Proceedings of the 2010 International Conference on Industrial Engineering and Operations Management
Dhaka, Bangladesh, January 9 – 10, 2010

Stochastic Investment Decision Making with Dynamic Programming

Md. Noor-E-Alam[1] and John Doucette[1,2]
[1]Dept. of Mechanical Engineering, University of Alberta, Edmonton, AB, Canada
[2]TRLabs, Edmonton, AB, Canada

Abstract

Proper investment decision making is key to success for every investor in their efforts to keep pace with the competitive business environment. Mitigation of exposure to risk plays a vital role, since investors are now directly exposed to the uncertain decision environment. The uncertainty (and risk) of an investment is increasing with the increased number of competing investors entering to market. As a result, the expected return on investment (ROI) of a decision quite often carries a high degree of uncertainty. Our objective is to formulate a dynamic programming mathematical model for the investment decision with incorporating this uncertainty in a probabilistic manner. Policy iteration algorithm of the dynamic programming is adopted to solve the model. Our simulation result shows that the algorithm is able to help us in taking optimum investment decision.

Keywords
Investment decision, policy iteration, dynamic programming, uncertainty.

1. Introduction
At present, investment decision making is a critical task because every investment exhibits at least some amount of risk and uncertainty. These risks and uncertainties are the results of huge business competition and vibrant market economy. As a result, recent research in investment decision making is undergoing a paradigm shift with much integration of new techniques with existing methods to develop robust decision making processes [Heikkinen et al. (2009)]. Net present value (NPV) is the most common method in investment evaluation [Wang (1998)]. Alkaraan and Northcott (2006) have analyzed the use of conventional investment appraisal techniques such as payback, return on assets (ROA), return on investment (ROI), internal rate of return (IRR), NPV and risk analysis approaches such as sensitivity analysis, adjustment of the payback period, or discount rate. Handling risk and uncertainty in projects is currently one of the main topics of interest for researchers and practitioners working in the area of project investment decision making [Raz and Michael (2001)]. Topaloglou et al. (2002) and Rockafellar et al. (2000) solved Portfolio optimization problems with linear programs and they used Monte Carlo simulations to capture the risk in each associated investment. Xu-song and Jian-mou (2002) studied the investment decision-making of a project with deterministic dynamic programming. Yan and Bai (2009) formulated a deterministic dynamic programming model to allocate funds between stocks in a portfolio to maximize income. They captured the risk issues by incorporating the positive correlation between risks and returns of a stock to a large extent. Heikkinen and Pietola (2009) studied the use of stochastic programming approaches to make optimal investment decisions by modeling the problem as a Markov decision process. A dynamic uncertainty cost is presented with the modification of the classical expected value of perfect information to a dynamic setting. Dixit and Pindyck (1994) described the use of a *Markov decision process (MDP)* defined in continuous time and with a continuous state space for optimal investment decisions. Most of the prior research does not consider the inter-related dynamics of the systems that can be encountered by a stochastic dynamic investment model [Botterud et al. (2007)].The present research herein aims to handle this uncertainty and interrelated dynamics using infinite horizon stochastic dynamic programming and at the same time make optimal investment decisions for which the maximum total expected reward can be achieved.

1.1 Problem Description
A company may have number of ongoing projects, where each will provide some amount of fund available for reinvestment at end of the year. At the end of each year, every project has a certain amount of fund available to reinvest. The incremental amount of funds available to reinvest depends on the real return on investment for a specific project. Available funds can be reinvested back in the project from which it is originated or can be reinvested in other projects. Depending on the nature of the projects available, there are different levels of uncertainties related to the associated risks, which will ultimately affect the ROI. Depending on these uncertainties

associated with ROI, it is imperative to make decisions on how much one should reinvest in which project each year so that we can maximize the expected accumulated fund over an infinite time horizon. Therefore, our goal is to formulate the model for the above problem in a stochastic dynamic programming frame work, and solve it to find optimum investment decision at which we can maximize our long run expected fund. A brief outline of this paper is as follows: Section 2 describes dynamic programming and policy iteration algorithm used in our work, Section 3 describes a case study used to analyze the proposed method, and Section 4 ends with conclusions and future research avenues.

