JEBI

Journal of Economics and Business Issues

The Effect of Correlation in the CVaR Algorithm

Reza Habibi[1,*], Somayeh Maddapour[2]

[1] Iran banking Institute; r_habibi@ibi.ac.ir
[2] Iran banking Institute; smadadpour@gmail.com
* habibi1356@gmail.com

Abstract: According to existing literature, the CVaR (conditional value at risk) method outperforms, therefore in this study CVaR is applied as a constraint to transform portfolio optimization problem into a linear one. Linearization of portfolio optimization plays a central role in financial studies, since linear problem allows for performing sensitivity analysis. Sensitivity analysis makes it possible to measure the variation of parameters due to the variation of one parameter in a linear problem, without solving the problem from the scratch. The objective function coefficient of mentioned method for a portfolio includes average of asset returns, which are highly correlated. Therefore, principal component analysis is employed to tackle the correlation of the functional relations. An example of stock market is employed to validate method. Finally, it is shown that the result of the presented method is closer to the ideal result.

Keywords: Correlated parameters, Sensitivity analysis, Linear portfolio optimization, Principal component analysis

Citation: Habibi, R. & Maddapour, S. (**2024**) The Effect of Correlation in the CVaR Algorithm. *Journal Of Economic and Business Issues*, *4(2)*, 35-45.

Received: 14/03/2024
Accepted: 04/07/2024
Published: 31/07/2024

1. Introduction

Studies about portfolio optimization could be categorized based on its risk measurement. The initial model uses variance as a risk measurement, thus considered as a quadratic programming problem. Regarding its compact covariance matrix there is computational difficulty associated with. Many scholars had attempted toward linearization of portfolio because linear problem have some advantages over the non-linear one. The related software and solution for a linear problem are effortless. Moreover, it makes the sensitivity analysis possible. As an alternative, the mean absolute deviation (MAD) that was first initiated by sharp (1971) was introduced to alleviate the shortcoming of initial model and plus takes its advantage. This method allows that large-scale optimization problem being solved on a real time. However, the assumption of dies-utility, which prevents detecting of risk from larger losses, reduces the popularity of this method.

CVaR method presented by Uryasev (2002) was highly populated due to the following advantages. It has been shown that CVaR is a convex risk measurement (Uryasev, 2002). The minimization of CVaR on the other hand leads to a portfolio with a small VaR. Moreover, CVaR could be involved either as a subjective function or as a constraint (Uryasev, 2002).

After proposition of linear portfolio optimization, the question raised that which algorithm has to be applied in order that the answers become more reliable. The simplex method finds an optimal solution for the LP problem, which has entered the algorithm. After the simple method, interior point methods (IPMs) were introduced to solve large-scale problems. This method solves problems in polynomial time. Simplex algorithm outperformed interior point methods (Potra, 2018). Therefore, this method has been widely used in portfolio optimization problems (Lim, 2020).

Later, the importance of measuring the perturbation of other parameters because of change in one parameter was considered. Such questions are addressed through sensitivity analysis (Bazaraa, 2009). However, the following sensitivity analysis techniques do not consider the correlation among parameters.

The ordinary sensitivity analysis proposed by Koltai and Terlaky (2019) is only solvable by simplex method and no simultaneous changes of OFC or RHS is allowed. The problem is tackled through two distinctively different methods. Each of these methods has some drawbacks. The 100% rule does not assure whether the optimal solution remain unchanged if this rule does not satisfy or not, and the tolerance approach could not be applied to large-scale problems due to the small range. The shortcoming of these methods intrigued. This method considered individual percentage change for every RHS parameters or OFC. However, small tolerance and no simultaneous perturbation among parameters are among weak points of this method. It is attempted to combine different sensitivity analysis through a parametric system. The Parametric programming method is useful when RHS parameters or OFC depend only on one parameter. The weak point of this study is that it is only applicable for diverse sensitivity analysis (Shahin, 2016). After that, Hladík (2021) presented another method, which was computationally attractive and applicable to broader cases. However, this method ignores the correlation among parameters. To eliminate the shortcomings of ordinary sensitivity analysis Hladík (2021) represented a noble method that addresses the correlation among parameters. The method then followed by Shahin 2016. In this study Shahin used Principal component analysis (PCA) to account for correlation among parameters.

