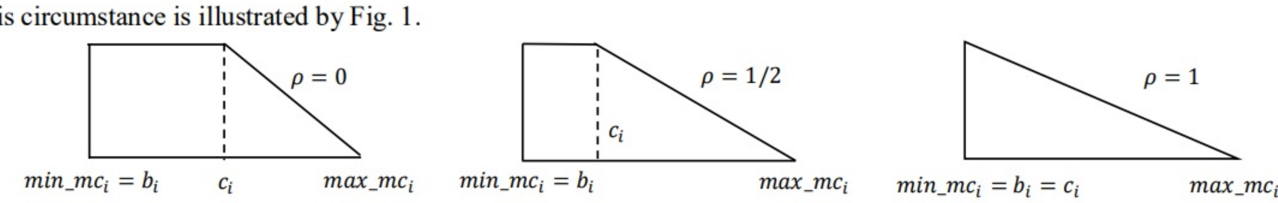

Fig. 1. Variance and kurtosis fuzzy numbers when $a_i = min\_mc_i$ and $d_i = max\_mc_i$

We set $x_{ij} := \widetilde{x}(A_{ij})$ and let $D = (x_{ij})$ be the $n \times 4$ decision matrix given below:

$$
D = \begin{pmatrix} x_{11} & \cdots & x_{14} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{n4} \end{pmatrix}.
$$

Thus the ith row of D corresponds to the ith asset which the investor chooses to allocate their initial wealth and the jth column of D corresponds to the jth criterion (return, variance, skewness, and kurtosis) with which asset performance is measured. The x_{ij} are the x-coordinates of the centroid of the trapezoidal fuzzy number which represents the maginal impact that the ith a sset has on the jth criteria. The SAW method requires a comparable scale for all elements in the decision matrix. This will be a normalized matrix $R = (r_{ij})$ which we define as follows: let ε_j and δ_j denote the minimum and maximum values in the jth column. For benefit criteria (meaning the larger the rating, the greater the preference), we set

$$
r_{ij} = \frac{x_{ij} - \varepsilon_j}{\delta_j - \varepsilon_j}.
$$

For cost criteria (meaning the smaller the rating, the greater the preference), we set

$$
r_{ij} = \frac{\delta_j - x_{ij}}{\delta_j - \varepsilon_j}.
$$

The weight of each criterion is obtained from the investor's normalized preference vector $\overline{\mathrm{s}} = \mathrm{s}/\|\mathrm{s}\|$. The performance score p_i of the ith asset is given by

$$
p_i = \sum_j \overline{\mathrm{s}}_j r_{ij}.
$$

Finally, set $\boldsymbol{w}_{\mathrm{s}} = \boldsymbol{p}/\|\boldsymbol{p}\|$ where $\boldsymbol{p} = (p_1, \ldots, p_n)^\top$. Then the investor whose preference is given by the s should the weight vector $\boldsymbol{w}_{\mathrm{s}}$.

Conclusion

In the paper, the authors devolped an interesting MCDA/MCDM approach to portfolio optimization. I found their use of fuzzy numbers to be very innovative. The fuzzy numbers were designed in a such a way so that the portfolio would be skewed towards the investor's extremes. For instance, if an investor's scheme was $s = (1 : 1 : 0 : 0)$, meaning they only wanted to maximize returns and minimize variance as much as possible and were not concerned with the higher moments, then this investor should use the weight vector $\boldsymbol{w}_s$ for their portfolio, where $\boldsymbol{w}_s$ is the weight vector obtained from SAW method. In