2. Stochastic Dynamic Programming

Dynamic programming is a systematic tool based on the simple idea of the principle of optimality [Bertsekas (2007)]. If a decision problem can be viewed as multiple stages with multiple states and known state transitions associated with each particular action, then the problem can be systematically solve with *deterministic dynamic programming* with the help of prominent Bellman equation [Bertsekas (2007)]. But if the state transition is probabilistic, then we can apply *stochastic dynamic programming*. Apart from this, if the number of stages is infinite (i.e., if we do not want to impose a limit on the number of stages), then the problem becomes an *infinite horizon stochastic dynamic programming problem (IHSDP)* [Bertsekas (2007)].

In our investment decision making, we are going to consider time as the various stages and the amount of accumulated funds as the state. The key factor for state transition is ROI and there is uncertainty associated with it that can be captured with a probability distribution, making up our state transition probabilities. As we are planning to make decision on how much one should reinvest in each project every year so that it maximizes the expected accumulated funds over an infinite time horizon, we should use IHSDP to make decisions based on longer-term potential benefits and avoid myopic reactions to short-term market fluctuations that may lead us to risky decision making.

Among the various approaches for IHSDP, the most widely used algorithms are *policy iteration, value iteration* and linear programming [Bertsekas (2007)]. In this paper we will use policy iteration technique to solve the investment decision problem. Throughout our work, the following notation will be used to describe this algorithm:

S = set of all states
$V^i(s)$ = value of the value function for state s and iteration i
$R(s,a)$ = reward for state s and any action a
$p[y|s,a]$ = transition probability from state s to y, if we take action a
α = discount factor
ε = tolerance parameter
μ^i = stationary policy at iteration i

2.1 Policy Iteration

The policy iteration algorithm calculates the optimum policy directly without the help of an optimal value function. Policy iteration converges faster in terms of the number of iterations (i), but every iteration will be computationally more expensive if we encounter a large number of states. This is due to the matrix inversion process for a large size transitional probability matrix. The following steps are involved in a *Policy Iteration Algorithm*:

Step 1: Start with any stationary action or policy we may have. These policies are state specific and state dependent.

Step 2: For that random action, perform policy evaluation, with equation-1. Equation-1 calculates the infinite horizon total expected reward. With following equation policy evaluation will check, whether the current policy is optimal or not.

$$V^{i}(s)=R(s,\mu^{i}(s))+\alpha\sum_{y\in S}p[y|s,a]V^{i}(y) \quad (1)$$

Step 3: If in step-2 the policy is found as not optimal, then perform policy improvement to select an optimum policy or action with the following operation. Following operation will select the optimum action from all available actions for which we can maximize our total reward.

$$\mu^{i+1}(s)=\arg\max_{a\in A(s)}\left[R(s,a)+\alpha\sum_{y\in s}p[y|s,a])V^{i}(y)\right] \quad (2)$$

Step 4: To check the optimality of the optimum action, check the following condition to see the improvement of value function:

$$\left\|V^{i}(s)-V^{i-1}(s)\right\|_{\infty}<\frac{1-\alpha}{\alpha}\varepsilon \quad (3)$$

If the above condition is met then stop, otherwise, return to step 2.

3. Case Study

We develop a case study by assuming reasonable test case values for all variables to use policy iteration algorithm for investment decision making:

A company has two ongoing projects. At the end of each year, every project has a certain amount of available funds to reinvest to its own project or other project. In this problem the amount of available funds for each project represents the state of the problem. For simplification suppose the company only has the capability to reinvest in a single project at a time. The reinvestment can be done to a particular project by utilizing the entire available fund of that project or by utilizing the entire available fund of that project plus some portion from the other project. Because, it is assumed that if there is no reinvestment in a project, then at least a $10,000 (or $10K) working capital is required for that project to run its own ongoing business. It is also assumed that there is no possibility to change the available fund level if there is no reinvestment in a project.

3.1 States

The amount of available funds in project-1 and project-2 represents the state of the problem. It is assumed that for each project the amount of available fund can be 10K, 11K,………50K. At some point, for instance, the state can be (13K, 20K), which means that the available funds in project-1 is 13K and in project-2 is 20K.

3.2 Actions

According to the above problem description there are four types of reinvestment actions we can consider for these two projects. For example, action-1 is to reinvest total amount of available fund in project-1 to project-1 only, i.e., if state is (13K, 22K) then reinvest the entire 13K to project-1 only. Action-2 is to reinvest total amount of available fund in project-1 and any excess amount above the working capital (assumed previously 10K) in project-2 to project-1 only i.e., if state is (13K, 22K) then reinvest 13K from project-1 and (22K-10K)=12k from project-2 to project-1. Action-3 and Action-4 are defined as accordingly.

