Optimal Security Liquidation Algorithms

Sergiy Butenko
Department of Industrial Engineering, Texas A&M University, College Station, TX 77843-3131, USA

Alexander Golodnikov
Glushkov Institute of Cybernetics, 40 Acad. Glushkov av., Kiev-22, 252650, Ukraine

Stanislav Uryasev†
Risk Management and Financial Engineering Lab, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL 32611, USA

Abstract. This paper develops trading strategies for liquidation of a financial security which maximize the expected return. The problem is formulated as a stochastic programming problem, which utilizes the scenario representation of possible returns. Two cases are considered, a case with no constraint on risk and a case when the risk of losses associated with trading strategy is constrained by Conditional Value-at-Risk (CVaR) measure. In the first case, two algorithms are proposed; one is based on linear programming techniques, and the other uses dynamic programming to solve the formulated stochastic program. The third proposed algorithm is obtained by adding the risk constraints to the linear program. The algorithms provide path-dependent strategies which sell some fractions of security depending upon price sample-path of security up to the current moment. The performance of the considered approaches is tested using a set of historical sample-paths of prices.

1. Introduction

Consider the following decision making problem arising in finance. Given M shares of a risky security, develop a *trading strategy* for the complete liquidation of the asset over the discrete set of times $1, 2, \ldots, T$ that would maximize the expected return, while constraining the risk of losses.

By a trading strategy we will mean the list $x_1, x_2, \ldots, x_T$, where $x_t,\ t = 1, 2, \ldots, T$, is the fraction of shares that are liquidated at time t. An example of a "naive" trading strategy is to sell the same fraction of shares at each time period, $x_t = M/T,\ t = 1, 2, \ldots, T$. Bertsimas and Lo (Bertsimas and Lo, 1998) derived the conditions on price dynamics under which an analogous strategy for acquiring M shares of a security minimizes the expected cost of execution. More advanced trading strategies are considered in papers by Almgren (Almgren, 2003), Almgren and Chriss (Almgren and Chriss, 2000). They proposed some

† Corresponding author

predefined sequences of fractions of positions to be sold depending upon assumptions on such parameters as risk aversion of a trader, temporary and permanent impact cost parameters, and security volatility. These sequences of fractions were optimized in a mean-return/variance framework. A potential disadvantage of this approach is that positions do not depend on the price trajectory of the security. Also, risk is estimated using variance which does not distinguish favorable and adverse price movements. Suppose that the security price had a big advance in the favorable direction. In this case, a trader may want to "lock-in" the achieved returns and sell the security as soon as possible (but taking into account liquidity considerations). Also, in case of large adverse price movements, a trader may be forced to sell the security with a high speed (to prevent excessive losses). The reader is referred to (Bertsimas et al., 2000; El-Yaniv et al., 2001) for other recent related papers.

In this paper we propose closing strategies that use, at each time moment, the security price trajectory up to this moment. The results of our numerical experiments suggest that such strategies tend to outperform simple trading strategies, which do not utilize information on price changes.

This paper was motivated by the following considerations. A trading strategy is available which provides time moments for opening positions for several securities. Historical sequences of prices of these securities (after opening these positions) were recorded. These price sample-paths are input data for considered closing strategies. The question was how to process these data in an optimal way in order to construct a strategy for closing positions which utilizes specific probabilistic characteristics of the available price sample-paths. We answer this question using a natural assumption that price changes exhibit correlations across time periods. To deal with uncertainty, we elaborate the scenario representation of possible returns (Birge and Louveaux, 1997).

An important issue associated with any trading strategy is concerned with the impact of trader's activities on the security's price (Bertsimas and Lo, 1998; Almgren, 2003; Rickard and Torre, 1999). Usually two major types of price impact from trading are assumed: *permanent market impact* and *temporary market impact*. The first of them refers to permanent changes in security's price fluctuation caused by transactions of the investor. On the other side, the temporary market impact paradigm deals with the short-term price changes caused by the investor's trading activities, that don't influence the price of the asset in the future time moments of the investor's trading cycle. Although the permanent market impact may effect the performance of a trading strategy considerably under certain conditions, unfortunately, up to this moment, there is no generally accepted model addressing this issue. Alternatively, the temporary market impact can be represented by increased transaction costs in a trading model. While we propose models assuming both linear and nonlinear temporary market impact in this paper, the completed numerical experiments are limited to the linear case only at this point.

First, we consider a situation with no constraints on risk. In this case, two approaches are proposed. One is based on using the stochastic programming methodology (Birge and Louveaux, 1997) to reduce the problem to a (large scale) linear program, which can be efficiently solved with modern optimization software packages. The second approach utilizes decision trees and dynamic programming techniques.

Another case that we study in this paper is when a trader wishes to restrict the risk associated with the decision making process. We propose an algorithm obtained by including the constraints, which restrict the Conditional Value-at-Risk, in the linear program formulated for the case without risk constraints.

The remainder of this paper is organized as follows. In Section 2 we state the problem. In Section 3 we propose a method for generating scenarios which will be used in the models discussed in this paper. Sections 4 and 5 present mathematical models for the cases with no risk constraints and with C-VaR constraints, respectively. An alternative, dynamic programming approach to solve the LP formulated in the case without risk constraints can be found in Section 6. The results of numerical experiments with the proposed algorithms are discussed in Section 7. Finally, a conclusion is made in Section 8.

2. Problem Statement

Recall that we have M shares of a single risky asset to be liquidated over the set of times $1, 2, \dots, T$. The major uncertainty we deal with is price of the asset at each time moment $t,\ t = 1, 2, \dots, T$. We describe this uncertainty by a random variable $R_t(\omega)$ which is the return at time moment t, defined as the fraction of the stock prices observed at times t and $t-1,\ 1 \leq t \leq T$. Assume that $\omega \in \Omega$ is a discrete space of random elements. Let $p(R_1(\omega))$ be the probability of the return $R_1(\omega)$ occurring at time $t = 1$, and $p(R_t(\omega)|R_1(\omega), R_2(\omega), \dots, R_{t-1}(\omega))$ be the conditional probabilities of the return $R_t(\omega)$ occurring at time t, provided that the returns up to time t were $R_1(\omega), R_2(\omega), \dots R_{t-1}(\omega),\ t \geq 2$.

