**OPTIMAL INVESTMENT PORTFOLIO SELECTION USING THE SNAPSACK MODEL**

**González Santoyo F., Flores Romero B, Flores Juan J.**

Universidad Michoacana de San Nicolas de Hidalgo. Mexico.

{fsantoyo, betyf, juanf}@umich.mx

**Gil Lafuente A.M.**

Universidad de Barcelona (Spain)

amgil@ub.edu

**ABSTRACT**

Nowadays it is of fundamental importance to use optimization methods and techniques in financial decision making. These methods provide (in real time) effective and efficient information to take the investor to success with a lesser amount of uncertainty, as to maximize the benefit levels under high-risk scenarios. In this work we use the knapsack model as an optimal investment portfolio selection strategy. This is accomplished through the analysis of diverse scenarios obtained in real time. These scenarios provide the decision maker with information to make rational and efficient decisions.

**KEYWORDS:** optimization, investment portfolios, Knapsack, decision making.

1. **INTRODUCTION**

Since the beginning of mankind, the human being started financial and commercial transactions. The responsibility in making the right investment decisions relies mainly on experience and a good criterion by the investor. There are investors whose experience and ability make him so efficient, and take him to the right choices most of the time. Nevertheless, an experience-based decision is still a qualitative and not quantitative circumstance. This does not entail the inexistence of an experience-based decision that yields excellent outcome; excellent is not necessarily optimal, though.

Optimization techniques present an option to evaluate quantitatively all those systems amenable to be modeled mathematically. Through computer implementations we can obtain in a systematic way the exact magnitudes of the variables that yield optimal levels of the variables subject to optimization.

The problems in finance, specifically investment options problems, are of course feasible to model through mathematical equations. Therefore, we can apply optimization techniques to find the exact mixture of investments to yield maximum benefits. This fact transforms decision making from a problem whose solution depends on experience and criteria of the decision maker to a problem that can be solved systematically by a computer. Such a computer program can take as input the changing conditions of investment options and return the optimal misture of the investment options that maximize the benefits for the investor; this software will act as an assistant in the decision making or even an autonomous decision making tool.

Through the years, many optimization techniques have been developed, each with its own characteristics. Some of the most know techniques are lineal programming, integer programming, dynamic programming, etc. (González S.F. 1995, 2005, 2011). Other optimization methods or techniques studies in the last few years solve problems of different complexity. An outstanding problem, which has been extensively studied is the knapsack problem. The study of this problem has produced a diverse range of solution techniques applicable not only to it, but to other problems as well. These techniques are the object of study of the work presented in this paper.

1. **KNAPSACK MODELS**

You wish to invest all or part of ***c*** investment units, considering ***n*** possible investment options; each investment options yields **** and requires **** as minimum investment amount. The solution to the optimal investment is the set of selected investment options  such that the total benefit  be the maximum possible. This kind of optimization problem is known as the knapsack problem.

# Evolution of solution algorithms for the knapsack problem

During the last decades, the knapsack problem (KP) has been studied from different approaches. The first algorithms were produced in the 1950’s through Bellman’s dynamic programming. Dynamic programming solves exactly the binary knapsack problem. In 1957 Dantzing generates an efficient method to determine the solution to the continuous relaxation problem, obtaining an upper bound for ***z***, which was used the following 2o years in almost every KP study [Martello y Toth, 1990].

En the 1950’s Gilmore and Gomory investigated the approximation of dynamic programming to the solution of KP problems. In 1967 Kolestar presents the algorithm *Branch and Bound.* In the 1970’s the approximation method of Branch and Bound is completely developed; this is the only method capable of solving problems with a large number of variables. Among the most known algorithms in that period are those developed by Horowitz and Sahni, with contributions by Ingargiola and Korsh with the first reduction procedure, Johnson the firs polynomial approximation scheme. Shany extends this result to the binary knapsack problem. Martello and Toth proposed the first upper bound dominating the continuous relaxation value.

