Mathematica
Balkanica

New Series Vol. 20, 2006, Fasc. 1

Dynamic Stochastic Optimization in Finance

Racho Denchev [2], *Krastyu Gumnerov* [1]

Presented at Conference on "PDE Methods in Appl. Math. and Image Proc."

Most financial activities involve optimal decision making under uncertainty. This paper is a survey concerning some stochastic optimization problems in finance and the related mathematics. Some of the topics covered are: Merton problem; financial planning systems; Hamilton-Jacobi-Bellman equation; two-stage and multistage linear recourse problems; decomposition and factorization methods. [1]

The financial activity like many other activities has two characteristics:

1. The decision-making is under uncertainty, e.g., it depends on the future values of parameters, unknown at the moment of the decision-making, so they are random quantities with respect to the information at the moment.

2. The decisions are optimal with respect to some objective.

Thus a system for financial planning (portfolio management) should have two modules:

1. a module describing the random quantities of the model and their evolution (scenario generator).

2. an optimization module for given objective function and variables evolution.

[1]**Acknowledgements** : The authors do express their gratitude to Bravo Consulting Ltd. for the financial support of this research.

This review examines some methods for building the second module, having in mind to use it when stable non-Gaussian processes give the variable evolution. It is based on the following sources: [4], [7], [14], [23], [24], [30], [40].

1. An optimal portfolio selection problem ([30]).

There are given the assets p_1 and p_2 with price processes, satisfying the equations

$$\frac{dp_1}{p_1} = adt + \alpha dB_t, \quad b < a, \tag{1}$$

$$\frac{dp_2}{p_2} = bdt, \quad b < a, \tag{2}$$

At the moment t let X_t be the investor's wealth. He divides it into two parts: $u_t X_t$ and $(1-u_t)\,X_t$, $0 \leq u_t < 1$. With the part $u_t X_t$ he buys the assets p_1, and with the part $(1-u_t)\,X_t$ - the assets p_2. In this way he composes the portfolio, which contains an amount $\frac{u_t X_t}{p_1(t)}$ of the asset p_1 and an amount $\frac{(1-u_t)X_t}{p_2(t)}$ of the asset p_2. The portfolio price increase dX_t will be

$$\begin{aligned} dX_t &= \frac{u_t X_t}{p_1} dp_1 + \frac{(1-u_t)X_t}{p_2} dp_2 = \frac{u_t X_t}{p_1}\left(p_1 adt + p_1 \alpha dB_t\right) + \frac{(1-u_t)X_t}{p_2} p_2 bdt \\ &= X_t[u_t a + (1-u_t)b]dt + X_t u_t \alpha dB_t. \end{aligned}$$

At the initial moment $s < T$ the investor's wealth X_s is determined $= X$. Let the performance N be an increasing concave function of the wealth X_t, for example $N\,(X_t)\,N\,(X_t) = X_t^r$, $0 < r < 1$.

The investor has an investment horizon T and he wants trading without loan to maximize the performance at the moment T, more exactly to maximize the quantity

$$J^u\,(s,x) = E^{s,x}\left[N\left(X_\tau^n\right)\right]$$

where $E^{s,x}$ is the expectation with respect to the probability law (the conditional probability density) of the process, which at the moment s has the value x; τ is the first exit time from the domain $G = \{(t,x) : t < T, x > 0\}$.

The problem is to find a function $\Phi\,(s,x)$ and a stochastic process u_t^*, $0 \leq u_t^* < 1$, to satisfy the conditions

$$\Phi\,(s,x) = \sup\left\{J^u\,(s,x) \qquad u \text{ is a stocahstic process, } 0 \leq u < 1\right\},$$

and $\Phi(s,x) = J^{u^*}(s,x)$. To solve this problem we compose the Hamilton-Jacobi-Bellman (HJB) equation (see the appendix for stochastic control). The infinitesimal generator of the process X_t is D^v,

$$(D^v f)(t,x) = \frac{\partial f}{\partial t} + x(av + b(1-v))\frac{\partial f}{\partial x} + \frac{1}{2}\alpha^2 v^2 x^2 \frac{\partial^2 f}{\partial x^2}.$$

The HJB equation is

$$\left\{ \ \sup_v \{(D^v\Phi)(t,x)\} = 0 \quad \text{for} \quad (t,x) \in G, \right. \tag{3}$$

From this equation for each (t,x) we find $\nu = u(t,x)$, so that the function

$$\eta(v) = D^v\Phi = \frac{\partial \Phi}{\partial t} + x(b + (a-b)v)\frac{\partial \Phi}{\partial x} + \frac{1}{2}\alpha^2 v^2 x^2 \frac{\partial^2 \Phi}{\partial x^2} \tag{4}$$

has a maximum. The function $\eta(v)$ is a polynomial of second order, hence if $\frac{\partial \Phi}{\partial x} > 0$ and $\frac{\partial^2 \Phi}{\partial x^2} < 0$, it gets a maximum for

$$v = u(t,x) = -\frac{(a-b)\frac{\partial \Phi}{\partial x}}{x\alpha^2 \frac{\partial^2 \Phi}{\partial x^2}}. \tag{5}$$