We assume that at each time t the number l_t of different possibilities of scenario development may occur. As an illustrative example we may choose two scenarios ($l_t = 2$ for all t), which represent the possibilities

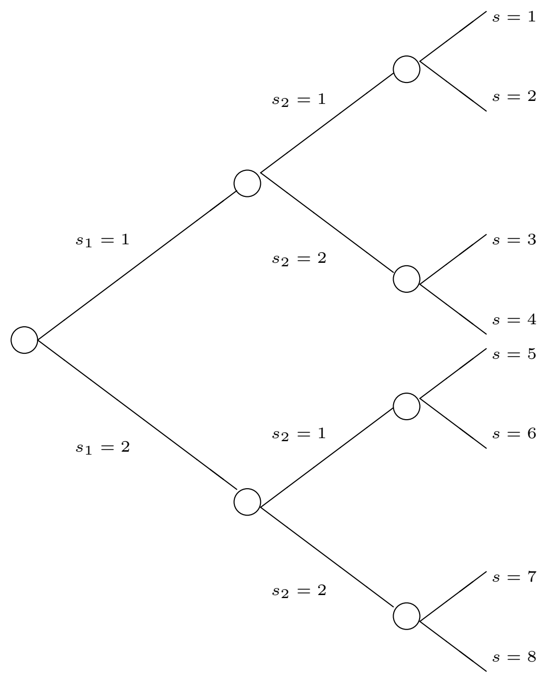

Figure 1. The scenario tree, $T = 3; l_t = 2, t = 1, 2, 3$.

of having nonnegative ($R_t \geq 1$) and negative ($R_t < 1$) return. Then, if $T = 3$, eight possible scenarios may occur. We enumerate the scenarios using an index $s = 1, 2, \ldots, 8$, which represents a set of outcomes ω that have common return in our model. Denote by $S = \{s(\omega), \omega \in \Omega\}$ the set of all scenarios. Now instead of the base element ω we can use more specific s. Each scenario s can be described as a 3-dimensional vector $s = (s_1, s_2, s_3)$, where

$$s_t = \begin{cases} 1, & \text{if } R(t, s) \geq 1, \\ 2, & \text{otherwise}; \end{cases}$$

$t = 1, 2, 3$. For example, $s = (1, 1, 1)$ for scenario 1, and $s = (2, 1, 2)$ for scenario 6. The eight scenarios are represented by the tree in Figure 1.

Tree of scenarios can be divided into branches, which are formed by sets of overlapping scenarios. For example, scenarios 1 and 3 have the same return for $t = 1$, therefore they belong to the same first branch.

Let the vector of returns corresponding to a scenario s be

$$(R_1, R_2, \ldots, R_T).$$

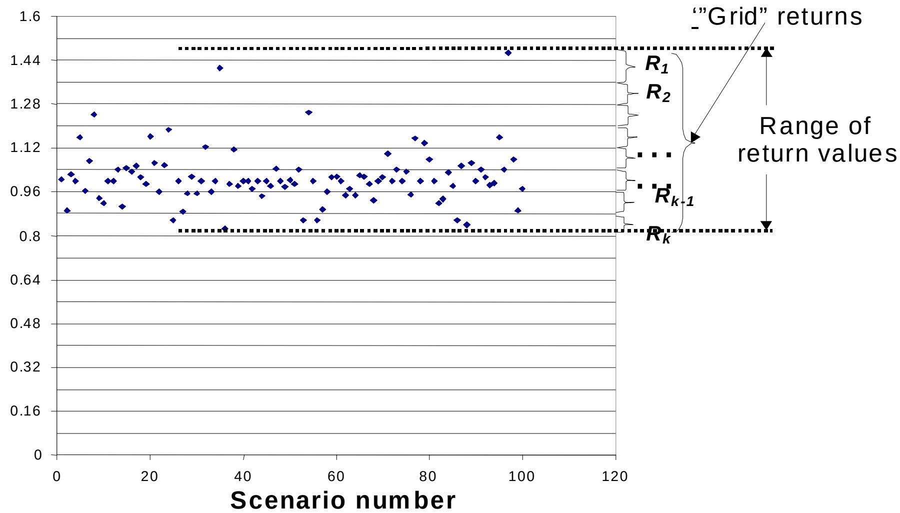

Figure 2. Projection of historical returns to discrete grid. Points designate historical returns at the first day for 100 scenarios.

Denote by $R(t, s)$ the return, in % of the initial asset price, for a scenario s at time t. Then

$$R(t, s) = R_1 \times R_2 \times \ldots \times R_t.$$

The probability of a scenario s can be determined from the following expression:

$$p(s) = p(R_1) \times p(R_2|R_1) \times \cdots \times p(R_T|R_1, R_2, \ldots, R_{T-1}).$$

3. Generating Scenarios and Estimating Probabilities

To utilize the approaches proposed in this paper, we need to generate for each step the set of possible outcomes and estimate probabilities of their occurring. For this purpose we can utilize historical data. The historical data we are working with consist of a number of the so-called sample paths, each of which reflects an asset price movement over discrete times.

Suppose that we are given a set of N sample paths of length T. A common way of transforming historical data into the model scenarios is to define a grid of returns and then project sample-path returns to the grid (see Figure 2). As a result, a set of historical returns $\{R_t(\omega),\ \ \omega \in \Omega\}$ occurred at time t is transformed into the set of "grid"

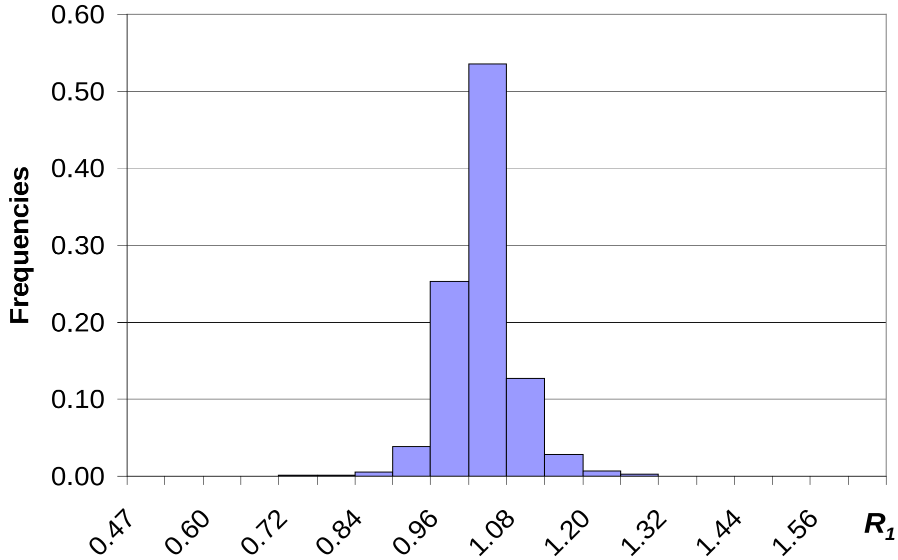

Figure 3. Frequencies of Occurring Different Values of Return R_1 at the First Step

returns $\{R_{t1}, R_{t2}, \ldots, R_{tk}\}$. According to this procedure of scenario generation the natural way of estimating probabilities is by calculating the frequencies. Frequencies can be calculated using the following formulas:

$$p(R_1) = n(R_1)/N,$$
$$p(R_2|R_1) = n(R_2|R_1)/n(R_1),$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$p(R_t|R_1, R_2, \ldots, R_{t-1}) = n(R_t|R_1, \ldots, R_{t-1})/n(R_1, \ldots, R_{t-1}),$$

where N is the total number of the historical sample paths; $n(R_1)$ is the number of sample paths with projected return equal to R_1 at time $t = 1$; $n(R_2|R_1)$ is the number of sample paths with projected return equal to R_1 at time $t = 1$ and with return equal to R_2 at $t = 2$; and so on.

