Research Article

A Class of Stochastic Programming Model in Investment Portfolio Based on Covering Rough Set

Lei Zhou and Dongli Zhang

School of Economics and Management, Hebei University of Science and Technology, Shijiazhuang 050018, China

Correspondence should be addressed to Lei Zhou; zhoulei19@126.com

Received 23 November 2021; Revised 17 January 2022; Accepted 21 January 2022; Published 11 February 2022

Academic Editor: Anna M. Gil-Lafuente

Copyright © 2022 Lei Zhou and Dongli Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to study the investment portfolio problem, this paper proposes a class of stochastic programming model with rough feasible region, where randomness and roughness coexist. Based on the covering rough set, the concept of the discrete degree covering is defined to divide the rough feasible region. Furthermore, the discrete degree covering stochastic rough programming model (DDC-SRP) is constructed depending on a synthesis effect function that considers discrete degree and the expectation and variance for random objective function. Properties of the DDC-SRP model are discussed. In addition, the convexity of the DDC-SRP model is obtained in some certain conditions. Considering the random rough simulation, a genetic algorithm is introduced. Finally, a numerical example is given to show the validity of the DDC-SRP model.

1. Introduction

Investment portfolio problem is very important in resource allocation with uncertainty. For dealing with such decision-making problems with uncertainty, many scholars have introduced some models including random or fuzzy variables to formulate the uncertainty. Mean-variance portfolio selection has been a central issue in finance since Markowitz’s pioneering work [1] on a single-period investment model. Fu et al. [2] derived explicit closed-form solutions, for the efficient frontier, optimal investment strategy and for the dynamic mean-variance portfolio selection problem under the constraint of a higher borrowing rate. Yu et al. [3] considered the effects of some frequently used utility functions in portfolio selection by comparing the optimal investment outcomes corresponding to these utility functions. Zhang et al. [4] proposed a new portfolio selection model with the maximum utility based on the interval-valued possibilistic mean and possibilistic variance, which is a two-parameter quadratic programming problem. Zhang [5] discussed the portfolio selection problem for bounded assets based on upper and lower possibilistic means and variances. The mean-standard deviation model for portfolio selection can be transformed to a linear programming under possibility distributions, so this methodology can be used to solve large-scale portfolio selection problems. Guo et al. [6] considered a fuzzy multiperiod portfolio selection problem with V-shaped transaction cost. Within the framework of credibility theory, a mean-variance model is formulated with the objective of maximizing the terminal return under the total risk constraint over the whole investment. Giner [7] proposed a new and simple variance decomposition based on splitting the domain space, and through the law of total variance and the first and second moments of each truncated distribution, the requested decomposition formulas are deduced. In [8], the improved artificial bee colony (ABC) algorithm with an external population is applied to the optimization of the blockchain investment portfolio.

These above studies considered the investment portfolio problem with randomness and fuzziness. However in real life, when making portfolio selection, we choose one stock as the key investment object (core stock) and others as noncore stocks to form a portfolio. Since each combination is a partition of an alternative stock set, the alternative stock set can be regarded as a rough set. The traditional classification by regression analysis can only deal with randomness to a
certain extent, but cannot deal with the roughness caused by portfolio classification of decision-makers' subjective consciousness. So here we apply rough set theory to deal with portfolio. When we select a stock as an investment objective, we must consider many related factors, such as the national policy, the enterprise benefit, the prospects for the development of the enterprise, and so on. Because the investment risk is bigger when investors select only one stock, they often select a stock portfolio, which means that they will select a class of stocks. In order to disperse risk in investment, investors tend to choose the obvious difference enterprises or the enterprises whose futures are not related and use their stocks as an investment portfolio. For example, we choose a stock that is more affected by the market and several stocks that are less affected by the market to form a portfolio. So if one cost of stock falls, it will not affect the other stocks. Here, a core stock and other stocks associated with it are used to form a portfolio. Taking each stock as the core stock can form a portfolio. These portfolios just cover the set composed of all optional stocks, that is, all portfolios are equivalent to the division of this optional stock set. Roughness is the uncertainty caused by the division of sets, so rough set theory is used to analyze stock investment problems. When selected a stock as the investment core, the investment portfolio with the core has roughness due to the difference result of classification. Therefore, the feasible region of the considered decision problem is rough set. In this article, a class of stochastic programming model with rough feasible region is put forward in investment portfolio. X is the set of investment stocks that can be selected. In accordance with the alternatives, expected return attribute that is numerical attribute, stocks can be classified. Stocks can also be classified in accordance with the different enterprise futures that is categorical attribute. On the investment object classification, we must consider both the numerical attribute and the categorical attribute. The stocks that are in the same class will be with similar relation N. \( f(x) \) is the comprehensive return value with the core investment object x in each portfolio. \( (x)_N \) is a set in which these elements have similar relation with x. These are two investment objects in the investment portfolio. Investors' goal is to choose the best investment portfolio in order to get the maximum expected profit, and this similar relation has the roughness based on covering.