3.3 State Transitions

As mentioned earlier there are different levels of uncertainties for ROI for each individual project. For this case study it is assumed that for project-1, there is a 20% probability that ROI is 5%, 60% probability that ROI is 10% and 20% probability that ROI is 20%. On the other hand, for project-2, there is 20% probability that ROI is 2.5%, 50% probability that ROI is 10% and 30% probability that ROI is 15%.

Suppose that the current state of the projects is (X,Y), which means the available funds in project-1 is X and project-2 is Y. For better understanding, state transition associated with action-1 for different probability (Pr) scenarios are shown in the following figure with following state transition rule equation-4:

$$X_{t+1}=X_t(1+i/100) \qquad (4)$$
$$Y_{t+1}=Y_t$$

; where i represents ROI, t represents present year

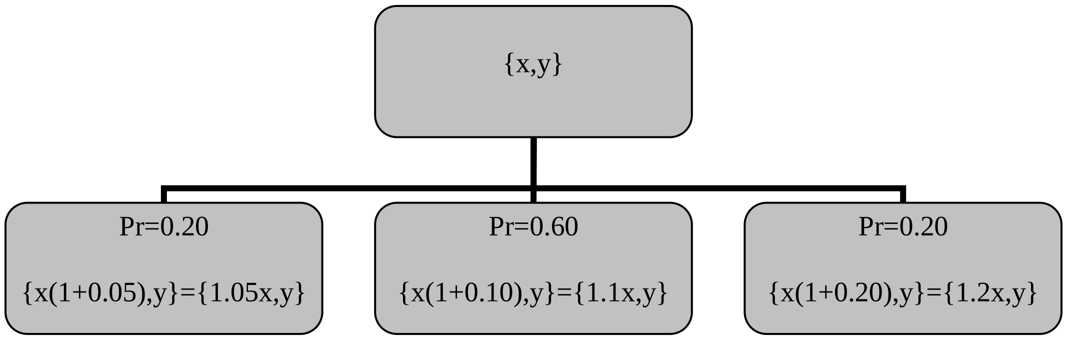

Figure 1: state transition diagram for action-1

3.4 Reward Calculation

We can gain a reward for each associated reinvestment decision depending on the state and the action we took. The following equation is used to calculate the expected reward.

$$\text{Expected Gain} = \sum (\text{Probability of ROI} \times \text{ROI} \times \text{Amount Reinvested}) \qquad (5)$$

3.5 Objective

Our objective is to select the optimal action (i.e., reinvestment decision) among these four options for each state to maximize the total gain or reward.

3.6 Result Analysis

Using MATLAB™, the transition probability matrix and reward matrix for each associated action is developed according to the problem description and transition rules. With the transition probability matrix and reward matrix, the problem is solved for discount factor α=0.9 and tolerance parameter ε=0.000001, using the policy iteration algorithm. To help determine whether the optimum policy works well or not, we can start from any state with a randomly chosen ROI and move with optimum policy which has been chosen previously. Following figure shows an example of state transition with optimum action.

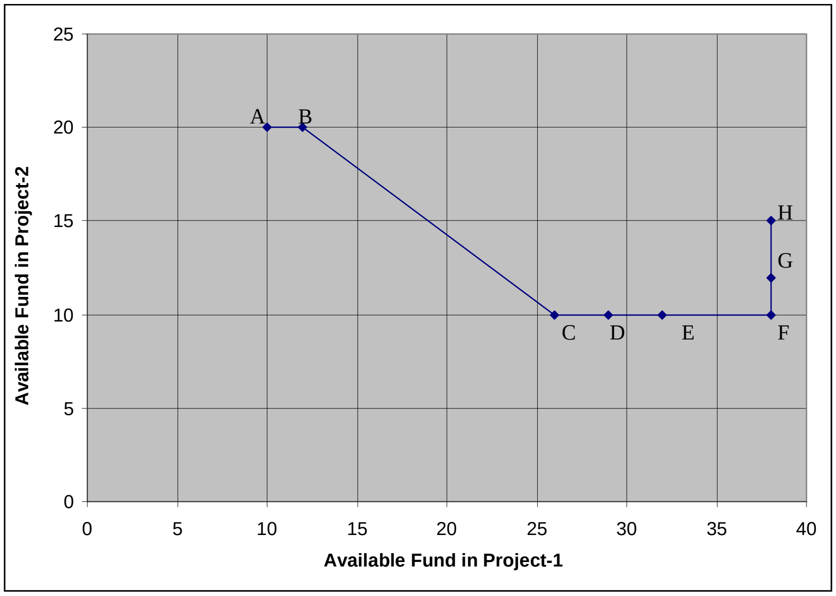

Figure 2: sample realization for optimum action (state-**A** starts with 10K, 10K)

