Journal of Engineering
Research and Education
Vol. 9 (2017) 1-10

Structural Learning of Two Layer Boltzmann Machine and Its Application to Power System Investment Planning

M. T. Siti Hajar[1], S. B. Yaakob[*1] and A. Amran[2]

[1]*School of Electrical System Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia.*
[2]*Institute of Engineering Mathematics, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia.*

ABSTRACT

In order to solve a problem efficiently, a structural learning of Boltzmann machine had been proposed and this method enables researcher to solve the problem defined in terms of mixed integer quadratic programming. From this proposed method, an effective selection of results was obtained. In this research, an analysis was performed by using the concepts of the reliability and risks of units evaluated using a variance-covariance matrix. In addition, the effect and expanses of replacement are also measured. Mean-variance analysis is formulated as a mathematical programming with two objectives to minimize the risk and maximize the expected return. Then, a Boltzmann machine was employed to solve the mean-variance analysis efficiently. Findings from this study show that the result of the structural learning of Boltzmann machine method was exemplified. For this reason, the effectiveness of the decision making process can be enhanced.

Keywords: Mean-Variance Analysis, Two Layer Boltzmann Machine, Power System Investment Planning.

INTRODUCTION

A structural learning of Boltzmann machine is proposed as a method to enable researcher to solve problems defined in terms of mixed integer quadratic programming. The Boltzmann machine [3-4] is an interconnected neural network that is proposed by G. E. Hinton. The Boltzmann machine is a model that improves a Hopfield network using probability rule to update the state of a neuron and its energy function. So, the energy function of the Boltzmann machine hardly falls into a local minimum.

Portfolio theory treats a mathematical allocation problem of a given amount of money among several different available investments such as stocks, bonds and stocks. H. Markowitz originally proposed and formulated the mean-variance

[*] Corresponding author. Tel : 604-988 5601, Fax : 604-988 5602. *E-mail address* : *shamshul@unimap.edu.my* (S.B. Yaakob)

approach to the portfolio selection problem[1-2]. The concepts of a Boltzmann machine are applied to solve the portfolio selection problem efficiently. Therefore, if the objective function of a portfolio selection problem is transformed into an energy function of the Boltzmann machine, the portfolio selection problem is enabled to solve as its highly approximate solution. The output value of each unit represents the investing rate to each stock. In the conventional method to solve portfolio selections, the investing rate of each stock is decided to realize the minimum risk under the constraints that the goal rate of an expected return given by a decision maker should be guaranteed.

In this paper, a portfolio selection problem is formulated as a mathematical programming with two objectives which are to minimize risk and to maximize the expected return.

RESEARCH METHODOLOGY

Mean-Variance Analysis

As previously mentioned in the Introduction, mean-variance analysis is widely used in investment theory as proposed by H. Markowitz. In the formulation of mean-variance analysis, H. Markowitz started the discussion with the assumption that almost all decision makers have aversion to risk even if the return might be less. However, it is difficult to identify a utility function because each utility functions have a different utility structure. Therefore, H. Markowitz has formulated a mean-variance analysis as the following quadratic programming problem under the restriction that the expected return rate must be more than certain amount.

[Formulation 1]

$$\text{minimize} \sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j \tag{1}$$

$$\text{subject to} \sum_{i=1}^{n}\mu_i x_i \geq R \tag{2}$$

$$\sum_{i=1}^{n} x_i = 1 \tag{3}$$

$$x_i \geq 0 \quad (i = 1, 2, \cdots, n) \tag{4}$$

where R denotes an acceptable least rate of the expected return, σ_{ij} a covariance between stock i and stock j, μ_i an expected return rate of stock i, and x_i an investing rate to stock i.

In Formulation 1, the optimal solution with the least risk is searched under the constraint that the expected return rate should be more than the value a decision-maker arbitrarily gives. The investing rate to each of the stocks is decided for the solution with the least risk to the given expected return rate. Since the risk is estimated with fixed expected return rate thus, the decision-maker cannot be fully satisfied of this solution. Therefore, the following Formulation 2 is proper and reasonable than the Formulation 1.