Because this roughness based on covering, rough set, and covering rough set theory must be introduced. Rough set theory [9], a tool for data analyzing, was first proposed by Pawlak in 1982. As a new mathematical theory to model incomplete knowledge, it can deal with problems with inexact data or imprecise information for complicated systems effectively. Now many researchers are studying on rough set with randomness. Zhu [10] studied a type of generalized rough sets based on covering and the relationship between this type of covering-based rough sets and the generalized rough sets based on binary relation. In [11], several uncertainty measures of neighborhood granules are proposed, which are neighborhood accuracy, information quantity, neighborhood entropy and information granularity in the neighborhood systems. Zhou et al. [12] defined a new neighborhood rough set relation with adapted neighbors named the gap relation and propose a new online streaming feature selection method based on this relation, named OFS-A3M. With the "maximal-dependency, maximal-relevance and maximal-significance" evaluation criteria, OFS-A3M can select features with high correlation, high dependency, and low redundancy. Many researchers have used probabilistic and fuzzy set methods to study rough sets and try to formulate axiomatic systems for classical rough sets from different viewpoints [13–18].

Based on rough set for the knowledge partition, rough programming is an equivalent classification method to determine the approximate domain of a given problem. Youness [19] proposed the concept of a rough programming and defined the surely optimal solution and possibly optimal solution. By equivalent classification method of rough set, the global rough optimal solution and local rough optimal solution are defined, in order to solve the problem of rough programming. Liu [20] introduced a concept of rough variable and further gave the concept of random rough variable, rough expected value model, and rough chance-constrained model. They provide a useful basis to solve the uncertainty of randomness and roughness. Xu et al. [21, 22] put forward a class of multiobjective linear programming models with random rough coefficients and gave some characters and solution methods. In [23], the concept of “rough interval” is introduced in the modeling framework to represent dual-uncertain parameters. Osman et al. [24] proposed that rough programming will be divided into three types in accordance to the roughness of the feasible domain and the objective function. They discussed the characteristics of feasible domain, optimal value, and optimal solution of the rough programming. Tao et al. [25] proposed a class of rough multiple objective programming and used it to solid transportation problem. Many researchers have studied rough programming from different viewpoints [26, 27].

The above literature puts forward some basic concepts for rough programming, but the practical application is narrow, and most of them should be discussed theoretically. This study studies the roughness of talent in the process of investment decision making, constructs an effect based rough programming model, and discusses its solution process.

Based on the above analysis, this paper made some aspects of the work: (1) by the investment portfolio problem, a class of rough programming model with discrete degree covering is proposed; (2) based on the random information and discrete degree covering roughness information compound quantitative, a class of rough programming model based on the synthesis effect and discrete degree covering is established, which is named discrete degree covering stochastic rough (DDC-SRP) model; (3) the convexity of the DDC-SRP model is analyzed; and (4) the DDC-SRP model is used into investment portfolio problem and the features of the solution are analyzed.
2. Preliminaries

In this section, we will introduce covering rough set model based on neighborhood, and the concepts of covering and neighbor can be showed in [10, 28–34]. Here, \( U \) is a universe.

**Definition 1.** Given arbitrary \( x_i \in U \) and \( B \subseteq C \), the neighborhood \( \delta_B(x_i) \) in feature space \( B \) is defined as

\[
\delta_B(x_i) = \{ x_j | x_j \in U, \Delta_B(x_i, x_j) \leq \delta \},
\]

where \( \Delta_B \) is a distance function. For any \( x_1, x_2, x_3 \in U \), it usually satisfies: (1) \( \Delta_B(x_1, x_2) \geq 0 \), \( \Delta_B(x_1, x_2) = 0 \) if and only if \( x_1 = x_2 \); (2) \( \Delta_B(x_1, x_2) = \Delta_B(x_2, x_1) \); (3) \( \Delta_B(x_1, x_3) \leq \Delta_B(x_1, x_2) + \Delta_B(x_2, x_3) \).

**Definition 2.** Let \( B_1 \subseteq A \) and \( B_2 \subseteq A \) be numerical attributes and categorical attributes, respectively. The neighborhood granule of sample \( x \) induced by \( B_1 \), \( B_2 \) and \( B_1 \cup B_2 \) is defined as (1) \( \delta_{B_1}(x) = \{ x_j | \Delta_{B_1}(x, x_j) \leq \delta, x_j \in U \} \); (2) \( \delta_{B_2}(x) = \{ x_j | \Delta_{B_2}(x, x_j) = 0, x_j \in U \} \); (3) \( \delta_{B_1 \cup B_2}(x) = \{ x_j | \Delta_{B_1 \cup B_2}(x, x_j) \leq \delta \Delta_{B_1}(x, x_j) \leq 0, x_j \in U \} \), where \( \Delta \) means “and” operator.