Substituting (5) in (4) and (3) we obtain the following nonlinear boundary problem for $\Phi(t,x)$:

$$\frac{\partial \Phi}{\partial t} + bx\frac{\partial \Phi}{\partial x} - \frac{(a-b)\left(\frac{\partial \Phi}{\partial x}\right)^2}{2\alpha^2 \frac{\partial^2 \Phi}{\partial x^2}} = 0 \tag{6}$$

and

$$\Phi(t,x)\,|_{t=T} = N(x), \quad \Phi(t,x)\,|_{x=0} = N(0),$$
$$(.. \quad \Phi(t,x)\,|_{\Gamma} = N(x), \quad \Gamma = (T,x) \cup (t,0)).$$

Let $N(x) = x^r$, $0 < r < 1$ and let look for $\Phi(t,x)$ in the form

$$\Phi(t,x) = f(t, x^r).$$

By substituting in (6) we get

$$\Phi(t,x) = e^{\lambda(t-x)}x^r, \tag{7}$$
$$\lambda = br + \frac{(a-b)^2 r}{2\alpha^2(1-r)}.$$

Now from (5) and (7) we obtain

$$u*(t,x) = \frac{a-b}{\alpha^2(1-r)} \tag{8}$$

If

$$0 < \frac{a-b}{\alpha^2(1-r)} < 1$$

(8) is the solution of the problem (u^* is in fact a constant).

This result means that in the portfolio practical management, the investor invests his capital at the initial moment in the proportion

$$\frac{a-b}{\alpha^2(1-r)}, \quad 1 - \frac{a-b}{\alpha^2(1-r)}$$

and he does not change it up to horizon T.

If $u^*(t,x)$ depends on t and x the investor rebalances the portfolio at each moment t (practically in discrete moments), and at the moment $t+dt$ he observes at the market the increases dp_1 and dp_2 of both assets prices, calculates the increase

$$dX_t = X_t\left[u(t)\frac{dp_1}{p_1(t)} + (1-u(t))\frac{dp_2}{p_2(t)}\right]$$

of his wealth and composes a portfolio of both assets in proportion u^*_{t+dt}, $1 - u^*_{t+dt}$, where

$$u^*_{t+dt} = u^*(t+dt, X_t + dX_t) = u^*(t+dt, X_{t+dt}).$$

2. The Merton problem ([14], [23]).

There are traded at the market $N+1$ assets p^0, p^1, $\blacksquare$, p^N, with the following price processes:

$$\frac{dp^0_t}{p^0_t} = rdt \tag{9}$$

The investor has a horizon T and at an arbitrary moment $t \in [0,T]$ he possesses a portfolio of the assets p^0, p^1, $\blacksquare$, p, in quantities respectively θ^0_t,

θ_t^1, ..., θ_t^N, doing a consumption at rate $C_t \geq 0$, where θ_t^0, θ_t^1, ..., θ_t^N, C_t are stochastic processes. At the initial moment his wealth is $X_0 \geq 0$, and he trades and consumes without exterior incomes. That means that his wealth X_t at the moment t is

$$X_t = \sum_{i=1}^{N} \theta_t^i p_t^i = X_0 + \sum_{i=1}^{N} \int_0^t \theta_\tau^i dp_\tau^i - \int_0^t c_\tau d\tau \geq 0, \quad t \in [0, \tau].$$

It follows that

$$dX_t = \sum_{i=0}^{N} \theta_t^i dp_t^i - c_t dt. \tag{10}$$

This is a "budget equation".

It is convenient instead of processes

$$\theta_t^i$$

to introduce the processes

$$\alpha_t^i = \frac{\theta_t^i p_t^i}{X_t} = \frac{\theta_t^i p_t^i}{\sum\limits_{j=0}^{N} \theta_t^j p_t^j}, \quad i = 0, ..., N, \quad \sum_i \alpha^i = 1,$$

which represent the part of the wealth included in p_t^i, i.e. the proportion in which the wealth is distributed in the assets.

Substituting (9) in (10) the budget equation gets the form of

$$\begin{aligned} dX &= \sum_{i=1}^{N} \theta^i \left(\mu_i p^i dt + \sum_{j=1}^{N} \sigma_{ij} p^i dB^j \right) + \theta^0 r p^0 dt - c dt = \\ &= \sum_{i=1}^{N} \mu_i \alpha^i X dt + \sum_{i,j=1}^{N} \sigma_{ij} \alpha^i X dB^j + r\alpha^o X dt - c dt. \end{aligned}$$

i.e.

$$dX = \left(X \sum_{i=1}^{N} \alpha^i \mu_i + X\alpha^0 r - c \right) dt + X \sum_{j=1}^{N} \left(\sum_{i=1}^{N} \alpha^i \sigma_{ij} \right) dB^t. \tag{11}$$

The investor's "performance" is defined by the consumption done in the period $[0, \tau]$ and by the wealth possessed at the moment τ. More exactly it is given by the formula

$$J^{(c,\alpha)}(t,x) = E^{t,x}\left[\int_t^T L(\tau, X_\tau; c\tau, \alpha_\tau)d\tau + K(X_\tau)\right],$$

where $E^{t,x}$ is the expectation with respect to the probability law of the process X, which begins at the moment t with the value x.