PCA method in some research is applied as a creator of factors (Victor, 2021; Heij, 2021). This is also a useful method in portfolio optimization. They used the forenamed method for portfolio optimization problem and showed the superiority of using PCA method in portfolio optimization to portfolio optimization without using PCA. PCA method diversifies portfolio; thus, the risk decreases because stocks are chosen from those with lower risks.

In addition, sensitivity analysis in portfolio optimization has been concerned in a number of studies (Arbaiy, 2019). Nonetheless, neither of them has considered correlation among stock's price.

In this paper we aim at performing sensitivity analysis for a linear portfolio optimization problem under CVaR method and simplex algorithm. CVaR method objective function consists of the average returns of the stocks and these parameters are highly correlated. To the best of authors' knowledge, among the previous studies there is no study to deal with this problem. We will use the PCA method to tackle the problem and then the sensitivity analyses based on the related formulas are computed.

The paper is organized as follows. In section 2 the methodology of sensitivity analysis in the presence of correlation among parameters is introduced. Section 3 illustrates the details of sensitivity analysis details while considering correlation among OFC. Section 4 presents portfolio optimization using CVaR as a constraint. In Section 5, the sensitivity analysis of real example of portfolio optimization is examined. Finally, the conclusion is presented in sections 6.

2. Materials and Methods

In this paper, to construct a linear portfolio under CVaR method, first, the historical data of four stocks (MNST, MAR, FISV, SON.SG) are derived from the related websites and then the logarithmic returns are calculated. Next, the returns divided into two parts the latest (newest) data, which is regarded as future data, and the historical data. For solving portfolio optimization with CVaR method some scenarios are needed. We use Mina & Yi Xiao method to generate scenarios, which is a historical data simulation. The historical monthly prices are used to derive monthly returns. The problem is solved without the latest data.

Next, we use the covariance matrix as the input of a multivariate statistical method called principal component analysis (PCA) in order to convert correlated parameters (OFC) into independent ones, introduced by some functional relations. Because the objective function constructs of average of returns; thus, the PCA should be applied over average of data. The historical returns for each stock in each 3 months put into one group -season data- and the average of each group is computed. Then the difference between historical return of one stock with and without the latest data is calculated. For the rest of the stocks, the deviation is calculated by the help of the equation 5.

Then the 100% rule should be applied to see whether the basic variables remain unchanged. If the 100% rule is satisfied, we know that the current optimal solution still remains optimal even though all the RHS parameters or OFC have changed. Then to evaluate the volatility and the prediction capability of this method, the result is compared with two other problems. The first one as we call the ideal result includes the whole data. Second, we consider a case in which instead of latest price, the historical price average is implied except for one stock which takes its latest data because we assume we know one stock price and not the rest of the stock price. To obtain it, the historical price of data without the latest data is considered and then average of each stock is calculated. Next, the result of linear portfolio optimization as before is derived. Finally, these three results are compared. The following flowchart shows the conceptual framework of this paper.