In the 1980’s the main results in this area aimed to solve large-scale problems. For this kind of problems sorting time take most of the execution time. In 1980 Balas and Zemel present a new approximation to solve the problem sorting (in many cases) only a small subset of variables, which represent the main part of the problem.

**MATHEMATICAL MODELS**

The knapsack problem can be stated mathematically, numbering variables from 1 to *n* and introducing a vector of binary variables  (*j=1, ... ,n*), where:

|  |  |
| --- | --- |
|  | (1) |

 represents the expected yield for investment option ,  represents the minimum amount required by that investment option, and  is the available capital in monetary units (knapsack size). The problem is to select from the binary vector all variables  that satisfy the constraint:

|  |  |
| --- | --- |
|  | (2) |

and maximize the objective function:

|  |  |
| --- | --- |
|  | (3) |

Equation (2) indicates that the sum of investments cannot be greater than the available capital. Equation (3) states that from all possible options we will choose the assignment that maximizes the total benefit.

One way to solve this problem is to analize all possible combinations and choose the one that satisfies Equations (2) and (3). Nonetheless, this procedure would be inefficient and intractable for large-size problems. In the last decades several exact and approximate algorithms have been developed to solve the knapsack problem.

# TYPES OF KNAPSACK PROBLEMS

Depending on the generality or specificity of the solution to the knapsack problem, several types of problems are derived. In general, we assume that the benefits (), the required amounts per investment option (), and the available capital (), are general integers. Nevertheless, the results can be extended to the real-valued case, and in most cases to non-positive values.

The different types of problems that can be generated are:

1. Binary Knapsack Problem (KP 0-1)
2. Multiple choice Knapsack Problem
3. Restricted Knapsack Problem
4. Unrestricted Knapsack Problem
5. Subset Addition Problem
6. Change Machine Problem
7. Binary Multiple Knapsack Problem
8. Generalized assignment Problem

This paper addresses the Binary Knapsack Problem KP 0-1 problem, which will be defined in the next section.

# BINARY KNAPSACK PROBLEM

The Binary Knapsack Problem is the most important knapsack problem and one of the most studied in discrete optimization. Mathematically, it can be described by Equations (1) TO (3). The reasons of its importance are the following:

* It can be seen as the simplest integer-programming problem.
* It appears as a sub-problem in many more complex problems.
* It can represent a large number of practical situations.

The solution to the Binary Knapsack Problem can be obtained using several exact and approximate algorithms, such as: Horowitz-Sahni’s greedy algorithm, Martello-Toth’s dinamyc programming, dynamic programming and polynomial approximation developed by Balas-Zemel, and Fayard-Plateau, among others. This type of problems is NP-Complete.

In this paper we use the greedy algorithm to solve the study case for the investment portfolios optimal selection.

# GREEDY ALGORITHM

The most immediate way to determine an approximate solution to the binary knapsack problem exploits the fact that the solution vector of selected investment options  has only one non-integer variable  (critical element: investment option that will not fit whole in the investment portfolio or knapsack; i.e., it would make the solution more expensive than the available capital). Unselecting this option () we can obtain an acceptable  very close to the optimal solution .

|  |  |
| --- | --- |
|  | (4) |

One would normally expect the approximate solution  to be very close to the exact solution . Nevertheless if the benefit the critical  yields, is relatively large one can obtain an improved heuristic solution  considering:

|  |  |
| --- | --- |
|  | (5) |

We assume the investment options to be sorted by yield value per monetary unit per investment option. That is,

|  |  |
| --- | --- |
|  | (6) |

1. **PROGRAM DEVELOPED**

This section presents Matlab® code of the implementation of the greedy algorithm to solve the binary knapsack problem.