The investor's goal is by trading and consuming, to maximize $J^{(c,a)}(t,x)$. Let

$$\Phi(t,x) =^{def} \sup_{(c,\alpha)} J^{(c,\alpha)}(t,x) = J^{(c*,\alpha*)}(t,x),$$

where c^*, α^* are the optimal consumption and investment strategy. The function $\Phi(t,x)$ satisfies the HJB equation and the boundary condition

$$\Phi(t,x_\tau) = K(x_\tau).$$

Let us compose the equation HJB. The infinitesimal generator of the process according to (11) is

$$D^{(\tau,\alpha)} = \frac{\partial}{\partial t} + \left(x\sum_{i=1}^{N}\alpha^i\mu_i + x\alpha^0 r - c\right)\frac{\partial}{\partial x} + \frac{1}{2}x^2\sum_{j=1}^{N}\left(\sum_{i=1}^{N}\alpha^i\sigma_{ij}\right)^2\frac{\partial^2}{\partial x^2}.$$

The HJB equation is

$$\sup_{(c,\alpha)}\left\{D^{(c,\alpha)}\Phi(t,x) + L(t,x)\right\} = 0,$$

$$\Phi(t,x) = K(x).$$

To solve the problem we follow the procedure: having arbitrary fixed t, x, we calculate the supremum of the function

$$\eta(c,\alpha) = \frac{\partial\Phi(t,x)}{\partial t} + \left(x\sum_{i=1}^{N}\alpha^i\mu_i + x\alpha^0 r - c\right)\frac{\partial\Phi}{\partial x}(t,x)$$
$$+\frac{1}{2}x^2\sum_{j=1}^{N}\left(\sum_{i=1}^{N}\alpha^i\sigma_{ij}\right)^2\frac{\partial^2\Phi}{\partial x^2}(t,x) + L(t,x;c,\alpha).$$

Setting equal to zero the derivatives with respect to c and α of the function $\eta(c,\alpha)$, we obtain a linear system of equations, from which we obtain c and a as functions of

$$\frac{\partial\Phi}{\partial x},\ \frac{\partial^2\Phi}{\partial x^2},$$

i.e.

$$c = c(\Phi_x, \Phi_{xx}),$$
$$\alpha^i = \alpha^i(\Phi_x, \Phi_{xx})$$

Substituting in $\eta\,(c, \alpha)$ we obtain the equation

$$\eta\,(c\,(\Phi_x, \Phi_{xx})\,, \alpha\,(\Phi_x, \Phi_{xx})) = 0 \tag{12}$$

which is a nonlinear partial differential equation.

In the special case when $L = 0$ and $K\,(x) = x^s$, $0 < s < 1$ the problem can be solved explicitly. We look for a solution of (12) in the form of $\widehat{O}\,(t, x) = f\,(t)\,x^s$ and obtain ordinary differential equation of first order for $f\,(t)$.

3. A Model of Financial Planning (Assets/Liabilities Management)

Recent papers analyze the effects of asset return predictability on asset allocation decision of long-term investors. These papers investigate how the investor's horizon or the uncertainty of the estimated parameters affects the allocation decision. There has been a growing interest in the development of multiperiod stochastic models for asset and liability management (ALM). Kusy and Ziemba in [20] developed a multistage stochastic linear programming model for Vancouver City Savings Credit Union. Another successful application of multistage stochastic programming is the Russell-Yasuda Kasai model in [9]. The investment strategy suggested by the model resulted in extra income of $79 million during the first two years of its application (1991 and 1992). For other examples see in [24], [34] and [35].

Here we consider a model taken from [24]. Assume the investor, having a determined goal at the horizon T, invests in given assets by rebalancing his portfolio until the moment T according to the received information and his expectations for the prices at the market. The moments, in which the portfolio is rebalanced are numbered with $\{0, 1, 2, ..., \tau, \tau + 1, ..., T\}$. Asset investment categories are defined by the set $A = \{1, 2, ..., I\}$, with category 1 representing cash. The possible scenarios are numbered with $\{1, 2, ..., s, ..., S\}$. Two different scenarios can coincide up to a moment. Hence the decisions based on these scenarios must coincide till that moment. This scenarios characteristic is called *nonanticipativity*.

There are two types of variables, which determine the model. The variables, which don't depend on the investor, are called parameters, the others depending on the investor are called decision variables.