[Formulation 2]

$$\text{maximize} \sum_{i=1}^{n} \mu_i x_i \tag{5}$$

$$\text{minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j \tag{6}$$

$$\text{subject to} \sum_{i=1}^{n} x_i = 1 \tag{7}$$

$$x_i \geq 0 \quad (i = 1, 2, \cdots, n) \tag{8}$$

Formulation 2 is a quadratic programming problem with two objective functions of an expected return rate and a degree of risk.

Boltzmann Machine

A Boltzmann machine is an interconnected neural network proposed by G. E. Hinton [3-4]. This model is based on a Hopfield network. The Boltzmann machine is a model that improves a Hopfield network by means of probability rules which are employed to update its state of the neuron and the energy function. If $V_i(t+1)$ is an output value of neuron i in next time $t+1$, $V_i(t+1)$ is 1 according to the probability P which is shown in the following. On the other hand, $V_i(t+1)$ is 0 according to the probability $1-P$.

$$P[V_i(t+1)] = f(\frac{u_i(t)}{T}) \tag{9}$$

where $f(\cdot)$ is a sigmoid function, $u_i(t)$ is the total of input values to neuron i shown in equation (10) and T is a parameter which is called temperature of the network.

$$u_i(t) = \sum_{j=1} w_{ij} V_i(t) + \theta_i \tag{10}$$

where w_{ij} is a weight between neuron i and neuron j, θ_i is a threshold of neuron i. The energy function proposed by J. J. Hopfield is written in the following equation:

$$E = \frac{1}{2} \sum_{ij=1}^{n} w_{ij} V_i V_j - \sum_{i=1}^{n} \theta_i V_i \tag{11}$$

J. J. Hopfield has shown that this energy function simply decreases as the neural network progress [5-6]. There is possibility that this energy function converges to a local minimum. However in the case of the Boltzmann machine, the energy function can increase with minute probability. Therefore, the energy function hardly falls into a local minimum [7-9].

Boltzmann Approach to Mean Variance Analysis

This section will discuss on how to solve a mean-variance analysis using a Boltzmann machine. The mean-variance model shown by Formulation 1 or 2 is transformed into energy function of the Boltzmann machine as follows. First, the objective function in equation (1) or (6) is transformed into energy function as shown in the following equation (12):

$$E = -\frac{1}{2} (\sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j) \tag{12}$$

The next step is a condition that show the total of investment rates of all stocks is 1 (Note that the investment rate of each stock is not less than 0). The condition can be expressed as follows:

$$(\sum_{i=1}^{n} x_i - 1)^2 \tag{13}$$

And equation (13) can be rewritten as follows:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j - 2 \sum_{i=1}^{n} x_i + 1 \tag{14}$$

The equation (12) can be transformed into equation (15), equation (12) is rewritten as follows:

$$E = -\frac{1}{2}(\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j + 2\sum_{i=1}^{n}\sum_{j=1}^{n}x_i x_j) + 2\sum_{i=1}^{n}x_i \quad (15)$$

Finally, the expected return was considered by using the equation of the expected return as in equation (2) or (5). Therefore, (2) or (5) were transformed into equation (16):

$$E = -\frac{1}{2}(\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j + 2\sum_{i=1}^{n}\sum_{j=1}^{n}x_i x_j) + 2\sum_{i=1}^{n}x_i + K\sum_{i=1}^{n}\mu_i x_i \quad (16)$$

where K is a real number which is not less than 0.

If value K is set to a larger number, the expected return is evaluated much more than the risk. If $K = 1.0$, then the Boltzmann machine converges into a problem of minimizing its risk. When the energy function of the Boltzmann machine described in this section converges into the global minimum, the investing rate to stocks by the output value of each unit can be obtained.