**Definition 3.** Given a set of objects \( U \) and a neighborhood relation \( N \) over \( U \), we call \( \langle U, N \rangle \) a neighborhood approximation space. For any \( x \in U \), any two subsets of objects, called \( N \)-lower approximation and \( N \)-upper approximation of \( x \) in \( \langle U, N \rangle \), are defined as

\[
\begin{align*}
N(x) &= \{ x_j | \delta_B(x_j) \subseteq x, x_j \in U \}, \\
\overline{N}(x) &= \{ x_j | \delta_B(x_j) \cap X \neq \phi, x_j \in U \}.
\end{align*}
\]

Obviously, we can see that the elements of \( N(X) \) are in \( X \) and the elements of \( U \setminus \overline{N}(X) \) are not in \( X \). But the elements of \( \overline{N}(X) \setminus N(X) \) are neither in \( X \) nor in \( U \setminus X \). So in approximation space \( \langle U, N \rangle \), \( X \) can be described by \( (N(X), \overline{N}(X)) \). We shall employ the following denotations: \( pos_N(X) = N(X) \) is \( N \)-positive region of \( X \); \( neg_N(X) = U \setminus \overline{N}(X) \) is \( N \)-negative region of \( X \); \( bn_N(X) = \overline{N}(X) \setminus N(X) \) is \( N \)-borderline region of \( X \); and \( a_N = |\overline{N}(X)|/|N(X)| \) is \( N \)-approximation precision of \( X \). Here, \( X \neq \phi \), \( |A| \) denotes the number of the elements in \( A \). In order to describe clearly, we note that \( x \in N(X) \) if and only if \( x \in N(X) \) and \( x \notin \overline{N}(X) \). In other words, \( x \in N(X) \) if and only if \( x \in X \) and \( x \notin \overline{N}(X) \).

3. Discrete Degree Covering Stochastic Rough (DDC-SRP) Model

3.1. Stochastic Rough Programming Model Based on Covering Approximation Relation

In optimization problems, the core is to make a decision that can make objective function optimal in a given feasible region. The model is as follows:

\[
\begin{align*}
\text{max } f(x), \\
\text{s.t. } x \in M,
\end{align*}
\]

Here, \( M \) is a set on \( U \); \( f(x) \) is a numerical or non-numerical function on \( U \). This is said \( M \) is the feasible region; \( f(x) \) is the objective function.

In according to the difference of feasible region and objective function, model (1) can be divided into crisp programming and uncertain programming. In this paper, we will discuss a class of uncertain programming whose objective function contains random variables and feasible region is a rough set based on similar relation. This model is called rough programming model based on similar relation.

There are too many decision problems with the above characteristics such as equipment renewal, resource management, complex system optimization, and so on. In different problems, the effect of \( (x)_u \) is different, but the decision processes involve the elements that have relation \( N \) with \( x \) in \( M \). That is the \( N \)-upper approximation of \( X \) in \( \langle U, N \rangle \). Therefore, the model of rough programming model based on similar relation is as follows:

\[
\begin{align*}
\text{max } E(f(x, \xi)), \\
\text{s.t. } x \in NM.
\end{align*}
\]

Here, \( X \) is a set on \( U \); \( f(x, \xi) \) is a numerical or non-numerical random function on \( U \). The effect on \( x \) can be described by the expectation and variance of \( f(x, \xi) \), which is called direct effect; \( N \) is a similar relation on \( U \). \( x \in N \subseteq M \). If we only consider the expectation of random variable, we can get the expectation value model:

\[
\begin{align*}
\text{max } E(f(x, \xi)), \\
\text{s.t. } x \in N \subseteq M.
\end{align*}
\]

The effect on \( x \) can be also described by \( (x)_N \) which is called indirect effect. Therefore, when we make a decision by rough programming, direct effect and indirect effect must be considered together. Indirect effect can be represented by a function \( G: U/N \rightarrow [0, \infty) \). Then direct effect and indirect effect can be considered as a function \( S(u, v, w): [0, \infty) \times [0, \infty) \times [0, \infty) \rightarrow [0, \infty) \), which named synthesis effect function. Synthesis effect function \( S(u, v, w) \) should satisfy the following conditions.

**Condition 1.** For any given \( v, w \in [0, \infty) \), \( S(u, v, w) \) is monotone nondecreasing in each \( u \).

**Condition 2.** For any given \( u, v \in [0, \infty) \), \( S(u, v, w) \) is monotone nondecreasing in each \( v \).

**Condition 3.** For any given \( u, v \in [0, \infty) \), \( S(u, v, w) \) is monotone nondecreasing in each \( w \).

**Condition 4.** \( S(0, 0, 0) \) is monotone increasing in each \( u \).

Let \( u = E(f(x, \xi)), v = D(f(x, \xi)), w = G((x)_N) \), model (2) can be converted into the following model (4):

\[
\begin{align*}
\text{max } S(E(f(x, \xi)), D(f(x, \xi)), G((x)_N)), \\
\text{s.t. } (x)_N \subseteq M.
\end{align*}
\]

Model (4) is named rough synthesis effect model. The following two kinds of functions are synthesis effect function.

\[
S_1(u, v, w) = u(1 + k l w)^{1-(1 + l w)}, l \leq 0, 0 \leq a, b < \infty.
\]
We see that rough programming can be seen as composite of a stochastic programming and a similar relation on \( U \). With different synthesis effect operators, the rough programming model can be converted into different crisp programming models.