In this particular study we are working with 5400 sample paths, each of which consists of five data points, corresponding to five trading days. The results of calculating frequencies based on this dataset are shown in Figures 3-4. The histogram shown in Figure 3 represents the frequencies of realization of different values of return at time $t = 1$. It has a shape of the plot of a lognormal probability density function.

The histogram shown in Figure 4 represents the estimate of conditional probabilities of occurring different values of return R_2 at the second step provided that the return at time $t = 1$ was $R_1 = 0.9$. An analysis of Figure 4 reveals that there are bins which contain insufficient return occurrences to derive reliable estimate of appropriate

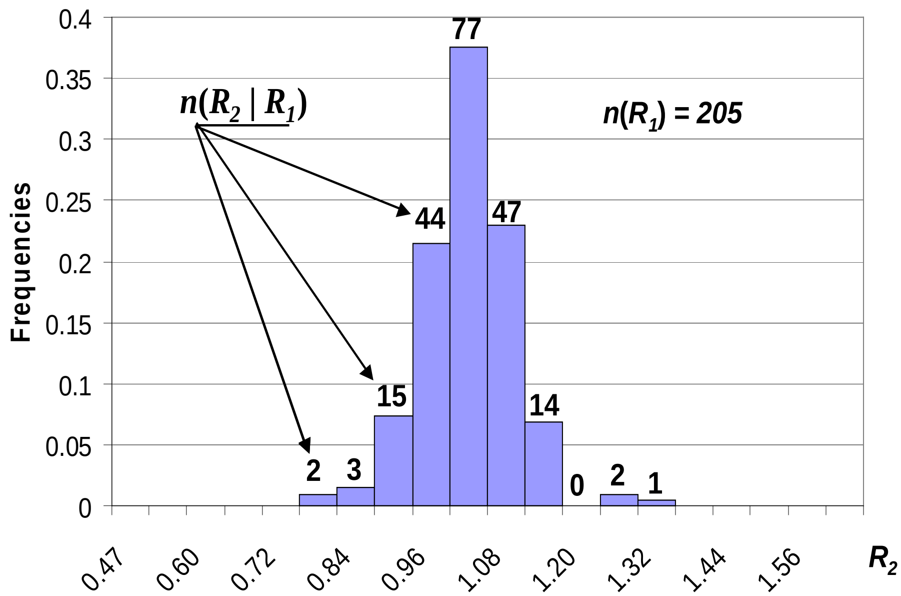

Figure 4. Conditional Frequencies of Occurring Different Values of Return R_2 at the Second Step Provided that Realization of Return at the First Step is $R_1 = 0.9$

probabilities; for example, there are zero frequencies for some bins. This may lead to an incorrect (too rough) model. In order to overcome this difficulty, we will use an alternative approach based on probabilistic modeling.

Following the common practice in finance applications, we assume that returns observed at times $1, 2, \ldots, T$ have a non-singular multivariate lognormal distribution. This means that the random vector $\xi = (\xi_1, \xi_2, \ldots, \xi_T)$, $\xi_t = \ln(R_t)$, $1 \leq t \leq T$, is multivariate normal with the mean

$$\mu = (\mu_1, \mu_2, \ldots, \mu_T)$$

and the covariance matrix

$$\Sigma = (\sigma_{ij},\ i, j = 1, 2, \ldots T).$$

The parameters of lognormal distribution can be estimated using the historical data. The estimate of vector of means can be determined by the following expressions:

$$\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_T), \quad \widehat{\mu}_i = \frac{1}{N} \sum_{n=1}^{N} x_{ni},$$

where, as before, N is the number of scenarios in the dataset; x_{ni} is the logarithm of return at time $t = i$ under the n-th scenario. This estimate does not take into consideration the time-dependent trend, and can be used as a rough estimate of μ. However, more sophisticated techniques based on time series analysis can be used for calculating the estimates of mean values (Anderson, 1971). Unbiased estimates of the components σ_{ij} of the covariance matrix Σ are

$$\widehat{\sigma}_{ij} = \frac{1}{N-1}\sum_{n=1}^{N}(x_{ni} - \widehat{\mu}_i)(x_{nj} - \widehat{\mu}_j).$$

Using the described probabilistic model we can estimate the probabilities

$$p(R_1), p(R_2|R_1), \ldots, p(R_t|R_1, R_2, \ldots, R_{t-1}).$$

To obtain estimates of $p(R_1)$ we use the marginal distribution of $\ln(R_1)$, which is normal with mean $\widehat{\mu}_1$ and variance $\widehat{\sigma}_{11}$.

To obtain estimates of conditional probabilities

$$p(R_t|R_1, R_2, \ldots, R_{t-1}),\ 1 < t \leq T,$$

we use the conditional distributions of $\ln(R_t)$, which are normal with the conditional mean (see (Anderson, 1958)):

$$E[\ln(R_t)|\ln(R_1), \ldots, \ln(R_{t-1})] = \upsilon(x^{(1)}) = \mu_t + \Sigma_{t,t-1}\Sigma^{-1}_{t-1,t-1}(x^{(1)} - \mu^{(1)}) \quad (1)$$

and variance

$$\tilde{\sigma}_{tt} = \sigma_{tt} - \Sigma_{t,t-1}\Sigma^{-1}_{t-1,t-1}\Sigma_{t-1,t}, \quad (2)$$

where

$$x^{(1)} = (x_1, x_2, \ldots, x_{t-1}) = (\ln(R_1), \ln(R_2), \ldots, \ln(R_{t-1}));$$
$$\mu^{(1)} = (\mu_1, \mu_2, \ldots, \mu_{t-1}); \quad \Sigma_{tt} = (\sigma_{ij},\ \ i, j = 1, \ldots, t);$$
$$\Sigma_{tt} = \begin{pmatrix} \Sigma_{t-1,t-1} & \Sigma_{t-1,t} \\ \Sigma_{t,t-1} & \sigma_{tt} \end{pmatrix};$$
$$\Sigma_{t,t-1} = (\sigma_{t1}, \sigma_{t2}, \ldots, \sigma_{t,t-1}) = \Sigma'_{t-1,t}.$$

Probabilities $p(R_t = R_{tj}|R_1, \ldots, R_{t-1})$ are found from the obtained conditional distributions.