*clear all*

*close all*

*% Program that implements the greedy algorithm*

*% to solve the Binary Knapsack Problem*

*%*

*% Reading data from keyboard*

*%*

*disp('Program that implements the greedy algorithm');*

*disp(' ');*

*n=input('Number of investment options n: ');*

*c=input('Total available investment capital c: ');*

*p=input('Benefit vector for each investment option {p}: ');*

*w=input('Vector of required amounts for each investment option {w}: ');*

*% The program produces zg (Optimal total benefit) and {x} (vector of selected options)*

*ct=c;*

*zh=0;*

*ja=1;*

*for j=1:n,*

 *if w(j)>ct*

 *x(j)=0;*

 *else*

 *x(j)=1;*

 *ct=ct-w(j);*

 *zh=zh+p(j);*

 *end*

 *if p(j)>p(ja)*

 *ja=j;*

 *end*

*end*

*if p(ja)>zh*

 *zh=p(ja);*

 *for j=1:n,*

 *x(j)=0;*

 *end*

 *x(ja)=1;*

*end*

*disp('Optimal total yield:')*

*zh*

*disp('Selected options vector:')*

*x*

1. **CASE ANALYSIS**

Case 1.

As a study case we will take a hypothetical example. Let “W” be the investor, with a financial resource of $25,000,000.00 mexican pesos, who whishes to obtain an optimal financial yield, given the investment options in the financial market. As a result of an investigation at the time of the investment decision, the weighing or importance levels for investment options 1, 2, 3, 4, 5, 6, and 7 in a per unit basis:  = (0.5, 0.02, 0.3, 0.431, 0.25, 0.49, 0.41) and the yield levels per unit are  = (1.5, 2, 5, 0.5, 3, 0.1, 1.1). The investor would like (a priori) to invest in all options. Nevertheless, to make the final investment decision he would like to know the optimal solution. The greedy algorithm determines quantitatively what are the most convenient options.

To solve this case using the greedy algorithm for the knapsack method, we need to sort the investment options according to Equation (6). After sorting, the investment options are shown in Table 1.

Table 1. Data for study case 1

|  |  |  |  |
| --- | --- | --- | --- |
| OPTION | **WEIGHT**  | **YIELD**  |  |
| 3 | $ 7,500,000.00 | $ 37,500,000.00 | 5.0 |
| 5 | 6,250,000.00  | 18,750,000.00 | 3.0 |
| 2 | 500,000.00 | 1,000,000.00 | 2.0 |
| 1 | 12,500,000.00 | 18,750,000.00 | 1.5 |
| 7 | 10,250,000.00 | 11,275,000.00 | 1.1 |
| 4 | 10,775,000.00 | 5,387,500.00 | 0.5 |
| 6 | 12,250,000.00 | 1,225,000.00 | 0.1 |

Once sorted, data is fed to the program shown in Section 3. The output results are the following:

**Total optimal benefit  $ 68, 525, 000**

**Selected Options: 2, 3, 5, and 7**

Table 2. Investment Portfolio for case 1

|  |  |
| --- | --- |
| INVESTMENT OPTION | AMOUNT |
| 2 | 500 000 |
| 3 | 7 500 000 |
| 5 | 6 250 000 |
| 7 | 10 250 000 |
| TOTAL | **24 500 000** |

**Case 2.**

As a second case let us assume the yield amounts per unit of investment are modified due to changes in the investment options. The investment options for this case are the following:  = (2, 1.1, 1.5, 0.7, 1, 2.5, 2.5). After sorting the investment options are shown in Table 3.

Table 3. Data for case 2.

|  |  |  |  |
| --- | --- | --- | --- |
| OPTION | **WEIGHT**  | **YIELD**  |  |
| 7 | $ 17,250,000.00 | $ 43,125,000.00 | 2.5 |
| 6 | 12,250,000.00  | 30,625,000.00 | 2.5 |
| 1 | 12,500,000.00 | 25,000,000.00 | 2.0 |
| 3 | 7,500,000.00 | 11,250,000.00 | 1.5 |
| 2 | 500,000.00 | 550,000.00 | 1.1 |
| 5 | 6,250,000.00 | 6,250,000.00 | 1.0 |
| 4 | 10,775,000.00 | 7,542,500.00 | 0.7 |

The application of the greedy algorithm to data in Table 2 yields the following results:

**Total optimal benefit  $ 54, 375, 000**

**Selected options: 3 and 7**

Table 4. Investment Portfolio for case 2

|  |  |
| --- | --- |
| INVESTMENT OPTION | AMOUNT |
| 3 | 7 500 000 |
| 7 | 17 250 000 |
| TOTAL | **24 750 000** |

Changing the yield values of the investment options, the selected options in the optimal investment portfolio change as well. The total optimal benefit also changes, but it is still optimal.