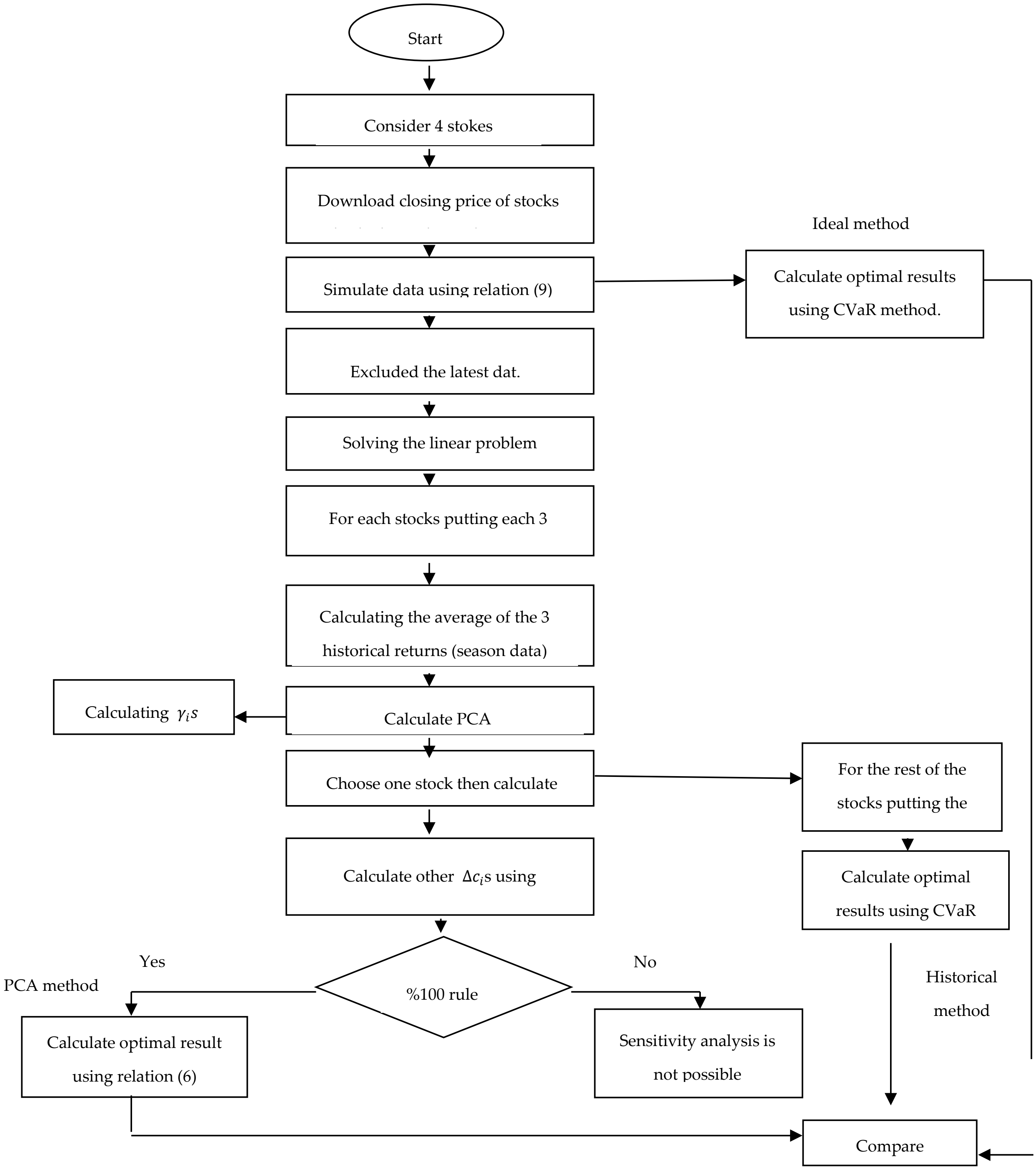

Fig. 1. The flowchart of sensitivity analysis in the presence of correlation

Source: Author's research

3. Sensitivity analysis:

In this section, we briefly discuss the formulas of sensitivity analysis of LP problem in the presence of correlation among OFC by Shahin (2016). It should be noted that these formulas rely upon a series of assumptions. First, they are valid for local perturbation, that is, the acceptable range for parameters change should be small and within ε-neighborhood of the estimated parameters. Second, formulas are only applicable when a basic optimal non-degenerate solution is available. The following LP program is considered as a basis on which other formulas are derived.

$$\begin{aligned} &\text{Min } \boldsymbol{c}^T \boldsymbol{x} \\ &\text{s.t. } \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\ &\boldsymbol{x} \geq \mathbf{0}, \end{aligned} \quad (1)$$

where $\boldsymbol{x}$ is a vector with n variables, **A** is an m*n matrix, $\boldsymbol{c}$ is the OFC vector with n variables, and $\boldsymbol{b}$ is the right hand side (RHS) vector with m parameters.