4. Mathematical Programming Models

Let $x(t, s)$ be a portion of the total number of shares to be liquidated at time t under a scenario s. We introduce a generally nonlinear objective function, which is the expectation of the total return, minus the transaction cost $c(t, s, x(t, s))$ (including the temporary market impact):

$$\max \sum_{s \in S} p(s) \sum_{t=1}^{T} \{R(t, s)x(t, s) - c(t, s, x(t, s))\}. \qquad (3)$$

Obviously, $c(t, s, x(t, s))$ depends on the number M of shares at hand, as well as on the total amount of the security's shares available in market. Konno and Wijayanayake (Konno and Wijayanayake, 2002) considered the transaction cost function, as a function of the traded amount x, concave up to a certain point x', and convex for $x \geq x'$. However, in modern market conditions, the concave part, corresponding to transactions of small amounts, can be neglected. This leaves us with a convex *slippage cost* function. With this assumption, the objective function (3) is obviously a concave function, and its global maximum, subject to linear constraints (which we will introduce later), can be efficiently computed using the well-established methods of convex optimization. We can also linearize this model as follows. Concave functions $f_{ts}(x) = R(t, s)x(t, s) - c(t, s, x(t, s)$ can be well approximated by some piecewise-linear concave functions for each t, s:

$$f_{ts}(x) = R(t, s)x(t, s) - c(t, s, x(t, s) \approx \min_{i=1,\ldots,k} d_i(t, s)x(t, s) + r_i(t, s),$$

where k is the number of linear pieces in each function, and $d_i(t, s)$ and $r_i(t, s)$ are some coefficients. Therefore, the objective function in (3) can be replaced by the following linear function subject to linear constraints:

$$\max \sum_{s \in S} p(s) \sum_{t=1}^{T} z(t, s) \qquad (4)$$

subject to

$$d_i(t, s)x(t, s) + r_i(t, s) \geq z(t, s). \qquad (5)$$

For the sake of simplicity of presentation, in the remainder of this paper we assume a *linear transaction cost*,

$$c(t, s, x(t, s)) = k_c(t, s)x(t, s)$$

for some coefficient $k_c(t, s)$. This assumption transforms the general objective (3) to the following linear function:

$$\max \sum_{s \in S} p(s) \sum_{t=1}^{T} (R(t, s) - k_c(t, s))x(t, s). \qquad (6)$$

Denoting by $\mathbf{R}(t,s)$ the modified return,

$$\mathbf{R}(t,s) = R(t,s) - k_c(t,s),$$

we obtain the following linear objective function:

$$\max \sum_{s\in S} p(s) \sum_{t=1}^{T} \mathbf{R}(t,s)x(t,s). \tag{7}$$

The following constraints have to be satisfied.

a) Balance constraints, which assure that the asset will be completely liquidated by the time T:

$$\sum_{t=1}^{T} x(t,s) = 1, \quad \forall s \in S. \tag{8}$$

b) Nonanticipativity constraints, linking the separate scenarios which have common paths up to a certain time:

$$\text{If } \ s'_t = s''_t \ \ \forall t \le t_0 \ \text{ then } \ x(t,s') = x(t,s''), \ \ t \le t_0. \tag{9}$$

c) Nonnegativity constraints:

$$x(t,s) \ge 0, \qquad 1 \le t \le T, \quad \forall s \in S. \tag{10}$$

Problem (7)-(10) is a linear program (LP) whose solution determines, for each scenario s and each time t, the fraction $x(t,s)$ of the total number of shares to be liquidated.

If the trader is interested in liquidating the asset in a single transaction, this can be modeled by the following additional constraints:

$$x(t,s) \in \{0,1\} \quad \forall t,s. \tag{11}$$

Thus, we obtain an integer programming problem (7)-(11).

Next, we will show that all extreme points of the feasible region of the above LP are integer. In order to show this, we will represent an arbitrary noninteger feasible point as a linear combination of two other feasible points. Let x be a feasible point of (7)-(10) such that $x(t',s') \notin \{0,1\}$ for some t', s'. Without loss of generality, we also assume that t' is the earliest time moment corresponding to a noninteger component of $x(\cdot, s')$ for the scenario s'. We define two feasible points, $\bar{x}$ and $\hat{x}$, as follows:

$$\bar{x}(t,s) = \begin{cases} 0, & \text{if } (t,s) = (t',s'); \\ \frac{x(t,s)}{1-x(t',s')}, & \text{if } t > t' \text{ and } s_i = s'_i \ \forall i \le t'; \\ x(t,s), & \text{otherwise}; \end{cases}$$

$$\hat{x}(t,s) = \begin{cases} 1, & \text{if } (t,s) = (t',s'); \\ 0, & \text{if } t > t' \text{ and } s_i = s'_i \ \forall i \le t'; \\ x(t,s), & \text{otherwise.} \end{cases}$$

Then for any t, s, $x(t,s) = (1-x(t',s'))\bar{x}(t,s)+x(t',s')\hat{x}(t,s)$, therefore x is not a vertex of the polyhedron defined by the LP constraints. Thus, all extreme points of the feasible region are integer, so (7)-(10) will always have a 0-1 optimal solution.

5. Adding the CVaR Constraints

In addition to constraints (8)-(10) a trader could wish to restrict the risk associated with his decision. There are several approaches for managing risk. We choose the Conditional Value-at-Risk (CVaR) to measure the risk, due to its clear engineering meaning and possibility to implement it using linear constraints (Rockafellar and Uryasev, 2000; Rockafellar and Uryasev, 2002).

Assume that the trader liquidates a certain portion $x(1,s),\ s \in S$, of the shares at time $t = 1$, then the value of the portfolio consists of the cash obtained for the liquidated shares, and the cost of the shares left.

The set of returns associated with the cash obtained at time $t = 1$ is $\{\mathbf{R}(1,s) \cdot x(1,s), s \in S\}$, and the set of returns associated with the remaining shares is

$$\{\mathbf{R}(1,s) \cdot (1 - x(1,s)),\ s \in S_1\} = \{[\mathbf{R}(1,s) \cdot \sum_{t=2}^{T} x(t,s)],\ s \in S\},$$

because of the constraint (8). Therefore, the set of returns associated with all shares at time $t = 1$ is

$$\{\mathbf{R}(1,s),\ \ s \in S\}.$$

For $t > 1$ the set of returns associated with the asset is

$$\{\ [\sum_{\tau=1}^{t-1} \mathbf{R}(\tau,s) \cdot x(\tau,s) + \sum_{\tau=t}^{T} \mathbf{R}(t,s) \cdot x(\tau,s)],\ \ s \in S\ \}.$$

We define the loss function at time t as the negative value of return:

$$f_t(x,s) = -[\sum_{\tau=1}^{t-1} \mathbf{R}(\tau,s) \cdot x(\tau,s) + \sum_{\tau=t}^{T} \mathbf{R}(t,s) \cdot x(\tau,s)], \tag{12}$$

$$s \in S, \;\; 1 \leq t \leq T,$$

and impose the requirement, that for some confidence level α, $\alpha-$CVaR of this loss functions at each step $1 < t \leq T$ has to be no greater than some predefined threshold ω. To express this as linear constraints, let us introduce the following function (Rockafellar and Uryasev, 2000):

$$F_\alpha(x, \varsigma_t, t) = \varsigma_t + (1-\alpha)^{-1} \sum_{s \in S} [f_t(x,s) - \varsigma_t]^+ p(s) \tag{13}$$

where $[t]^+ = \max\{0, t\}$.