**CONCLUSIONS**

This paper uses an optimization technique known as the knapsack problem. This technique is applied to the optimal selection of investment portfolios as an exact quantitative technique.

The use of mathematical algorithms amenable to implementation to solve the optimal investment portfolio selection allows automation of decision-making. This implementation works in real time, taking into consideration the changing conditions of the investment options at every time.

When the problem is too big, i.e., when the number of investment options is big, the experience of a financial adviser may not be enough to produce the best selection of investment options. This problem gets worse in dynamic scenarios, where investment options change continuously. In these conditions a computer implementation of a greedy algorithm is of great help in the solution of the problem of optimal selection of investment options.

The application of these kind of modes is indispensable for real time decision-making, when the basic information of references of the investor is known.

**REFERENCES:**

Gil Aluja J., González Santoyo F., Flores R B., Flores R. J., *Techniques and Metodologies for Modelling and Simulation of Systems.* AMSE,UMSNH. México. (2005).

González S.F., Flores R. B., Gil Lafuente A.M., *Modelos y teorías para evaluación de inversiones empresariales.* UMSNH. México. (2010).

González Santoyo F., Flores Romero B., Gil Lafuente A.M., *Procesos para la toma de decisiones en un entorno globalizado.* Editorial Universitaria Ramón Areces. Madrid España. (2011).

González Santoyo F, Alfaro Calderón G., Flores Romero B., *Modelos clásicos para la selección de Carteras de Inversión*. Ciencia Nicolaita No. 15. Agosto 1997. pp.102-118. Morelia México. UMSNH. (1997).

González S.F., Alfaro C.G., *Las carteras de inversión en las decisiones financieras*. 1er Foro Nacional de Inv. En Disciplinas Administrativas. FCA-UNAM. México. D.F. (1996).

González Santoyo F., *Planeación de la producción usando técnicas de descomposición Jerárquica*. Tesis Doctoral (Doctor en Ingeniería-Investigación de Operaciones). Facultad de Ingeniería- División de Estudios de Posgrado UNAM, México. (1995).

González S.F.,*Optimización en el Diseño de Carteras de Inversión y su aplicación en la Industria*. Reporte interno de Inv. Coordinación de la Investigación Científica - UMSNH. Morelia, México. (1995).

Markowitz H , *Portfolio selection: efficient diversification of investments.* Jhon Wiley sons, New Cork. (1959).

Martínez Cárdenas F., González Santoyo F., Flores Romero V., Vargas Uribe G. ,*Selección óptima de carteras de inversión usando el problema Knapsack.* XIV International Conference AEDEM. Morelia México. (2005).

Silvano Martello; Paolo Toth, *Knapsack Problems Algorithms and Computer Implementations.* John Wiley & Sons. (1990).

Flores, Juan J.; Avila, Javier; Gonzalez, Federico; Flores, Beatriz; Gil Lafuente, Anna M. Fuzzy Near-Optimal Algorithm for the Determination of Investment Options. International Journal of Financial Decision Making. Vol.3. No.1. pp 1-12. Greece. (2007).

Flores, Juan J.; Avila, Javier; Gonzalez, Federico; Flores, Beatriz. Heuristic Analysis of Near-Optimal Algorithm for the Determination of Investment Options. Revista Matemática: Teoría y Aplicaciones (ISSN 1409-2433). Editorial de la Universidad de Costa Rica. Vol.13. No.2. pp 125-138. Costa Rica. http://revista.emate.ucr.ac.cr/index.php/revista/issue/view/27. (2006).