In Figure 2, the horizontal axis represents the total amount of funds available in project-1 and vertical axis represents the total amount of funds available in project-2. In above figure **A, B, C, D, E, F, G, H** are representing different states. Among them **A** and **H** are initial and final states, and rest of them are intermediate states. To show the effect of optimum action we have started from state-**A** with its state dependent optimum action and reached to state-**B**. Again in state-**B**, we have taken its optimum action and reached to state-**C**. Similarly, we have reached to the final state-**H**. It is clearly observed that every time we move to a better state (i.e., the total amount of funds available in the two projects is higher than the previous year), that optimum policy is optimum for all states. This is because our objective is to maximize total reward, which is directly proportional to the value of state variables. In addition to that, it is seen from the above figure that our optimal policy is different for different states as we move forward; this is because in our offline implementation (i.e., selecting optimum policy with policy iteration algorithms), the optimum policy is chosen considering all uncertainty related to ROI and state transitions.

4. Concluding Discussion

In traditional methods, we can select the project with only the greater expected ROI, but it will often lead us to sub-optimal decisions as the expected return on investment (ROI) of a decision quite often carries a high degree of uncertainty with interrelated dynamics. This research aims to handle this uncertainty and interrelated dynamics using infinite horizon stochastic dynamic programming and at the same time make optimal investment decisions for which the maximum total expected reward can be achieved. To keep this in mind, we have formulated investment decision problem in a stochastic dynamic programming framework and used policy iteration technique to get the optimum investment decision. To test the feasibility of our proposed method we have run simulation with a case study by assuming reasonable test case scenario. Our simulation result shows that the algorithm is able to help us in taking optimum investment decision. In this method, we can move from state to state through a series of state-specific optimum policies obtained from policy iteration algorithm, leading to improved states at each step. This method of investment decision analysis makes investment decision by taking care of uncertainty in ROI and the state transition dynamics simultaneously. Therefore, we see that our optimal policy is different for different states as we move forward. Further investigation can be done to check the efficiency of policy iteration technique in terms of run time statistics and required number of iteration with other stochastic dynamic programming algorithms such as value iteration and linear programming.

References

1. Alkaraan, F., and Northcott, D., (2006) “Strategic capital investment decision-making: A role for emergent analysis tools? A study of practice in large UK manufacturing companies”, The British Accounting Review, vol. 38, pp. 149–173.
2. Bertsekas, D. P., (2007), “Dynamic Programming and optimal Control”, Third Edition, Volume II., Athena Scientific, Belmont, Massachusetts.
3. Botterud, A., and Korpa˚s, M. A., (2007), “A stochastic dynamic model for optimal timing of investments in new generation capacity in restructured power systems”, International Journal of Electrical Power and Energy Systems, vol. 29, pp. 163–174.
4. Dixit, A., and Pindyck, R., (1994), “Investment under Uncertainty”, Princeton University Press, Princeton.
5. Heikkinen, T., and Pietola, K., (2009), “Investment and the dynamic cost of income uncertainty: The case of diminishing expectations in agriculture”, European Journal of Operational Research, vol. 192, pp.634–646.
6. Raz, T., and Michael, E., (2001), “Use and benefits of tools for project risk management”, International Journal of Project Management, vol. 19, pp. 9–17.
7. Rockafellar, R.T., and Uryasev, S., (2000), “Optimization of conditional value-at-risk”, Journal of Risk, vol. 2, pp. 21–41.
8. Topaloglou, N., Vladimirou, H., and Zenios, S., (2002), “CVaR models with selective hedging for international asset allocation”, Journal of Banking and Finance, vol. 27, pp.1535-1561.
9. Wang, S.Q., (1998), “Evaluation and competitive tendering of power plant project in china”, Journal of Construction Engineering and Management, vol. 124, pp. 333–341.
10. Xu-song, X., and Jian-mou, W., (2002), “A Dynamic Programming Algorithm on Project-Gang Investment Decision-Making”, Wuhan University Journal of Natural Sciences vol. 7, no. 4, pp. 103-107.
11. Yan, F., and Bai, F., (2009), “Application of Dynamic Programming Model in Stock Portfolio-under the Background of the Subprime Mortgage Crisis”, International Journal of Business and management, vol.9, no 3, pp. 178-182.