Then $\alpha-$CVaR can be determined from the formula

$$\phi(x,t) = \min_{\varsigma_t \in R} F_\alpha(x, \varsigma_t, t).$$

Using this notation, the constraints on risk can be stated as follows:

$$\phi(x,t) \leq \omega \;\; \text{for } 1 < t \leq \mathrm{T}. \tag{14}$$

According to (Rockafellar and Uryasev, 2000), the problem (7)-(10), (14) can be rewritten in the form

$$\min_{x, \varsigma_2, \ldots \varsigma_T} [-\sum_{s \in S} p(s) \sum_{t=1}^{T} \mathbf{R}(t,s) x(t,s)] \tag{15}$$

subject to constraints (8)-(10) and

$$\varsigma_t + \frac{1}{J_t(1-\alpha)} \sum_{s \in S} [f_t(x,s) - \varsigma_t]^+ \;\leq \omega \;\; \text{for} \;\; 1 < t \leq T. \tag{16}$$

By introducing additional variables we reduce constraints (16) to

$$\varsigma_t + \frac{1}{J_t(1-\alpha)} \sum_{j=1}^{J_t} z_{jt} \leq \omega \tag{17}$$

$$z_{jt} \geq f_t(x, s_{jt}) - \varsigma_t, \quad z_{jt} \geq 0, \;\; s_{jt} \in S, \;\; 1 < t \leq T. \tag{18}$$

Thus, we obtain the problem with the objective (15) subject to constraints (8)-(10), (17), and

$$z_{jt} \geq -[\sum_{\tau=1}^{t-1} \mathbf{R}(\tau, s_{jt}) \cdot x(\tau, s_{jt}) + \sum_{\tau=t}^{T} \mathbf{R}(t, s_{jt}) \cdot x(\tau, s_{jt})], \tag{19}$$

$$z_{jt} \geq 0, \;\; s_{jt} \in S, \; 1 < t \leq T. \tag{20}$$

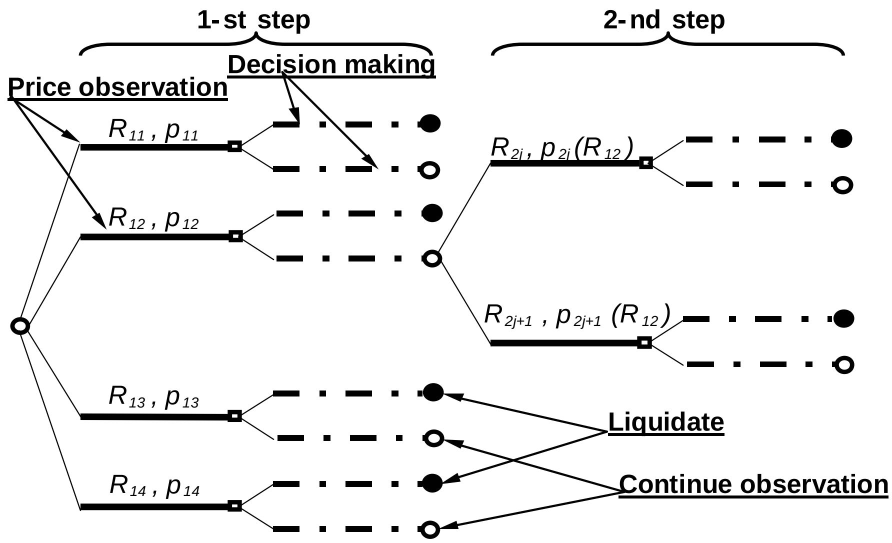

Figure 5. Example of a decision tree. Solid lines designate possible realizations of random returns, dashed lines designate decision opportunities.

6. Dynamic Programming Approach

In this section, we discuss how the stochastic program (7)-(11), can be solved using dynamic programming techniques. The main advantage of using this method is that it allows us to substantially reduce the size of the problem, as will be discussed below. However, this approach is applicable only for the model without risk constraints.

The problem (7)-(11) can be represented by the following decision tree, which is a graphical model of the stochastic programming problem. Figure 5 illustrates a two-step decision making process. There are two types of nodes: circles represent realizations of random returns and squares represent decision making. Solid lines show the results of observation of random returns (outcomes) and dashed lines show the decision opportunities. Possible outcomes and probabilities of their occurrence are shown above the corresponding solid lines.

At each time moment t we first observe one of the l_t possible realizations of the random return (including the transaction costs), which occur with given probabilities. Then we should choose one of the two possible actions: (1) to liquidate the asset completely, or (2) to continue observations. If we decide to liquidate the asset after observing the return $\mathbf{R}(t, j)$ at time t, then the wealth increases $\mathbf{R}(t, j)$ times.

To analyze a decision tree, we start with its right side, and work backwards to the left. At each step, for a given decision node, the method calculates the profit of each decision opportunity and selects the decision with the largest benefit. Then, using recursive formulas, the calculations are propagated towards the left side.

6.1. Reduction of the Problem Size

One of the main disadvantages of both the stochastic programming and dynamic programming approaches is dramatic increase of the size of the tree with increasing accuracy of the model (number of outcomes at each step). In order to handle large-scale cases we develop a special approach.

Let t be a predefined time of the asset liquidation. We introduce the following notations:

$$\begin{aligned}
\mu^{(2)} &= (\mu_2, \ldots, \mu_t);\\
\Sigma_{tt} &= \begin{pmatrix} \sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix};\\
\Sigma_{22} &= (\sigma_{ij}, \;\; i, j = 2, \ldots, t);\\
\Sigma_{12} &= (\sigma_{12}, \ldots, \sigma_{1t}) = \Sigma_{22}';\\
\xi^{(2)} &= (\ln(R_2), \ldots, \ln(R_t)).
\end{aligned}$$

Let $D = (1, \ldots, 1)$ be a $(t-1)$-dimensional vector. Then the conditional mean of the logarithm of the combined return obtained at times in the interval $[2, t]$ is expressed by (see (Anderson, 1958))

$$z_t = E\{\ln(R_2 \cdot \ldots \cdot R_t)|R_1\} = D[\mu^{(2)} + \Sigma_{21}\sigma_{11}^{-1}(\ln(R_1) - \mu_1)], \quad (21)$$

and the variance is equal to

$$\sigma_{z_t}^2 = D(\Sigma_{22} - \Sigma_{21}\sigma_{11}^{-1}\Sigma_{12})D.' \quad (22)$$

According to a property of lognormal distribution, the conditional mean of return obtained during at times in the interval [2,t] is equal to

$$E(\prod_{j=2}^{t} R_j|R_1) = \exp(z_t + \sigma_{z_t}^2/2). \quad (23)$$

From (1)-(21) it follows, that a trading strategy liquidating the security at time t, $t > 1$ yields a higher expected return than the strategy liquidating the asset at time $t = 1$, if the following inequality holds

$$\begin{aligned}
&z_t + \sigma_{z_t}^2/2 = D[\mu^{(2)} + \Sigma_{21}\sigma_{11}^{-1}(\ln(R_1) - \mu_1)]+ \\
&+[D(\Sigma_{22} - \Sigma_{21}\sigma_{11}^{-1}\Sigma_{12})D']/2 > 0.
\end{aligned} \quad (24)$$

Therefore, an optimal trading strategy does not liquidate the security at time $t = 1$, if

$$\max_{2 \le t \le T} \{z_t + \sigma_{z_t}^2/2\} > 0. \tag{25}$$

We obtain the following theorem.