3.2. Description of DDC-SRP Model. In Section 3.1, in accordance to several typical decision-making problems, general form of the rough programming model is given and rough synthesis effect model is proposed. In different problems, \((x)_N\) represents different means. So how to build the indirect effect metric mode is a key part in solving rough programming problem. In this section, we use dispersion as indirect effect and propose the DDC-SRP model.

Definition 4. If \( d(x_j) \) means the value of attribute \( x_j \), then

\[
C_D((x)_N) = \sum_{y_i, y_j \in (x)_N} |d(y_i) - d(y_j)|
\]

Means discrete degree of \((x)_N\).

Using \( C_D(x) \) as indirect effect of \((x)_N\), model (4) can be converted into the following model (5), which named the DDC-SRP model:

\[
\begin{align*}
\text{max} & \ E(f((x)_N, \eta)), D(f((x)_N, \eta)), C_D((x)_N), \\
\text{s.t.} & \ (x)_N \subset M.
\end{align*}
\]

Here, \( \eta \) is a random vector.

Specially, if \( S(u, v, w) = \{w \geq \alpha, 0 < w < \alpha\} \), model (5) can be converted into model (6)

\[
\begin{align*}
\text{max} & \ E(f((x)_N, \eta)), D(f((x)_N, \eta)), \\
\text{s.t.} & \ (x)_N \subset M, \text{and } C_D((x)_N) \geq \beta.
\end{align*}
\]

Here, \( \beta \in [0, +\infty) \) means the lowest level of discrete degree of \((x)_N\).

Let \( S_2(u, v, w) = \{ue^{-av}(1 + bw), 0 \leq a, b < \infty\} \), model (5) can be converted into model (7)

\[
\begin{align*}
\text{max} & \ E(f((x)_N, \eta))e^{-aD(f((x)_N, \eta))}(1 + bC_D((x)_N)), \\
\text{s.t.} & \ (x)_N \subset M.
\end{align*}
\]

3.3. Features of the DDC-SRP Model

Theorem 1. If we select a class of special synthesis effect functions, the DDC-SRP model (9) can be converted into the expectation model (7) of rough programming.

Proof. We use a class of synthesis effect functions \( S_2(u, v, w) = \{ue^{-av}(1 + bw), 0 \leq a, b < \infty\} \). If we let \( a = b = 0 \), the DDC-SRP model (9) can be converted into the expectation model (7) of rough programming.

\[
\text{Theorem 2.} \text{ If random variable function } g(x, \xi) \text{ is joint convex with respect to } (x, \xi), \text{ and probability measure } Pr \text{ is pseudo-concave, then for any given confidence level } \alpha, 0 \leq \alpha \leq 1, \text{ chance constrained set } X = \{x | Pr[g(x, \xi) \leq \alpha] \geq \alpha\} \text{ is a convex set.}
\]

Proof. Let \( x_1, x_2 \in X \), then \( Pr[g(x_1, \xi) \leq \alpha] \geq \alpha \) and \( Pr[g(x_2, \xi) \leq \alpha] \geq \alpha \). So

\[
Pr[\lambda g(x_1, \xi) + (1 - \lambda)g(x_2, \xi) \leq 0] \geq \min\{|Pr[g(x_1, \xi) \leq 0], Pr[g(x_2, \xi) \leq 0]| \geq \alpha
\]
Therefore, \( \Pr \left[ g(\lambda x_1 + (1-\lambda)x_2, \xi) \leq 0 \right] \geq \alpha \). So \( \lambda x_1 + (1-\lambda)x_2 \in X \).

It shows that chance constrained set \( X = \{ x | \Pr [ g(x, \xi) \leq 0 ] \geq \alpha \} \) is a convex set.

**Definition 6.** Let \( X \subseteq \mathbb{R}^n \) be a convex set, \( \xi \) be a given random variable on probability space \((\Omega, F, \mathbb{P})\), and \( f(x, \xi) \) be a random variable function on \( X \).

(1) If \( D(f(\lambda x_1 + (1-\lambda)x_2), \xi) \leq \lambda D(f(x_1, \xi) + (1-\lambda)f(x_2, \xi)) \) for any given \( x_1, x_2 \in X \) and \( \lambda \in [0, 1] \), then \( f(x, \xi) \) is stochastic convex with respect to \( x \) on \( X \);

(2) If \( D(f(\lambda x_1 + (1-\lambda)x_2), \xi) \geq \lambda D(f(x_1, \xi) + (1-\lambda)f(x_2, \xi)) \) for any given \( x_1, x_2 \in X \) and \( \lambda \in [0, 1] \), then \( f(x, \xi) \) is stochastic concave with respect to \( x \) on \( X \).

**Definition 7.** Let \( X \subseteq \mathbb{R}^n \) be a convex set. If \( C_D(\{ \lambda x_1 + (1-\lambda)x_2 \} \cap \Omega) \leq \lambda C_D(\{ x_1 \} \cap \Omega) + (1-\lambda)C_D(\{ x_2 \} \cap \Omega) \) for any given \( x_1, x_2 \in X \) and \( \lambda \in [0, 1] \), then \( C_D(\{ x \} \cap \Omega) \) is convex with respect to \( x \) on \( X \).