“Parameters”:

$r_{i,t}^{s} = 1 + \rho_{i,t}^{s}$where $\rho_{i,t}^{s}$ is the return for asset i, in moment t, under scenario s;

π_s Probability that scenario s occurs, $\sum\limits_{s=1}^{S} \pi_s = 1$;

w_0 Wealth in moment 0;

$\sigma_{i,t}$ Transaction costs incurred in rebalancing asset i at the beginning of time period t (cost of selling equals cost of buying);

β_t^s Borrowing rate in period t under scenario s;

“Decision variables ”:

$x_{i,t}^{s}$ Amount of money for asset category i, in time period t, under scenario s, after rebalancing;

$v_{i,t}^{s}$ Amount of money for asset category i, in the beginning of time period t, under scenario s, before rebalancing;

w_t^s Wealth at the beginning of time period t, under scenario s;

$p_{i,t}^{s}$ Amount of asset i, purchased for rebalancing in period t, under scenario s;

$d_{i,t}^{s}$ Amount of asset i, sold for rebalancing in period t, under scenario s;

b_t^s Amount borrowed in period t, under scenario s.

The Model:

$$\max (z) = \sum_{s=1}^{S} \pi_s f(w_\tau^s), \tag{13}$$

s.t.

$$\sum_i x_{i,0}^{s} = w_0, \quad \forall s \in S, \tag{14}$$

$$\sum_i x_{i,0}^{s} = w_\tau^s, \quad \forall s \in S, \tag{15}$$

$$v_{i,t}^{s} = r_{i,t-1}^{s} x_{i,t-1}^{s},\ \forall s \in S,\ t = 1, ..., \tau,\ i \in A, \tag{16}$$

$$x_{i,t}^{s} = v_{i,t}^{s} + p_{i,t}^{s}(1 - \sigma_{i,t}) - d_{i,t}^{s},\ \forall s \in S,\ i \neq 1,\ t = 1, ..., \tau, \tag{17}$$

(18)
$$x^s_{1,t} = v^s_{1,t} + \sum_{i\neq 1} d^s_{i,t}(1-\sigma_{i,t}) - \sum_{i\neq 1} p^s_{i,t} - b^s_{t-1}(1-\beta^s_{t-1}) + b^s_t,\ \forall s \in S,\ t = 1, ..., \tau,$$

(19)
$$x^s_{i,t} = x^{s'}_{i,t}$$

for all scenarios s and s' with identical past up to time t.

The model works as follows:

Assume we have solved the optimization problem, e.g. according to the given scenarios for the parameters (from the scenario generator), we have determined the optimal "decision variables" (see the appendix for stochastic programming).

Assume in the moment $t-1$ the scenario s' is realized and after rebalancing the investor invests (places funds) in the assets i the sum of

$$x^{s'}_{i,t-1}.$$

Assume in the following moment t the scenario s is realized. It is identical with s' till the moment $t-1$ and because of that (taking into consideration the nonanticipativiy)

$$x^s_{i,t-1} = x^{s'}_{i,t-1}$$

Then we determine $v^s_{i,t}$ using (16), $v^s_{i,t} = r^s_{i,t-1}x^s_{i,t-1} = r^s_{i,t-1}x^{s'}_{i,t-1}$ and from (17) and (18) we determine $x^s_{i,t}$.

In conclusion we wood like to mention that investors are richly rewarded for identifying superior investment strategies as compared to their benchmark portfolios. High frequency data are readily available on a global basis. Powerful computers are also easily found to conduct the optimization search. The obstacles for applying stochastic optimization models are quickly receding.

Still the resulting financial optimization models are large, difficult to understand and evaluate with respect to performance, and can be expensive to build and operate. Mathematical programmers will be able to overcome at least some of these barriers.

Appendix

Mathematics for Stochastic Optimization

To understand the stochastic optimization it is useful to keep in mind the analogy with some more simple problems:

1) The elementary problem of finding the conditional extremum of a function;

2) The problems of the calculus of variations and the variational approach in the classical mechanics and mathematical physics;

3) The problems of the deterministic optimal control.

In general, the solution of these problems, including the stochastic control problems, is reduced to some optimality conditions, in the form of equations (more often). These are the equations of the considered system: they describe the evolution of the parameters defining the system. Characteristic examples are: the equation for the stationary points of a function of numerical variables, the Kun-Taker equations, the Euler-Lagrange equations in the calculus of variations, the equations of the mathematical physics, the Hamilton-Jacobi equation, the Hamilton-Jacobi-Bellman equation.

The advantage of this approach is that it allows to see sometimes the patterns of the system in consideration, it permits sometimes a qualitative investigation of the solutions of the problem. The stochastic programming actually gives us methods for numerical solution of these equations, but in principle other methods are also possible and in some special cases it is possible to obtain solutions in explicit form. An example is the Merton problem.

I. Stochastic Control. The Hamilton-Jacobi-Bellman equation

In the following we give a brief presentation of a basic theorem of the stochastic control. We closely follow [30].

Assume the system in question is described in a probability space (Ω, F_t, P) by the Ito process of the form

$$dX_t = dX_t^u = b\,(t, X_t, u_t)\,dt + \sigma\,(t, X_t, u_t)\,dB_t \tag{20}$$

where $X_t \in \mathbb{R}$, $b : \mathbb{R} \times \mathbb{R}^n \times U \times \mathbb{R}^n$, $\sigma : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}^{n\times m}$, B_t is the $m-$dimensional Brownian motion and $u_t \in U \subset \mathbb{R}^k$ is a parameter, whose values in a given set U we can choose at each moment t to control the system. Thus

$$u_t = u\,(t, \omega)$$

is also a stochastic process, $F_t^{(m)}$–adapted. It will be called "control".