The algorithm of the Boltzmann machine is executed according to the Algorithm 1.

[Algorithm 1]

Step 1: Give the initial value of all units optionally.

Step 2: Choose a certain unit (i) out of all units at random.

Step 3: Compute a total of the input $u_i(t)$ into the chosen unit $i(1 \leq iN)$.

Step 4: Add a sufficiently small value to the output value $V_i(t+1)$ of unit i according to the probability P shown in equation (9). And subtract a sufficiently small value from the output value with probability $1-P$. But the output value isn't varied in the case $u_i(t) = 0$.

Step 5: The output value of units j except i aren't varied.

Step 6: After iterating from Steps 2 to 5, compute the probability of each unit for all units.

Two Layer Boltzmann Machine

The two layer Boltzmann machine is a neural network model, which is proposed by J. Watada *et al.* [10]. This model deletes the units of lower layer, which are not selected in the upper layer in its execution. Then, the lower layer is restructured using the selected units. Because of this feature, a two layer Boltzmann machine converges more efficiently than a conventional Boltzmann machine. This is an efficient method for solving a portfolio selection problem by transforming its objective function into the energy function since the Hopfield and Boltzmann machines converge at the minimum point of the energy function.

The two layer Boltzmann machine mentioned above converted the objective function into the energy functions of the two components that are upper layer (Hopfield network) and the lower layer (Boltzmann machine) as described below:

Upper Layer

$$E_u = -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}s_i s_j + k_u\sum_{i=1}^{n}\mu_i s_i \tag{17}$$

Lower Layer

$$E_l = -\frac{1}{2}(\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}x_i x_j + 2\sum_{i=1}^{n}\sum_{j=1}^{n}x_i x_j) + 2\sum_{i=1}^{n}x_i + K_l\sum_{i=1}^{n}\mu_i x_i \tag{18}$$

where K_u, K_l are weights of the expected return rate for each layer and S_i is the output value of the i^{th} unit of the upper layer [11-13].

The algorithm of two layers Boltzmann machine is described as follows.

[Algorithm 2]

Step 1: Set each parameter to its initial value.
Step 2: Input the values of K_u and K_l.
Step 3: Execute the upper layer.
Step 4: If the output value of a unit in the upper layer is 1, add some amount of value to the corresponding unit in the lower layer. Execute the lower layer.
Step 5: After executing the lower layer the constant number of times, decrease the temperature.
Step 6: If the output value is sufficiently large, add a certain amount of value to the corresponding unit in the upper layer.
Step 7: Iterate from Step 3 to Step 6 until the temperature reaches the restructuring temperature.
Step 8: Restructure the lower layer using the selected units of the upper layer.
Step 9: Execute the lower layer until reaching at the termination.

RESULTS AND DISCUSSION

Data

This explanatory example was employed to illustrate the power system investment planning problem. The maintenance cost consists of fifteen substations. In this analysis, the maintenance cost rate for five years are employed to analyze the expense investment over fifteen substations.

Table 1: Result of simulation in investment rate for each substation

Substation	K=0.3	K=0.5	K=0.7	K=1.0
Substation A	-	-	-	-
Substation B	-	-	-	0.0019
Substation C	0.5165	0.4999	0.4669	0.4581
Substation D	-	-	-	-
Substation E	-	-	-	-
Substation F	-	-	-	-
Substation G	0.2174	0.2371	0.2593	0.2671
Substation H	-	-	-	-
Substation J	-	-	0.0092	0.0101
Substation K	-	-	-	-
Substation L	0.2661	0.2561	0.2401	0.2278
Substation M	-	-	-	-
Substation N	-	0.0069	0.0123	0.0191
Substation P	-	-	-	-
Substation Q	-	-	0.0122	0.0159

Table 2: Expected invest rate and risk

K	Investment Rate	Risk
0.3	4.8211	0.0499
0.5	4.8999	0.0581
0.7	4.9045	0.0601
1.0	4.9143	0.0682

Result

From Table 1, in case where K = 0.3, Substation C should be invested by 51.7%, Substation G by 21.7% and Substation L by 26.6%. Other substations which are not included in the list of units after restructuration will not get any investment. In case where K = 0.5, 0.7, and 1.0, it can be concluded that the number of selected substations in restructured list will increase if K is increase.