**Theorem 3.** Let \( X \subseteq \mathbb{R}^n \) be a convex set, \( \xi \) be a given random variable on probability space \((\Omega, F, \mathbb{P})\). If

(1) \( S(u, v, w) \) is joint convex with respect to \((u, v, w)\) on \( \mathbb{R}^3 \);

(2) for any given \( \xi \), \( f(x, \xi) \) is convex with respect to \( x \) on \( X \);

(3) \( f(x, \xi) \) is stochastic concave with respect to \( x \) on \( X \), and

(4) \( C_D(\{ x \} \cap \Omega) \) is convex with respect to \( x \).

Then the DDC-SRP model (9) is a convex programming.

**Proof.** From Condition 2, we can see that for any given \( \xi \),

\[
E(f(\lambda x_1 + (1-\lambda)x_2, \xi)) \leq \lambda E(f(x_1, \xi)) + (1-\lambda)E(f(x_2, \xi)).
\]

It shows that \( E(f(x, \xi)) \) is convex with respect to \( x \) on \( X \). Because \( S(u, v, w) \) is monotone nondecreasing in each \( u \), we can get

\[
S(E(f(\lambda x_1 + (1-\lambda)x_2, \xi)), v, w) \leq S(\lambda E(f(x_1, \xi)) + (1-\lambda)E(f(x_2, \xi)), v, w). \tag{14}
\]

From Condition 3, we know that

\[
D(f(\lambda x_1 + (1-\lambda)x_2, \xi)) \geq \lambda D(f(x_1, \xi)) + (1-\lambda)D(f(x_2, \xi)). \tag{15}
\]

So

\[
S(u, D(f(\lambda x_1 + (1-\lambda)x_2, \xi)), w) \leq S(u, \lambda D(f(x_1, \xi)) + (1-\lambda)D(f(x_2, \xi)), w). \tag{16}
\]

\( S(u, v, w) \) is monotone nondecreasing in each \( w \), so

\[
S(u, v, w(\lambda x_1 + (1-\lambda)x_2)) \leq S(u, v, \lambda w(x_1) + (1-\lambda)w(x_2)). \tag{17}
\]

From Condition 1, we can see that \( S(u, v, w) \) is joint convex with respect to \((u, v)\) on \( D \subseteq \mathbb{R}^3 \). So

\[
S(E(f(\lambda x_1 + (1-\lambda)x_2, \xi)), D(f(\lambda x_1 + (1-\lambda)x_2, \xi)), C_D(\{ \lambda x_1 + (1-\lambda)x_2 \} \cap \Omega)) \leq S(\lambda E(f(x_1, \xi)) + (1-\lambda)E(f(x_2, \xi)), D(f(x_1, \xi)) + (1-\lambda)D(f(x_2, \xi)), C_D(\{ x_1 \} \cap \Omega) + (1-\lambda)C_D(\{ x_2 \} \cap \Omega)) \tag{18}
\]

\[
\leq S(\lambda E(f(x_1, \xi)) + (1-\lambda)E(f(x_2, \xi)), D(f(x_1, \xi)) + (1-\lambda)D(f(x_2, \xi)), \lambda C_D(\{ x_1 \} \cap \Omega) + (1-\lambda)C_D(\{ x_2 \} \cap \Omega)) \leq S(\lambda E(f(x_1, \xi)) D(f(x_1, \xi)), C_D(\{ x_1 \} \cap \Omega) + (1-\lambda)S(E(f(x_2, \xi)), D(f(x_2, \xi)), C_D(\{ x_2 \} \cap \Omega)).
\]
Therefore, the objective function of model (5) is a convex function. So model (5) is a convex programming.

If the DDC-SRP model is a convex programming, we can use 0.618 method, Newton method, restricted function method, etc. to solve it. But when we deal with some complex problems, the DDC-SRP model is a nonconvex programming, and the above methods cannot be used. Genetic algorithm is an effective method to solve such programming problems. In part 4, we will give the solution procedure by genetic algorithm, and an example about investment portfolio problem will be given in part 4. \[ \square \]

4. Solution Procedure

In what follows, we will give the solution procedure as chart 1 to solve stochastic rough programming problem. Here we will establish the DDC-SRP model by synthesis strategy of objective and discrete degree of \((x)_N\).

Step 1 Input the initial programming model. Based on the investment portfolio problems, a class of stochastic programming model with rough feasible region is established.

Step 2 Select the proper a discrete degree \(C_D(x)\) and a synthesizing effect function \(S(u, v, w)\). The discrete degree that we select must follow the Definition 4 and the synthesizing effect function that we select must meet Conditions 1–4 in Section 3. We can also select a class of synthesizing effect functions such as \(S_j(u, v, w) = ae^{-\alpha u} (1 + bw), 0 \leq a, b < \infty\). With different coefficients \(k, l, a, b\), we can get synthesizing effect functions.

Step 3 Convert the rough programming model to the DDC-SRP model. Through proper synthesizing effect functions, we can convert the stochastic programming model with rough feasible region to the DDC-SRP model. The DDC-SRP model is a programming model without uncertainty.

Step 4 Solve the DDC-SRP model by genetic algorithm. In following, we introduce the genetic algorithm based on real coding and random rough simulation.