Let $\left\{X_h^{s,x}\right\}_{h\geq s}$ be the solution of (20), such that $X_{t|t=s} = x$, i.e.,

$$X_h^{s,x} = x + \int_s^h b(r, X_r^{s,x}, u_r)dr + \int_s^h \sigma(r, X_r^{s,x}, u_r)dB_r, \quad h \geq s.$$

Let $L : \mathbb{R} \times \mathbb{R}^n \times U \to \mathbb{R}$ and $K : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be given continuous functions, $G \subset \mathbb{R} \times \mathbb{R}^n$ be a domain and let

$$\widehat{T}$$

be the first exit time after s from G for the process $\{X_r^{s,x}\}_{r\geq s}$, i.e.

$$\vec{T} = \vec{T}^{s,x}(\omega) = \inf\{r : r > s, (r, X_r^{s,x}(\omega) \notin G\}.$$

We define the quantity "performance"

$$J^u(s,x) = E^{s,x}\left[\int_s^{\vec{T}} L(r, X_r, u_r)dr + K(\vec{T}, X_{\vec{T}})\chi_{\vec{T}<\infty}\right] \tag{21}$$

To simplify the notations we introduce

$$Y_t = \left(s+t, X_{s+t}^{s,x}\right)$$

for

$$t \geq 0, \qquad Y_0 = (s,x) = y, \qquad T = \widehat{T} - s.$$

Then (20) becomes

$$dY_t = dY_t^u = b(Y_t, u_t)dt + \sigma(Y_t, u_t)dB_t.$$

The problem is for each $y \in G$ to find a control $u^* = u^*(y,t,\omega)$, so that

$$J^{u^*}(y) = \sup J^u(y);$$

$J^{u^*}(y)$ is called "optimal performance" and u^*– "optimal control".

The optimality condition is given by some equations, formulated in the theorem that follows.

For $v \in U$ and $g \in C_0^2(\mathbb{R} \times \mathbb{R}^n)$ we define

$$(D^v g)(y) = \frac{\partial g}{\partial s}(y) + \sum_{i=1}^n b_i(y,v)\frac{\partial g}{\partial x_j}(y) + \sum_{i,j=1}^n a_{ij}(y,v)\frac{\partial^2 g}{\partial x_i \partial x_j}$$

where

$$a_{ij} = \frac{1}{2}(\sigma\sigma^r)_{ij}.$$

For each choice of the function $u : \mathbb{R} \times \mathbb{R}^n \times U$ the operator A, defined by

$$(Ag)(y) = \left(D^{u(y)}g\right)(g) \qquad g \in C_0^2(\mathbb{R} \times \mathbb{R}^n)$$

is an infinitesimal generator of a process Y_t, which is a solution of the equation

$$dY_t = b(Y_t, u(Y_t))dt + \sigma(Y_t, u(Y_t))dB_t.$$

Theorem. *Let the function* $\Phi(y) = \sup J^u(y)$ *be bounded and belong to* $C^2(G) \cap C(\overline{G})$, $u(t,\omega)$. *Let*

$T^\wedge < \infty$ *a.s. for each* $y \in G$ *and let exist an optimal control* u^*. *Then*

(22)

$$\sup\{L(y,v) + D^v\Phi(y)\} = 0 \qquad y \in G, v \in U$$

and

$\Phi(y) = K(y)$, $y \in \partial G$.

The supremum in (22) is obtained when $v = u^*(y)$, *where* u^* *is an optimal control, i.e.*

$$L(y, u^*(y)) + D^{u^*(y)}\Phi(y) = 0 \qquad y \in G.$$

The equation (22) is called Hamilton-Jacobi-Bellman equation (HJB).

II. Two-Stage Stochastic Linear Programs with Fixed Recourse (2S-SLPR)

The stochastic programming is practically the most important approach for optimal decision-making under uncertainty.

Stochastic programming, as the name implies, is mathematical (i.e. linear, integer, mixed-integer, nonlinear) programming but with a stochastic element present in the data. By this we mean that:

- in deterministic mathematical programming the data (coefficients) are known numbers;

- in stochastic programming these numbers are unknown, instead we may have a probability distribution present.

Stochastic programming therefore deals with situations where we have uncertainty present.

In the common case the stochastic programs are generalizations of the deterministic optimization programs, in which some uncontrollable data are unknown surely. Their typical features are many decision variables with many potential values, discrete time periods for the decisions, the use of expectation functionals for the goals and known (or partially known) distributions.

In the simplest model of this type we have two stages:

- in the first stage we make a decision;

- in the second stage we see realizations of the stochastic elements of the problem but are allowed to make further decisions to avoid the constraints of the problem becoming infeasible.

In other words, in the second stage we have *recourse* to a further degree of flexibility to preserve feasibility (but at a cost). Note particularly here that in this second stage the decisions that we make will be dependent upon the particular realization of the stochastic elements observed.