If equation (1) is solved, then the optimal value and the optimal solution are calculated as the following:

$$\begin{aligned} z^* &= \boldsymbol{c}_B^T \boldsymbol{B}^{-1}\boldsymbol{b}, \\ \boldsymbol{x}_{\boldsymbol{B}}^* &= \boldsymbol{B}^{-1}\boldsymbol{b}, \end{aligned} \quad (2)$$

where z^* is the optimal value of the objective function, $\boldsymbol{c}_{\mathbf{B}}$ is the OFC of basic variable and variable $\boldsymbol{x}^*_{\mathbf{B}}$ is the optimal solution. The subscript **B** denotes basic variables and the superscript $*$ indicates the optimal value. In addition, **B** is the constraint matrices of basic and N is the constraint matrix of non-basic variables.

If there is not a correlation between the components of vector c, we will have:

$$\frac{dz^*}{dc_k} = \begin{cases} x^*_{B_i}, & if\ c_k = c_{B_i} \\ 0\ , & if\ c_k \ \text{corresponds to nonbasic variables} \end{cases} \quad (3)$$

In the presence of functional relation between OFC parameters, Shahin (2014) extended the following formulas based on the PCA method in order to change dependent parameters into the independents parameters:

$$\Delta\overline{c}_j = \left[\frac{\gamma_{11}}{\gamma_{1j}} + \frac{\gamma_{21}}{\gamma_{2j}} + \cdots + \frac{\gamma_{n1}}{\gamma_{nj}}\right] \quad (5)$$

$$\Delta z^* = \left[\frac{\partial z^*}{\partial \overline{c_1}} + \frac{\partial z^*}{\partial \overline{c_2}}\left(\frac{\gamma_{11}}{\gamma_{12}} + \frac{\gamma_{21}}{\gamma_{22}} + \cdots + \frac{\gamma_{n1}}{\gamma_{n2}}\right) + \cdots + \frac{\partial z^*}{\partial \overline{c_n}}\left(\frac{\gamma_{11}}{\gamma_{1n}} + \frac{\gamma_{21}}{\gamma_{2n}} + \cdots + \frac{\gamma_{n1}}{\gamma_{nn}}\right)\right]\Delta\overline{c}_1 \quad (6)$$

Here γ_{ij}'s are the entries of the j-th eigenvector, which is calculated through PCA method.

4. CVaR Algorithm:

In this section, the CVaR algorithm is discussed. The portfolio optimization under CVaR method is solved by using scenario generation method. The optimal weights of stocks and optimal value of portfolio could be derived afterwards. Then by using the formulas which are presented in section 3, the portfolio sensitivity analysis based on the small changes in OFC which are mean return of stock's price will be performed and portfolio sensitivity analysis based on the a small changes of each stocks return is determined.

The first step of the CVaR calculation is to find the matrix of historical returns from the matrix of historical prices. Mina and Yi Xiao (2021) historical simulation method is applied, which considers logarithmic returns are applied. Logarithmic returns is the preferred method for return calculations in finance and it will make calculations simpler in later stages of the thesis. The general formula for logarithmic returns is as the following:

$$r_{it} = \ln\left(\frac{p_{it}}{p_{it-1}}\right) \quad (8)$$

where r_{it} is the return of stock i in day t and p_i indicates the initial price of the security, whereas P_{i+1} is the price in the next period.

Then by using formula (9), the scenarios of the next period of stock's price will be calculated based on the historical monthly returns i.e. r_{it-1}, r_{it-2}, ...

$$y_{ij} = q_i * exp(r_{it-j} * \sqrt{t}) \quad j = 1,2, \dots J \quad (9)$$

Here y_{ij} is the t+1 or the next-month price of stock i in the scenario j also is a random variable and q_i is price of stock i in month t.

The expected end-of-period $(t+1)$ stock i is derived from the following equation:

$$E[y_i] = \sum_{j=1}^{J} \pi_j \, y_{ij} = \frac{1}{J}\sum_{j=1}^{J} y_{ij} \quad (10)$$

We assume that all scenarios have equal probability.