Step 4.1 Input the parameters population size \(N\), crossover probability \(P_c\), mutation probability \(P_m\).

Step 4.2 Initialize \(N\)-size chromosomes and check feasibility of chromosomes.

Step 4.3 Update the chromosomes by crossover and mutation operations and check the feasibility of offspring.

Step 4.4 Select the chromosomes by spinning the roulette wheel.

Step 4.5 Make the crossover operation and the mutation operation.

Step 4.6 Repeat the Step 2 to the Step 5, until completes the cycle index.

Step 4.7 Get the best chromosome as the optimal solution and obtain the optimum value.

Step 5 Output the solution.

A flowchart for the solution procedure is provided, as shown in Figure 1.

5. Number Example

A investor want to choose a portfolio among 20 given stocks. The total fund is 100, 000 dollars. The performance and relevance of the 20 stocks in the following Table 1. \(\xi_j\) that is the profit of jth stock \((j)\) is a random variable. Here, \(\xi_j\), \(j = 1, 7, 9, 16\) obey normal distribution whose expectation is 18 and variance is 4. \(\xi_j\), \(j = 5, 8, 14, 20\) obey normal distribution whose expectation is 12 and variance is 4. \(\xi_j\), \(j = 2, 3, 4, 6, 10, 11, 12, 13, 15, 17, 18, 19\) obey normal distribution whose expectation is 4 and variance is 2.

Remark 1. The data in Table 1 come from the data of China’s stock market in the past 10 years, which are obtained through the database of China stock network. Here, in the past 10 years, it is “good” if the stock market value fluctuation increases, and it is “poor” if the stock market value fluctuation decreases. \(N_2\) represents the relevance of the relationship between the rise and fall of the stock and the rise of the market, which also comes from the data of China’s stock market in the past 10 years. The worse the relevance, the greater the \(N_2\) value of the two stocks. Here, \(N_2\) is calculated according to the change relevance statistics of the rise and fall of the two stocks.

\[ U = \{(1), (2), \ldots, (20)\} \] is a set of all stocks that can be chosen. From Table 1, we can see that performance \((N_1)\) is a categorical attribute and relevance \((N_2)\) is a numerical attribute. Using categorical attribute \(N_1\), \(U\) can be divided into two parts. Then we can get the following rough programming model.

\[
\begin{align*}
\max z &= f \left( (x)_N, \eta \right), \\
\text{s.t.} & \quad (x)_N \in U.
\end{align*}
\]

Here, \(N = N_1 \cup N_2\). \((x)_N\) is a covering under \(N\). \(\eta\) is a random vector represent the profit of stock in \((x)_N\).

Each core investment object \(x_j, j = 1, 5, 7, 8, 9, 14, 16, 20\) must come from the class of “good” that is \{1, 5, 7, 8, 9, 14, 16, 20\}. The discrete degree of the portfolio \((x)_N\) must be as big as possible in order to disperse risk. Here let the relevance distances between every secondary investment object of \((x)_N\) and core investment
object are bigger than 2 \((d_i = 2)\). Then we can get the following coverings.

\[
\begin{align*}
\delta((1)) &= \{10, 11, 14, 15, 16\}, \\
\delta((5)) &= \phi, \\
\delta((7)) &= \{8, 9, 10, 11, 12, 14, 15, 16, 19, 20\}, \\
\delta((8)) &= \{2, 3, 4, 6, 7, 17\}, \\
\delta((9)) &= \{6, 7\}, \\
\delta((14)) &= \{1, 2, 3, 4, 6, 7, 17\}, \\
\delta((16)) &= \{1, 2, 3, 4, 6, 7, 17\}, \\
\delta((20)) &= \{4, 6, 7\}. \\
\end{align*}
\]

Furthermore, in the same investment portfolio, the relevance between every two secondary investment object can be closer than the relevance between every secondary investment object and core investment object. In a same \(\delta(x_i)\), the relevance distances between every two secondary investment object can be smaller than the relevance distances between every secondary investment object of \((x_i)_{N}\) and core investment object. So the relevance distances

<table>
<thead>
<tr>
<th>Stock code</th>
<th>Performance ((N_1))</th>
<th>Relevance ((N_2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Good</td>
<td>2</td>
</tr>
<tr>
<td>(2)</td>
<td>Bad</td>
<td>1.6</td>
</tr>
<tr>
<td>(3)</td>
<td>Bad</td>
<td>1.2</td>
</tr>
<tr>
<td>(4)</td>
<td>Bad</td>
<td>1</td>
</tr>
<tr>
<td>(5)</td>
<td>Good</td>
<td>4</td>
</tr>
<tr>
<td>(6)</td>
<td>Bad</td>
<td>0.6</td>
</tr>
<tr>
<td>(7)</td>
<td>Good</td>
<td>0.2</td>
</tr>
<tr>
<td>(8)</td>
<td>Good</td>
<td>6</td>
</tr>
<tr>
<td>(9)</td>
<td>Good</td>
<td>5</td>
</tr>
<tr>
<td>(10)</td>
<td>Bad</td>
<td>6.2</td>
</tr>
<tr>
<td>(11)</td>
<td>Bad</td>
<td>7.8</td>
</tr>
<tr>
<td>(12)</td>
<td>Bad</td>
<td>4.4</td>
</tr>
<tr>
<td>(13)</td>
<td>Bad</td>
<td>3.8</td>
</tr>
<tr>
<td>(14)</td>
<td>Good</td>
<td>6.6</td>
</tr>
<tr>
<td>(15)</td>
<td>Bad</td>
<td>7</td>
</tr>
<tr>
<td>(16)</td>
<td>Good</td>
<td>7.4</td>
</tr>
<tr>
<td>(17)</td>
<td>Bad</td>
<td>1.8</td>
</tr>
<tr>
<td>(18)</td>
<td>Bad</td>
<td>3.6</td>
</tr>
<tr>
<td>(19)</td>
<td>Bad</td>
<td>5</td>
</tr>
<tr>
<td>(20)</td>
<td>Good</td>
<td>5.2</td>
</tr>
</tbody>
</table>
between every two secondary investment object are bigger than 1 ($d_2 = 1$). Then we can get the following coverings.