II.1. Formulation ([7]) The two-stage stochastic linear program with fixed recourse takes the following form:

$$\min z = c^T x + E_\xi \left[\min q(\omega)^T y(\omega) \right] \tag{23}$$

s. t. $Ax = b$, $T(\omega)x + Wy(\omega) = h(\omega)$, $x \geq 0$, $y(\omega) \geq 0$, where:
c is a known vector in $\mathbb{R}^{n_1}$,
b is a known vector in $\mathbb{R}^{m_1}$,
A and W are known matrices of size $m_1 \times n_1$, $m_2 \times n_2$.
W is called the *recourse matrix*, which we assume here is fixed.

For each ω , $O(\omega)$ is $m_2 \times n_1$ matrix and $q(\omega) \in \mathbb{R}^{n_2}$, $h(\omega) \in \mathbb{R}^{m_2}$. Piecing together the stochastic components of the problem, we obtain a vector:

$\xi^T(\omega) = (q(\omega)^T, h(\omega)^T, T_1(\omega), ..., T_{m_2}(\omega))$, with $N = n_2 + m_2 + m_2 \times n_1$ components, where $T_i(\omega)$ is the i-th row of $O(\omega)$.

E_ξ represents the mathematical expectation with respect to ξ.

A distinction is made between the first stage and the second stage. The first- stage decisions are represented by vector x. Corresponding to x are the first-stage vectors and matrices c, b and A. In the second stage, a number of random events

$\omega \in \Omega$ (set of all random events) may realize. For a given realization ω , the second-stage problem data $q(\omega)$, $h(\omega)$ and $O(\omega)$ become known. Each component of q, T, and h is thus a possible random variable.

Let also $\Xi \subseteq R^n$ be the support of ξ , i.e. the smallest closed subset in $\mathbb{R}^n$, such that

$$\begin{aligned} \min z &= c^T x + E_\xi \left[\min q(\omega)^T y(\omega) \right], \\ P(\xi &\in \Xi) = 1. \end{aligned}$$

As just said, when the random event ω is realized, the second-stage problem data $q(\omega)$,$h(\omega)$ and $O(\omega)$ become known. Then the second-stage decisions $y(\omega, x)$ must be taken. The dependence of y on ω is of a completely different nature from the dependence of q or other parameters on ω . It is not functional but simply indicates that the decisions y are typically not the same under different realizations of ω .

The objective function of (23) contains a deterministic term $c^T x$ and the expectation of the second-stage objective $q(\omega)^T y(\omega)$ taken over all realizations of the random event ω . This second-stage term is the more difficult one because, for each ω , the value $y(\omega)$ is the solution of a linear program.

Using discrete distributions, the resulting model can be reformulated as a linear programming problem called deterministic equivalent program (DEP).

Problem (23) is equivalent to the DEP:

$$\min z = c^T x + \vec{Q}(x) \tag{24}$$

$s.t.$

$$Ax = b \qquad x \geq 0 \tag{25}$$

where:

$$\vec{Q}(x) \;=\; E_{\xi} Q(x, \xi(\omega)) \tag{26}$$

and for a given realization ω ,

$$Q(x, \xi(\omega)) = \min_{y} q(\omega)^T y / Wy = h(\omega) - T(\omega)x, y \geq 0$$

is the second-stage value function.

The expected value of the function of the second stage is given in (26). Because of its importance for the application, let's go into the situation, when ξ is finite discrete random variable, and precisely $\xi \in \Xi$, with Ξ a finite or countable set. The second-stage value function is then the weighted sum of the $Q\left(x, \xi\right)$ values for the various possible realizations of ξ. If the stochastic elements have continuous distributions then the mathematical problems associated with formulation and solution become formidable.

II.2. Approaches for Solving 2S-SLPR The current researches can be divided in three basic types: decomposition, direct and lagrangian. The decomposition methods, being the most computability efficient, we will present better description, and we shall only mention the concepts on which are built the direct methods.

Note here that what made our formulation of the problem simple was the fact that the stochastic element had a discrete distribution.

II.2.1. Decomposition Methods ([7]) The stochastic programs with recourse calculations focus in a two-stage problem with finite number implementations and similarly in multi-stage problems with finite number implementations in the stages. The two stages problem is also the primal model for calculations. The general model consists in the choice of some initial decision to minimize the current expenses adding the expected value of further recourse actions.

The basic idea is to approximate the recursive function $\vec{Q}(x)$ in the objective of these problems. $\vec{Q}(x)$ is of polyhedral structure. Cuts represent successive linear approximations of $\vec{Q}(x)$. Because of polyhedral structure of $\vec{Q}(x)$ this is a convergence process to an optimal solution of a finite number steps. Let $k =$

1, ■ , K index the possible realizations and let p_k be their probabilities. Under this assumption, we may now write the deterministic equivalent program in the extensive form. This form is created by associating one set of second-stage decisions, say, y_k to each realization ξ, i.e. to each realization of q_k, h_k and T_k.