In an optimization problem, CVaR can be considered as an either objective or constraint. If CVaR is considered as the objective, the risk of the portfolio measured by CVaR will be minimized based on the given required expected return. If CVaR is considered as a constraint, the expected return of the portfolio will be maximizing based on the given level of risk.

In this research, CVaR will be treated as the constraint of the portfolio optimization. Therefore, the objective of the optimization problem is to maximize the portfolio expected return. In other words, we could change the objective function into minimization by adding minus and minimize the portfolio expected loss based on the certain level of risk of the risk.

In this case, the portfolio optimization problem would be:

$$\min_{X,\zeta} \sum_{i=1}^{n} -E[y_i]x_i \quad (11)$$

subject to.

$$\zeta + (1-\alpha)^{-1}\sum_{j=1}^{J} \pi_j e_j \le w \sum_{i=1}^{n} q_i x_i^0 \quad (12)$$

$$e_j \ge \sum_{i=1}^{n}\left(-y_{ij}x_i + q_i x_i^0\right) - \zeta. \quad e_j \ge 0. \; j = 1, \dots, J \quad (13)$$

$$q_i x_i \le \upsilon_i \sum_{k=1}^{n} q_k x_k . \quad i = 1, \dots, n \quad (14)$$

where i is the number of stocks, E$[yij]$ is the average of stock's return in all scenarios, $\boldsymbol{x_i}$ is the number of stock i in portfolio, *w* is the coefficient of risk tolerance, $\boldsymbol{x_i^0}$ is the number of stock i in the initial portfolio, ej is the coefficient for changing CVaR into the linear variable, qi is the price of stock i at the end of month in scenario j and ζ is VaR.

5. Numerical examples:

In order to exemplify and mixed two proposed methods, a real example of stock market is solved in this section. Then, the results are compared with the results of two other problems where correlations among parameters have not been considered.

The data-sets used in this paper are historical monthly close prices of 4 stocks (SON.SG, FISV, MAR, MNST) from February 1, 2014 to September 1, 2015. The latest prices are excluded and the problem is solved by using CVaR method. For solving portfolio optimization with CVaR method, some scenarios are required. Thus, Mina & Yi Xiao method is used to generate scenarios. First, the historical monthly prices are used to derive monthly returns. Twenty scenarios are derived from according to equation (9) for each stock. We assume that all scenarios have equal likelihood. Then based on the CVaR algorithm, the portfolio optimization problem is solved with lingo software. The basic variables are shown in the following table:

Table 1: The basic value

ζ	e(20)	X(MNST)	X(FISV)	X(MAR)
692.33	28.53	13.87	22.53	13.59

Source: Author's research

Because objective function includes mean y, each of three monthly scenarios is put in one group, then the average of each group is computed. We regard these data as season data which are to generate PCA. To do so, first covariance matrix between season data is calculated and then eigen-value and eigenvector are derived.

Table 2: Eigen-values of season data

Component	Eigen-value
1	0.031548103
2	6.064279805
3	24.62035181
4	1187.936196

Source: Author's research

It turns out; the Eigen-value of the fourth element is notably bigger than other elements. Therefore, based on the equation (15) it can be concluded that 0.97 percent of data are explained by the fourth element.

$$s_1^2 + s_2^2 + \cdots + s_p^2 = l_1 + l_2 + \cdots + l_p \quad (15)$$

where s represents variance and $\boldsymbol{l}$ is the eigenvalue.

Table 3: Eigenvectors of season data

Eigenvectors				
Variable	1	2	3	4
SON1.SG	-0.85846	0.287165	0.380537	0.189148
MNST	0.184078	-0.32384	0.210371	0.903872
FISV	0.311176	0.898522	-0.11826	0.286081
MAR	-0.36378	-0.07294	-0.89272	0.255728

Source: Author's research

To derive PCA, eigenvectors should be arranged in descending order based on the Eigen-value.