\[
\begin{align*}
\bar{\delta}(1) &= \bar{\delta}(5) = \bar{\delta}(8) = \bar{\delta}(9) = \bar{\delta}(14) = \bar{\delta}(16) = \bar{\delta}(20) = \phi \\
\bar{\delta}_1(7) &= \{9, 11\}, \\
\bar{\delta}_2(7) &= \{11, 12\}, \\
\bar{\delta}_3(7) &= \{12, 14\}, \\
\bar{\delta}_4(7) &= \{12, 15\}, \\
\bar{\delta}_5(7) &= \{9, 16\}, \\
\bar{\delta}_6(7) &= \{12, 16\}, \\
\bar{\delta}_7(7) &= \{16, 19\}, \\
\bar{\delta}_8(7) &= \{16, 20\}. \\
\end{align*}
\]  

We next get the optimal portfolio from the \(\bar{\delta}_j(7), j = 1, 2, \ldots, 8\).

Using synthesizing effect function $S_2(u, v, w) = u e^{-aw} (1 + bw)$, $0 \leq a, b < \infty$, model (8) can be converted into the DDC-SRP model with discrete condition:

\[
\max_{\text{s.t. } (x)_N \subseteq M,} E \left( f \left( (x)_N, \eta \right) \right) e^{-a D \left( f \left( (x)_N, \eta \right) \right)} (1 + b C_D \left( (x)_N \right)),
\]

\[
\begin{align*}
|x - x_j| &> d_1, \forall x_j \in (x)_N/x, \\
|x_i - x_j| &> d_2, \forall x_i, x_j \in (x)_N/x.
\end{align*}
\]

(23)

Here $f \left( (x)_N, \eta \right) = \sum_{x \in (x)_N} w_j f \left( x, x_j \right)$, $w_j$ is the weight of the stock $j$.

Using genetic algorithm (its parameters setting are binary code mode; mutation probability is 0.001; crossover probability is 1; population size is 80; evolutionary generations is 1000), we can get Table 2.

Using genetic algorithm (its parameters setting are binary code mode; mutation probability is 0.01; crossover probability is 0.8; population size is 80; evolutionary generations is 1000), we can get Table 3.

Using genetic algorithm (its parameters setting are binary code mode; mutation probability is 0.1; crossover probability is 0.6; population size is 80; evolutionary generations is 1000), we can get Table 4.

By Tables 2–4, we can see that every portfolio contains three stocks (one core investment object and two secondary investment objects). Because the weights of core investment objects are 2 and weights of secondary investment objects are 1, the amount of investment of core investment object is 50,000 dollars, and amount of investment of each two secondary investment object is 25,000 dollar.

Obviously, in $S_2(u, v, w) = u e^{-aw} (1 + bw)$, $0 \leq a, b < \infty$, $a$ and $b$ are the parameters describing the uncertainty consciousness. When the weights of core investment objects are 2 and weights of secondary investment objects are 1, Table 2 shows the every solutions in different decision-making consciousness. Using expectation value model, the portfolio is $\bar{\delta}_5(7) = \{9, 16\}$. It means that stock (7) is core investment object, and stocks (9) and (16) are secondary investment objects. In the DDC-SRP model, when $a$ is smaller, the investor will pay more attention to the expectation of income in each stock. The result is same to the expectation value model. When $b$ is bigger, the investor will pay more attention to the risk of investment portfolio that can be represented by discrete degree.

With different mutation probability and crossover probability, the solution is different. When the mutation probability is smaller and the crossover probability is bigger, the kinds of solutions are smaller.