It is a large-scale linear problem that we can define as the *extensive form (EF)*:

$$\min\ c^T x + \sum_{k=1}^{K} p_k q_k^T y_k \qquad \text{s.t.} \quad Ax = b, \tag{27}$$
$$T_k x + W y_k = h_k, \quad k = 1, \cdots ..., \ K;$$
$$x \geq 0, \quad y_k \geq 0, \quad k = 1, \ ..., \ K.$$

$L-$shaped algorithm is a representative of Benders Decomposition (or outer linearization). It solve the linear program (28) – (30),

$$\min z = c^T x + \theta \qquad \text{s.t.} \quad Ax = b \tag{28}$$

$$D_l x \geq d_l, \qquad l = 1, 2, ..., r, \tag{29}$$

$$E_l x + \theta \geq e_l, \qquad l = 1, 2, ..., s, \tag{30}$$

with

$$x \geq 0, \qquad \theta \in \mathbb{R}.$$

The method consists in solving an approximation of (24) by using an outer linearization of $\vec{Q}(x)$. Two types of constraints are sequentially added:

(i) feasibility cuts (29) determining $\left\{x : \vec{Q}(x) < +\infty\right\}$; and

(ii) optimality cuts (30), which are linear approximations to $\vec{Q}(x)$.

It is constructively proved that constraints of the type (30) are supporting hyperplanes of $\vec{Q}(x)$ and that the algorithm will converge to an optimal solution provided the constraints (29) adequately define feasible points of K_2 .

For full description and modifications of L-shaped method see [7]. An alternative of the outer linearization or the feasibility cut approach is a column generation approach, known as a Dantzig-Wolfe decomposition or inner linearization (for other basic approaches to the big problems, see [16]).

In fact it is produced via L-shaped method, getting the dual problem, and both approaches are completely mutually dual.

II.2.2. Direct Methods ([4]) These are methods, using the standard optimization procedures, but without high-dimensional strategies as decomposition. They are concentrated on the stochastic programs specific structure. The computational methods in this domain are common for the linear models, so that we focus on the basic extreme point method (simplex) and on the interior point methods.

II.2.2.1. Extreme Point Methods At these methods we first treat the two-stage model with a discrete distributed support on

$$\Xi = \{\xi_1, ..., \xi_N\},$$

with associate probabilities $p_1, ..., p_N$. In the case of these hypothesis 2S-SLPR becomes

$$\begin{array}{ll} \min_{x,y} c^T x + p_1 q_1^T y_1 + \cdots + p_N q_N^T y_N \\ \text{s.t.} \quad Ax = b\,. \end{array} \tag{31}$$

The extreme point methods treat the matrix of (31) using a specific factorization (see [19] and [33]). This complete factorization is composed by separated submatrix factorizations corresponding to the diagonal blocks (with W components) and one single block.

In the multi-stage programs there are also possible basis factorization techniques extensions [2], but in this case it seems that the additional (at each new stage) specific structure outweighs every memory computational economies. But the results can be quite different for the interior point methods. The key can be the level up to which each method can be implemented in parallel.

For more detailed description of these methods see [7].

II.2.2.2. Interior Point Methods The factorization schemes are also very hopeful in the interior point methods where there are many speculations that their solutions efforts increase linearly with the problem dimension. For example, a discussion about the computational results appears in [21], [22]. The problems of the type (31) can be also solved via interior points methods. If we denote the constraints matrix in (31) with

$$\overline{A},$$

the effort for solving in most of the interior point methods is dominated by the calculations made with one matrix,

$$M = \overline{A} D^2 \overline{A}^T,$$

where D is a diagonal matrix, such as when $D_0 = diag\,(x^\nu)$, $D_N = diag\,(y_N^\nu)$, $k = 1, 2, ..., N$, we will obtain

$$M = \begin{pmatrix} AD_0^2A^T & AD_0^2T_1^T & \cdots & AD_0^2T_N^T \\ & & & \\ & & & \end{pmatrix}. \tag{32}$$

This one is a matrix much more dense than the original constraints matrix in (31). In this case a direct application of the interior point method that solves the systems with M is quite inefficient.

The factorization method is used for maintaining an invertible matrix, when M is singular (see [8]).

There are some other options for interior point methods. In fact all the results report improvements for large problems over direct simplex-based techniques. Decomposition methods, however, have often achieved even greater efficiencies.

III. Multistage Stochastic Linear Programs with Recourse (MS-SLPR)

III.1. Formulation ([7]) The multistage stochastic linear program with fixed recourse takes the following form:

$$\begin{gathered} \min z = c^1x^1 + E_{\xi^2}[\min c^2(\omega)x^2(\omega^2) + \cdots + E_{\xi^H}[\min c^H(\omega)x^H(\omega^H)]\cdots] \\ \text{s.t. } W^1x^1 = h^1, \\ T^1(\omega)x^1 + W^2x^2(\omega^2) = h^2(\omega), \\ \cdots \vdots \\ T^{H-1}(\omega)x^{H-1}(\omega^{H-1}) + W^Hx^H(\omega^H) = h^H(\omega), \\ x^1 \geq 0;\ x^t(\omega^t) \geq 0,\ t = 2,\ \ldots,\ H; \end{gathered} \tag{33}$$

where: c^1 is a known vector in $\mathbb{R}^{n_1}$, h^1 a known vector in $\mathbb{R}^{m_1}$, $\xi^t(\omega)^T = (c^t(\omega)^T, h^t(\omega)^T, T_1^{t-1}(\omega), ..., T_{m_t}^{t-1})$ is a random N_t -vector defined on $\left(\Omega, \Sigma^t, P\right)$, (where $\Sigma^t \subset \Sigma^{t+1}$ for all $t = 2, \blacksquare\,, H$) and each W^t is a known $m_t \times n_t$ matrix. The decisions x depend on the history up to time t, which we indicate by ω^t . We also suppose that Ξ^t is the support of ξ^t.