Table 4: PCA of data

γ_{ij}	j=1	j=2	j=3	j=4
i=1	0.189148	-0.38054	0.287165	0.858458
i=2	0.903872	-0.21037	-0.32384	-0.18408
i=3	0.286081	0.118264	0.898522	-0.31118
i=4	0.255728	0.89272	-0.07294	0.363778

Source: Author's research

Then for a selected stock, for the selected stock the difference between $\bar{y}$ with or without latest data is calculated, it is notable that this value should be in acceptable range, then the variation of other stock is calculated based on formula 5. These data should also be in acceptable range. We consider stock 4, the variation between two $\bar{y}$ is -0.242 and variation of other stocks are calculated using equation 5. The result is shown in table 6:

Table 5: Objective Coefficient Range (from lingo output)

Variable	Allowable Increase	Allowable Decrease
X_SON1.SG	INFINITY	58.96188
X_MNST	37.53784	INFINITY
X_FISV	2.339523	INFINITY
X_MAR	115.8636	2.22285

Source: Author's research

Table 6: The variation of correlated parameters

Stock (i)	Δc_i
i=1	0.199453733
i=2	1.382213411
i=3	-1.74051599

Source: Author's research

All of the above changes are in the accepted range for which the basis is unchanged. Now 100% rule should be examined. If it is satisfied, we can calculate the new z^*.

$$\sum_{i=1}^{4} \frac{\Delta \bar{c}_j}{\Delta \bar{c}_j^{max}} \leq 1 \qquad (16)$$

By replacing the value of $\Delta \bar{c}_j$ in the above equation, it can be observed that the 100% rule is satisfied. Therefore, the current optimal basis remains optimal.

In addition, by calculating $\frac{dz^*}{d\bar{c}_k}$ and replacing those into equation 6 new z^* can be calculated.

$$\Delta z^* = -23.3258$$

$$new\ z^* = oldz^* + \Delta z^* = -4073.013 + (-23.3258) = -4096.3388$$

For Comparison, two steps are taken. First, we include the latest data that has been excluded first time and the portfolio optimization with whole data is solved. We regard this case as the ideal result because we assume that the future can be predicted 100%. Second, which is regarded a historical method the forth stock variation considered as the first case (-0.242) while for the other stocks the average of historical return without the latest data is applied. The purpose of this case is that how the result would be different from the ideal results, if the correlation among parameters is ignored and instead the historical data replaced. Finally, we want to know the derivation of which case is considerable (considering correlation among parameters or using historical data) based on the ideal result. The summary results are presented in table 7.

Table7: The summary of results

	ζ	Non-zero e	x (SON1.SG)	x (MNST)	x (FISV)	x (MAR)	VaR	CVaR	Z*
PCA method	692.33	e(20)=28.53	0	13.87	22.53	13.59	0.16	0.17	-4096.33
Ideal method	468.69	e(21)=127.54	0	14.8018	21.96	13.95	0.11	0.25	-4110.35
Historical method	460.11	105.06	0	15.55135	22.83	11.61	0.111	0.23	-4133.02

Source: Author's research

As can show in table (7) the variation of PCA method and ideal method is 14.02 and the variation of historical method and the ideal method is 22.26, which can be realized that the PCA method has explained the correlation and has resulted in the closer result to the ideal result. The results of three different cases are depicted in Fig.2 and Fig 3.