THEOREM 6.1. *An optimal trading strategy does not liquidate the security at time $t = 1$, if inequality (25) holds. All scenarios satisfying (25) can be excluded from consideration without changing the optimal solution at time $t = 1$.*

The above theorem provides a method for reduction of the problem size. Theorem 6.1 says that in order to find the decision corresponding to an optimal trading strategy at $t = 1$, one does not have to analyze the whole decision tree. Only the scenarios for which (25) is not satisfied have to be analyzed; for the rest of the scenarios $x(1, s) = 0$ ("continue observation") in the optimal trading strategy. This point is discussed in more detail in the next section.

7. Case Study

Our case study is based on historical data consisting of 5-day sample-path scenarios for a set of stocks (each path is based on a 5-day opening price trajectory of a single stock). The total number of sample paths is 5400. Table I represents typical scenarios under consideration.

Table I. Examples of sample-path scenarios.

1 day	2 day	3 day	4 day	5 day
1.007	0.979	1.038	1.002	1.016
0.891	1.151	1.081	0.948	1.073
1.023	1.053	0.993	0.978	0.963
1.000	1.056	1.030	1.007	0.970
...	...	...	...	...

Based on the historical data, we estimated the mean values $E(\ln(R_t)),\ 1 \leq t \leq 5$, and the covariance matrix:

$$\widehat{\mu} = (0.0059,\ 0.0005,\ 0.0033,\ 0.0045,\ 0.0004),$$

$$\Sigma = \begin{pmatrix} 0.00309 & -0.00024 & -0.00034 & -0.00023 & -0.00016 \\ -0.00024 & 0.00265 & -0.00015 & -0.00020 & -0.00004 \\ -0.00034 & -0.00015 & 0.00224 & 0.00002 & 0.00001 \\ -0.00023 & -0.00020 & 0.00002 & 0.00227 & -0.00003 \\ -0.00016 & -0.00004 & 0.00001 & -0.00003 & 0.00226 \end{pmatrix}.$$

In all of our numerical experiments we used

$$k_c(t,s) = 0.01R(t,s)$$

in (6), which corresponds to transaction costs equal to 1% of the profit. This yields $\mathbf{R}(t,s) = 0.99R(t,s)$.

7.1. LP models

First, consider the linear programming models. We generated an instance of dimension $4 \cdot 4 \cdot 4 \cdot 4 \cdot 4$; for the model with CVaR constraints the confidence value $\alpha = 0.9$, and predefined level $\omega = 0.9$ were chosen.

For this instance, we solved the problem (15), (8)-(10), (17), (19), (20) (with risk constraints) and compared the results with the results of solving the corresponding problem (7)-(11) (without risk constraints). The optimal value of expected return for the case without risk constraints was equal to 1.01344, and for the case with risk constraints - to 1.00905. Figures 6 and 7 show the optimal trading strategies found for both cases.

7.2. Dynamic programming approach

Recall that we use the dynamic programming approach only for the model without risk constraints. We use Theorem 6.1 to reduce the size of the problem which needs to be solved in order to obtain the decision corresponding to an optimal trading strategy at $t = 1$. From Theorem 6.1 it follows that if inequality (25) is satisfied for scenario s, then $x(1,s) = 0$. We derived the following inequalities for (24):

$$\begin{aligned} 0.002281 - 0.07861 \cdot \ln(R_1) > 0, \\ 0.007186 - 0.18875 \cdot \ln(R_1) > 0, \\ 0.013043 - 0.26353 \cdot \ln(R_1) > 0, \\ 0.014768 - 0.31553 \cdot \ln(R_1) > 0. \end{aligned} \tag{26}$$

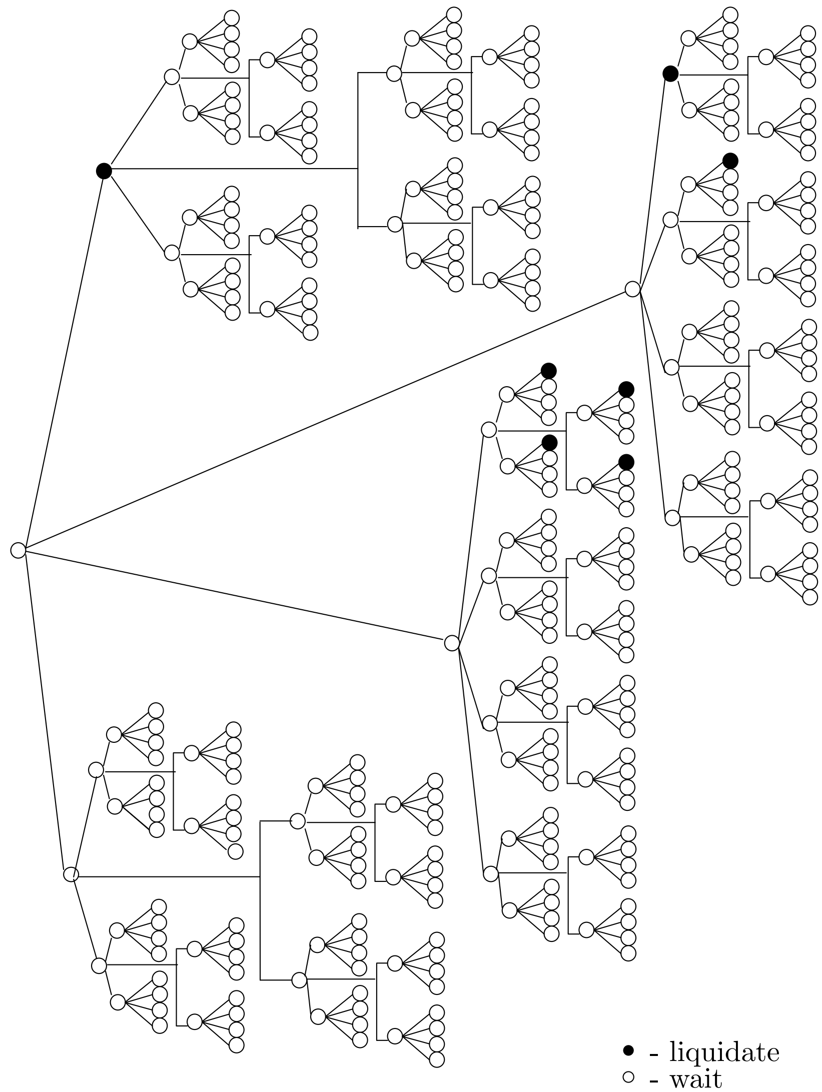

Figure 6. Solution of problem of dimension $4 \cdot 4 \cdot 4 \cdot 4 \cdot 4$ without constraint on risk.