As above analysis, if we let $d_1 = 2, 1 < d_2 < 1.2$, we can get the following coverings.

\[
\begin{align*}
\bar{\delta}(1) &= \bar{\delta}(5) = \bar{\delta}(8) = \bar{\delta}(9) = \bar{\delta}(14) \\
&= \bar{\delta}(16) = \bar{\delta}(20) = \phi, \\
\bar{\delta}_1(7) &= \{9, 11\}, \\
\bar{\delta}_2(7) &= \{11, 12\}, \\
\bar{\delta}_3(7) &= \{12, 15\}, \\
\bar{\delta}_4(7) &= \{9, 16\}, \\
\bar{\delta}_5(7) &= \{12, 16\}, \\
\bar{\delta}_6(7) &= \{16, 19\}. \\
\end{align*}
\]

(24)

These coverings does not contain $\bar{\delta}_1(7) = \{12, 14\}$ and $\bar{\delta}_5(7) = \{16, 20\}$. Through calculating the results for the DDC-SRP model (13) with discrete condition $d_1 = 2, 1 < d_2 < 1.2$ is same as the DDC-SRP model with discrete degree $d_1 = 2, d_2 = 1$. 


When we select the DDC-SRP model with discrete degree $d_1 = 2, 1.2 \leq d_2 < 1.4$, using genetic algorithm, we can get the Tables 5–7.

In Tables 5–7, we can see that using the expectation value model, the portfolio is $\delta_1 ((7)) = [(9), (11)]$. In the DDC-SRP model, with different parameters $a$ and $b$, the optimal solutions are $\delta_1 ((7)) = [(9), (11)]$ or $\delta_2 ((7)) = [(11), (12)]$, which contain different decision-making consciousness. In the DDC-SRP model, when $a$ is smaller, the investor will pay more attention to the expectation of income in each stock. So the result is same to the expectation value model. Other situations show the investors will pay more attention to the income risk and the risk of investment portfolio, which is the same as Table 2. With different mutation probability and crossover probability, the solution is different. When the mutation probability is smaller and the crossover probability is bigger, the kinds of solutions are smaller.

When we select the DDC-SRP model with discrete degree $d_1 = 2, 1.4 \leq d_2 < 1.7$, the feasible region is only one covering $\delta_1 ((7)) = [(11), (12)]$. Therefore, the optimal solutions both in the expectation value model and the DDC-SRP model is $\delta_1 ((7)) = [(11), (12)]$, with different mutation probability and crossover probability. When we select the DDC-SRP model with discrete degree $d_1 = 2, d_2 \geq 1.7$, the feasible region is $\phi$ with different mutation probability and crossover probability.
Table 5: The results for the DDC-SRP model ($d_1 = 2, 1.2 \leq d_2 < 1.4$, $P_m = 0.001$, $P_c = 1$).

<table>
<thead>
<tr>
<th>Expectation value model ($a = 0, b = 0$)</th>
<th>Optimal solution</th>
<th>Expectation of objective</th>
<th>Variation of objective</th>
<th>Discrete degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = 0.1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 10$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
</tbody>
</table>

Table 6: The results for DDC-SRP model ($d_1 = 2, 1.2 \leq d_2 < 1.4$, $P_m = 0.01$, $P_c = 0.8$).

<table>
<thead>
<tr>
<th>Expectation value model ($a = 0, b = 0$)</th>
<th>Optimal solution</th>
<th>Expectation of objective</th>
<th>Variation of objective</th>
<th>Discrete degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = 0.1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 10$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
</tbody>
</table>

Table 7: The results for the DDC-SRP model ($d_1 = 2, 1.2 \leq d_2 < 1.4$, $P_m = 0.1$, $P_c = 0.6$).

<table>
<thead>
<tr>
<th>Expectation value model ($a = 0, b = 0$)</th>
<th>Optimal solution</th>
<th>Expectation of objective</th>
<th>Variation of objective</th>
<th>Discrete degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a = 0.1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 0.1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 1, b = 10$</td>
<td>$\hat{y}_1((7)) = [(11), (12)]$</td>
<td>22</td>
<td>10</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 0.1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 1$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
<tr>
<td>$a = 10, b = 10$</td>
<td>$\hat{y}_1((7)) = [(9), (11)]$</td>
<td>29</td>
<td>11</td>
<td>15.2</td>
</tr>
</tbody>
</table>

Using the DDC-SRP model, investors can comprehensively consider expectation of income, income risk, and risk of investment portfolio. The above give optimal solutions of the DDC-SRP model with synthesizing effect function $S_2(u, v, w) = ue^{-av}(1 + bw), 0 \leq a, b < \infty$ and discrete degree $d_1 = 2, d_2 \geq 1$, which is the division way of feasible region. When we use different synthesizing effect functions and different division way of feasible region, we can get more results.

6. Conclusion

In this article, a class of the stochastic programming model with rough feasible region is proposed to study the investment portfolio problem. The DDC-SRP model is developed that contains the expectation value model and improves the stochastic programming method. We analyze the convexity of the proposed model. With real-coded and random rough simulation genetic algorithm, we verify the variable of the solutions. Then the advantages and disadvantages of different solutions in different decision-making consciousnesses are shown. This method is more concise and effective than traditional programming solution models in investment portfolio problem [35].

Data Availability

The data used to support the findings of this study are included within the article.
Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Youth Top Talent Project of Research Project of Humanities and Social Sciences in Colleges and Universities of Hebei Province (BJ2021088), the Soft Science Research Project of Innovation Capacity Promotion Program of Hebei Province (21557629D), and the Research Project on the Development of Social Sciences in Hebei Province in 2021 (20210201325).

References