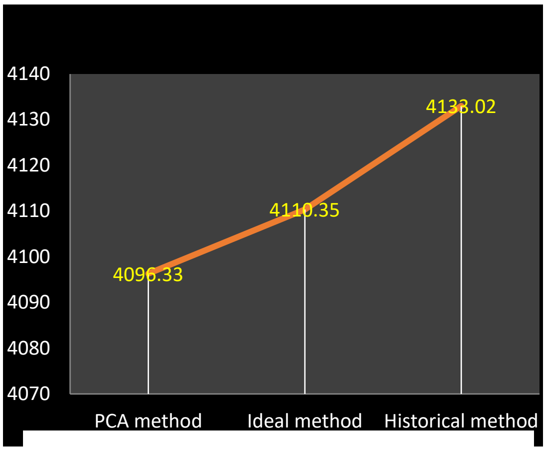

Fig. 2. The amount of Z* of three cases

Source: Author's research

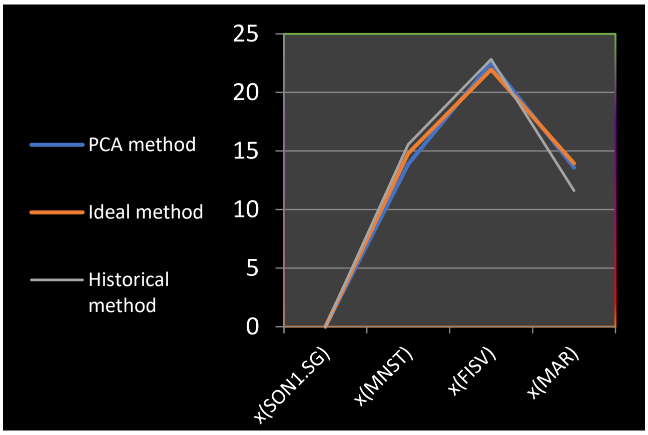

Fig. 3. The amount of x*s in optimal portfolio

The amount of Z* of three cases are compared by

$$\frac{\text{Historical}}{\text{Ideal}} = 1.005; \ \frac{\text{Ideal}}{\text{PCA}} = 1.003; \ \frac{\text{Historical}}{\text{PCA}} = 1.009$$

Clearly, PCA has better performance, In this way. By comparing the amount of x*s in optimal portfolio in four cases SON1.SG, MNST, FIVS, and MAR again the PCA is relatively better.

6. Conclusions:

In this study, we discussed a linear portfolio optimization regarding correlation among the average returns of four stocks. In case of presence of correlation among prices, the variation of one variable contributes to the variation of correlated variables. Current sensitivity analysis had failed to predict the exact changes in other variables. Here, we calculated the changes in correlated variables and compared the result of new method with two other cases.

Historical data set divided into two parts. The first part includes the latest historical returns and the rest of the data put into the second part. For performing portfolio optimization, the CVaR method is employed and PCA method is applied to account for the correlation among OFC. Then, sensitivity analysis of linear portfolio optimization has performed to calculate changes in parameters as a result of change in one parameter without solving problem from the beginning.

In order to evaluate the result, we considered two different cases. First case is called ideal result and obtained by considering the whole data (including latest data). We try to be as much as close to this result. Second case, this obtained through historical data.   The results then indicate that the result of new sensitivity analysis is closer to the ideal result whereas the historical method.

References

1. Arbaiy, N. (2019). A fractal optimization approach for possibility programming problem in fuzzy random environment. *International Journal of Artificial Intelligence and Soft Computing*, 3(4), 330-343.

2. Bazaraa, M. S. (2009). *Linear programming and network flows*. John Wiley & Sons.

3. Heij, C. (2021). Macroeconomic forecasting with matched principal components. *International Journal of Forecasting*, 24(1), 87-100.

4. Hladík, M. (2021). Tolerance analysis in linear systems and linear programming. *Optimization Methods and Software*, 26(3), 381-396.

5. Koltai, T., & Terlaky, T. (2019). The difference between the managerial and mathematical interpretation of sensitivity analysis results in linear programming. *International Journal of Production Economics*, 65(3), 257-274.

6. Lim, A. E. (2020). Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Operations Research Letters*, 39 (3), 163-171.

7. Mina, J., & Xiao, J. Y. (2021). *Return to RiskMetrics: the evolution of a standard*. Risk-Metrics Group.

8. Potra, F. A. (2018). Interior-point methods. *Journal of Computational and Applied Mathematics*, 124(1), 281-302.

9. Shahin, A. (2016). Sensitivity analysis of linear programming in the presence of correlation among right-hand side parameters or objective function coefficients. *Central European Journal of Operations Research*, *24*(3), 563-593.

10. Victor, S. (2019). Comparing the performance of Chinese banks: a principal component approach. *China Economic Review*,18(1), 15-34.