From Theorem 6.1, we conclude that if the return at $t = 1$ is no greater than $R^* = 1.0507$ for some scenario s, then $x(1, s) = 0$ in an optimal trading strategy.

Further, we examine how much this method reduces the problem size. For this purpose we generated a set of problems of different dimensions with scenarios of equal probability (see Table II). Consider as an example the problem with 10 outcomes at each step. For this problem 10^5 scenarios were generated with equal probabilities. Possible

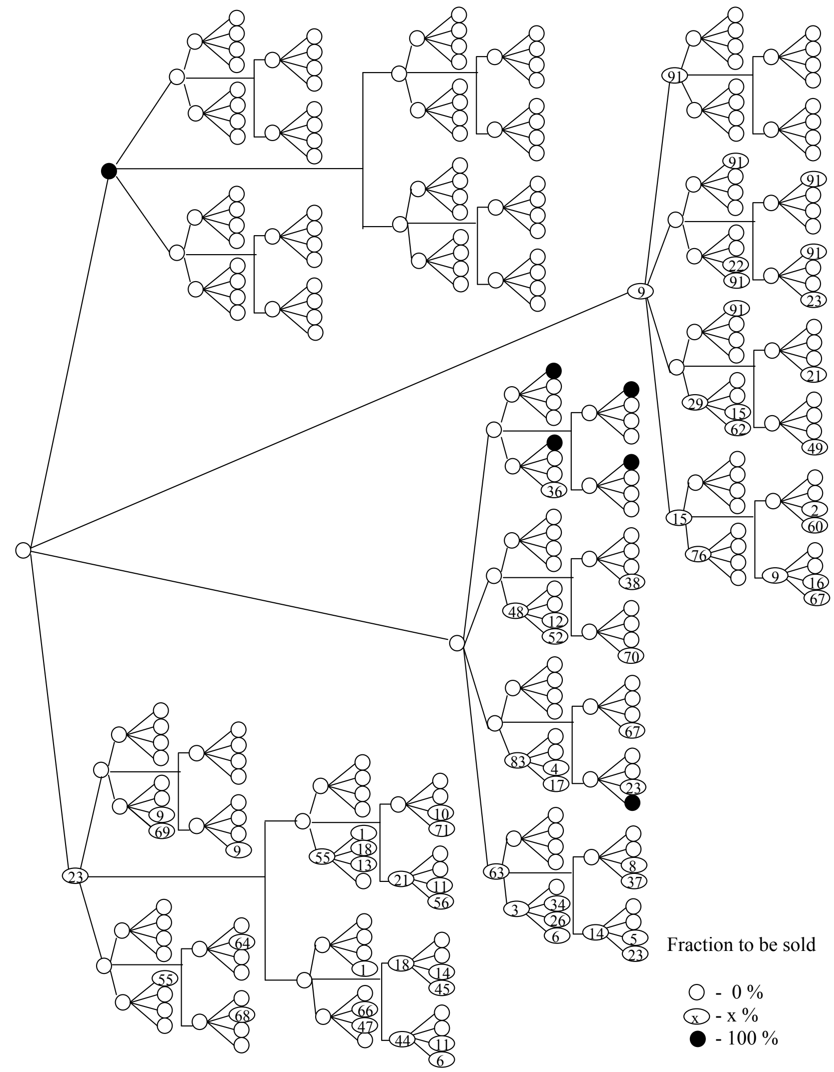

Figure 7. Solution of problem of dimension $4 \cdot 4 \cdot 4 \cdot 4 \cdot 4$ with constraint on risk.

outcomes at time $t = 1$ for this problem are 1.127, 1.047, 1.041, 1.030, 1.014, 0,999, 0,984, 0.968, 0.950, 0.914. As one can see, nine of these ten numbers are less than $R^* = 1.0507$. Therefore, if our goal is to find the optimal trading strategy for $t = 1$ only, the nine corresponding branches can be excluded from consideration and only one branch with return 1.127 at needs to be analyzed. Since according to (1) the conditional mean at time $t = T$ can be calculated using generated outcomes for the previous $T - 1$ steps, there is no need to generate outcomes for the time

Table II. Dimensions of generated problems

Dimensions of problems, $n_1 \cdot n_2 \cdot n_3 \cdot n_4 \cdot n_5$	Number of generated scenarios	Number of LP variables $5 \cdot n_1 \cdot n_2 \cdot n_3 \cdot n_4 \cdot n_5$
$4 \cdot 4 \cdot 4 \cdot 4 \cdot 4$	1024	5120
$10 \cdot 10 \cdot 10 \cdot 10 \cdot 10$	100000	50000
$14 \cdot 14 \cdot 14 \cdot 14 \cdot 14$	537824	2689120
$15 \cdot 15 \cdot 15 \cdot 15 \cdot 15$	759375	3796875
$20 \cdot 15 \cdot 15 \cdot 15 \cdot 15$	1012500	5062500
$25 \cdot 10 \cdot 10 \cdot 10 \cdot 10$	250000	1250000
$29 \cdot 15 \cdot 10 \cdot 10 \cdot 15$	652500	3262500

$t = 5$. To summarize, for this particular instance our method provides the dimension reduction from 10^5 scenarios to 10^3 scenarios at $t = 1$!

Calculations performed for other generated problems also demonstrated significant scale reduction.

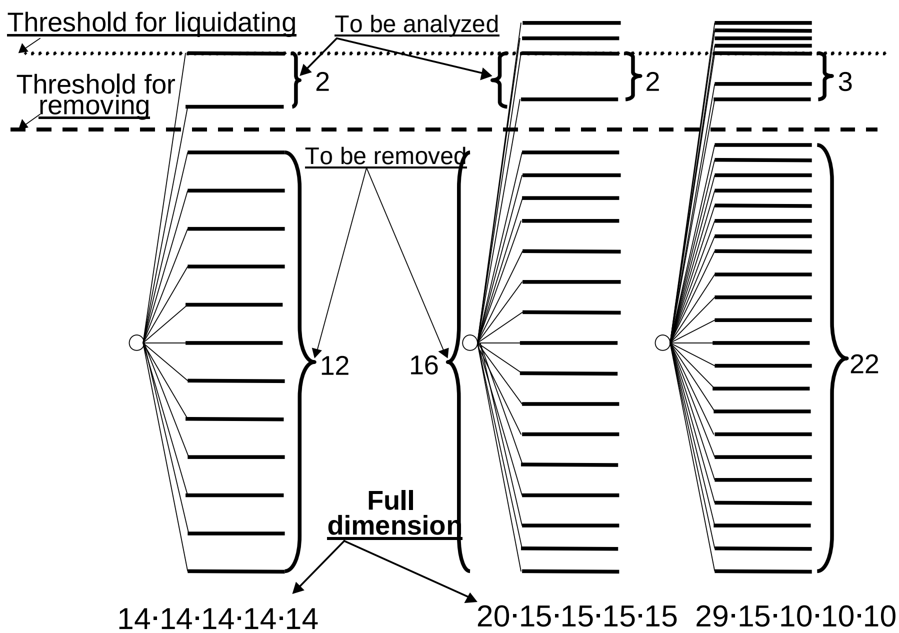

Figure 8. Removing of possible outcomes at time $t = 1$ for 3 decision trees of different dimensions