The expected investment rate and risk were calculated as shows in Table 2 indicating four different levels of risk aversion, values of K, and reflect decision maker different preference. When K is set to the large value, the solution was obtained with high investment rate and high risk.

CONCLUSION

A new proposal was presented to define an efficient method for investment of the maintenance cost based on mean-variance analysis. The proposed method is applied in concern with the portfolio method for investment. Using two layers Boltzmann machine for unit restructure, the basis for estimating the investment expense of maintenance was obtained. In order to evaluate the proposed method, a case study is performed which generated quantitative predictions of the invest expenses to the timeout failure. Calculation of the cost rate associated with maintenance cost is becoming an important factor as the computation of the reliability measures themselves. The mean-variance analysis which employs portfolio method and two layers Boltzmann machine approach can effectively deal with these types of power system investment planning problems. The effectiveness of the two layers Boltzmann machine approach applied to mean–variance analysis that is based on the uncertainties in the alternatives. The selection of substations for investment based on the maintenance cost has been discussed. Based on the allocated budget of the maintenance cost, the change in investment rate and types of results obtained were discussed. The results also demonstrate that this proposal is effective for the decision making process and for this reason, the effectiveness of the decision making process can be enhanced.

REFERENCES

[1] Markowitz, H. (1987). *Mean-variance analysis in portfolio choice and capital markets* (1st ed.). Oxford, OX, UK: B. Blackwell.

[2] Ellis, C., & Wilson, P. J. (2005). Can a Neural Network Property Portfolio Selection Process Outperform the Property Market. *Journal of Real Estate Portfolio Management*, 11(2), 105-121

[3] Bass, B. & Nixon, T. (2013). Boltzmann Learning. Reference Module in Earth Systems and Environmental Sciences.

[4] Hinton, G. (2007). Boltzmann machine. Scholarpedia, 2(5), 1668.

[5] Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. Proceedings of the National Academy of Sciences, 79(8), 2554-2558.

[6] A. Ashwood (June, 2013). Portfolio Selection Using Artificial Intelligence. Unpublished PhD thesis, Queensland University of Technology.

[7] Barra, A., Bernacchia, A., Santucci, E., & Contucci, P. (2012). On the equivalence of Hopfield networks and Boltzmann Machines. Neural Networks, 34, 1-9.

[8] Agliari, E., Barra, A., De Antoni, A., & Galluzzi, A. (2013). Parallel retrieval of correlated patterns: From Hopfield networks to Boltzmann machines. Neural Networks, *38*, 52-63.

[9] Xia, Y. & Wang, J. (2016). A Bi-Projection Neural Network for Solving Constrained Quadratic Optimization Problems. IEEE Transactions on Neural Networks And Learning Systems, *27*(2), 214-224.

[10] Watada, J and Oda, K. (2000). Formulation of a Two-Layered Boltzmann Machine for Portfolio Selection. *International Journal of Fuzzy Systems,* 2(1), 39-44.

[11] Yaakob, S. B., Watada, J., & Fulcher, J. (2011). Structural learning of the Boltzmann machine and its application to life cycle management. Neurocomputing, *74*(12-13), 2193-2200.

[12] Yaakob, S. B., Watada, J., Takahashi, T., & Okamoto, T. (2012). Reliability enhancement of power systems through a mean–variance approach. Neural Computing and Applications, 21(6), 1363-1373.

[13] Yaakob, S.B., & Watada, J. (2011). Solving Bilevel Quadratic Programming Problems and Its Application. *Knowledge-Based and Intelligent Information and Engineering Systems*. 6883(series Lecture Notes in Computer Science), 187-196.