Figure 8 shows three decision trees of different dimensions: $14 \cdot 14 \cdot 14 \cdot 14 \cdot 14$, $20 \cdot 15 \cdot 15 \cdot 15 \cdot 15$ and $29 \cdot 15 \cdot 10 \cdot 10 \cdot 10$. Only outcomes for the return at time $t = 1$ are presented for each problem. There are two horizontal dashed lines. The lower line corresponds to $R*$ used as a criterion for excluding branches from consideration, and the upper line corresponds to the value of return at $t = 1$, which, if exceeded, requires the complete liquidation of the security at time $t = 1$ for optimality. As one can see, 12 outcomes of the first tree, 16 outcomes of the second tree and 22 outcomes of the third tree are below R^*. According to the proposed method, there is no liquidation under these scenarios, therefore they should be excluded from consideration at $t = 1$. Only two upper outcomes of the first tree, 4 upper outcomes of the second tree and 7 upper outcomes of the third tree should be analyzed by the dynamic programming approach to get the solution at $t = 1$.

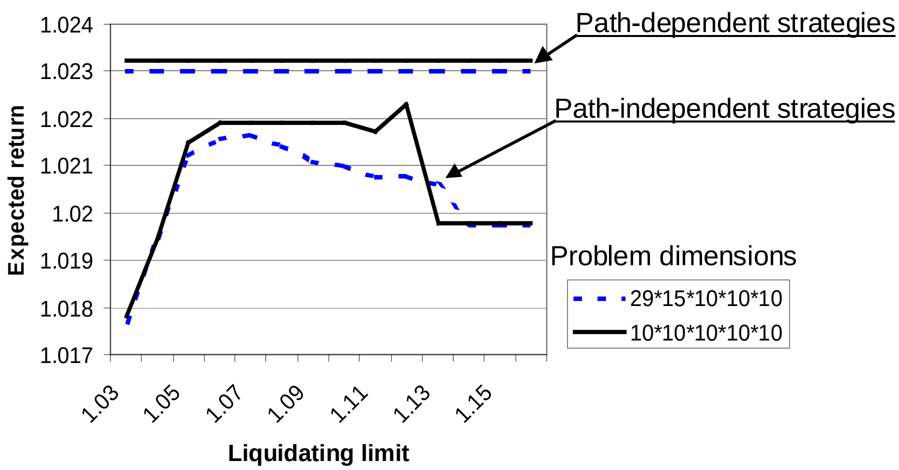

Figure 9. Comparison of different types of trading strategies

Let us now compare an optimal path-dependent strategy with simple path-independent strategies that utilize a fixed stopping limit for all steps of the decision making process. We generated two problems: one of the dimension $10 \cdot 10 \cdot 10 \cdot 10 \cdot 10$, and another of the dimension $29 \cdot 15 \cdot 10 \cdot 10 \cdot 10$. Then we applied to each of these problems a simple path-independent strategy which liquidates the security at the time moment when its return achieves or exceeds some fixed stopping limit and calculated corresponding expected return. This procedure was repeated for different values of stopping limits. The results are demonstrated in Figure 9, where expected return values are plotted as a function of the value of the stopping limit. These plots are compared with expected returns which are achieved if optimal path-dependent strategies are used in the generated problems.

The results presented in Figure 9 suggest that the path-dependent strategies outperform path-independent strategies for any fixed value of the stopping limit.

Finally, we compared the optimal value of expected return calculated for the generated model of dimension $10 \cdot 10 \cdot 10 \cdot 10 \cdot 10$ with the value of expected return calculated by simulating the optimal trading strategy on historical data. We obtained the following results:

- the expected return for historical data = 1.01349;
- the expected return for the model = 1.01298.

Therefore we can draw a conclusion that our model approximates historical data with a high level of accuracy.

8. Conclusion

We proposed several algorithms for liquidation of a financial security. The presented approaches utilize information about the security's price movements at previous time moments, which allows one to outperform simple path-independent strategies. We developed a parametric model for generating scenarios, which employs properties of multivariate log-normal distributions.

We considered cases with and without constraints on risk associated with trading. In the case without risk constraints, two solution methods were used: linear programming and dynamic programming. Although in some cases the dynamic programming approach enables us to reduce the problem size significantly by solving it step by step, this method is not applicable to the case with risk constraints.

There are still some practical issues to be addressed in the future. These include a realistic description of the permanent market impact caused by the trader's activities.

The approaches proposed in this paper can be extended to various financial applications, including optimal stopping rules and pricing of derivatives.

References

Almgren, R.: 2003, 'Optimal execution with nonlinear impact functions and trading-enhanced risk'. *Applied Mathematical Finance* **10**, 1–18.

Almgren, R. and N. Chriss: 2000, 'Optimal execution of portfolio transactions'. *Journal of Risk* **3**, 5–39.

Anderson, T. W.: 1958, *An Introduction to Multivariate Statistical Analysis*. John Willey & Sons, Inc.

Anderson, T. W.: 1971, *The Statistical Analysis of Time Series*. John Willey & Sons, Inc.

Bertsimas, D. and A. W. Lo: 1998, 'Optimal Control of Execution Costs'. *J. Financial Markets* **1**, 1–50.

Bertsimas, D., A. W. Lo, and P. Hummel: 2000, 'Optimal Control of Execution Costs for Portfolios'. *Computing in Science & Engineering* **1**, 40–53.

Birge, J. and F. Louveaux: 1997, *Introduction to Stochastic Programming*. Springer Verlag.

El-Yaniv, R., A. Fiat, R. Karp, and G. Turpin: 2001, 'Optimal Search and One-Way Trading Online Algorithms'. *Algorithmica* **30**, 101–139.

Konno, H. and A. Wijayanayake: 2002, 'Portfolio Optimization under D. C. Transaction Costs and Minimal Transaction Unit Constraints'. *J. Global Optimization* **22**, 137–154.

Rickard, J. and N. Torre: 1999, 'Optimal Transaction Implementation'. *J. Management Information Systems* **16**, 47–62.

Rockafellar, R. T. and S. Uryasev: 2000, 'Optimization of Conditional Value-at-Risk'. *Journal of Risk* **2**, 21–41.

Rockafellar, R. T. and S. Uryasev: 2002, 'Conditional Value-at-Risk for General Loss Distributions'. *Journal of Banking and Finance* **26**, 1443–1471.