A multistage stochastic program with recourse is a multi-period mathematical program where parameters are assumed to be uncertain along the time path. The term recourse means that the decision variables adapt to the different outcomes of random parameters at each time period. Different formulations of MS-SLPR are proposed in the literature, for example see [5] and [12].

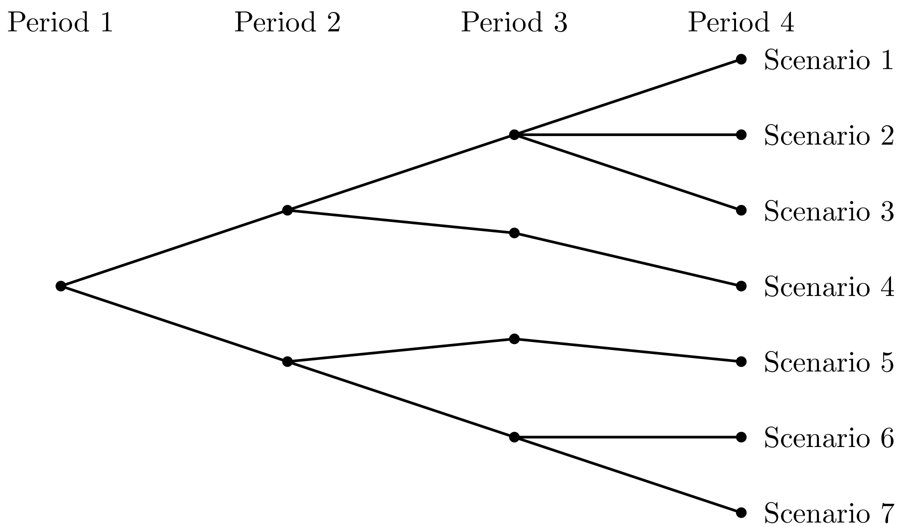

Figure 1: A tree of seven scenarios over four periods.

Using discrete distributions, the resulting model on every stage can be reformulated as a linear programming problem also called deterministic equivalent. In this way it is possible to present the realization of the random vector ξ^t by the so called scenarios tree. As an illustration, we present in Figure 1 an example of a scenarios tree.

We first describe the deterministic equivalent form of this problem in terms of a dynamic program. If the stages are 1 to H, we can define the states as $x^t\left(\omega^t\right)$. Noting that the only interaction between periods is through this realization, we can define a dynamic programming type of recursion. For terminal conditions we have

$$\begin{aligned} &Q^H(x^{H-1},\xi^H(\omega)) = \min c^H(\omega)x^H(\omega) \\ &s.t. \quad W^H x^H(\omega) = h^H(\omega) - T^{H-1}(\omega)x^{H-1}, \quad x^H(\omega) \geq 0. \end{aligned} \tag{34}$$

Letting

$$\vec{Q}^{t+1}(x^t) = E_{\xi^{t+1}}[Q^{t+1}(x^t,\xi^{t+1}(\omega))]$$

for all t, we obtain the recursion for $t = 2, ..., H-1$,

$$\begin{aligned} &Q^t(x^{t-1},\xi^t(\omega)) = \min c^t(\omega)x^t(\omega) + \vec{Q}^{t+1}(x^t) \\ &s.t. \quad W^t x^t(\omega) = h^t(\omega) - T^{t-1}(\omega)x^{t-1}, \quad x^t(\omega) \geq 0. \end{aligned} \tag{35}$$

where we use x^t to indicate the state of the system. Other state information in terms of the realizations of the random parameters up to time t, should be included if the distribution of ξ^t is independent of the past outcomes.

The value we seek is:

$$\min z = cx^1 + \vec{Q}\left(x^1\right) \qquad (36)$$
$$\text{s.t. } W^1 x^1 = h^1 \qquad x^1 \geq 0$$

which has the same form as the two-stage deterministic equivalent program.

It is plain that an obvious extension to the above simple 2S-SLPR is to have more stages. In such cases it is common that:
- the stochastic elements have a discrete distribution;
- the realizations of the stochastic elements are represented as a number of scenario's of the future.

Although the original problem had stochastic elements the use of scenarios to explicitly represent the set of all possible futures has enabled us to formulate the problem deterministically.

III.2. Approaches for Solving MS-SLPR **A stochastic programming implementation of a deterministic model means that one has to deal with two challenges:**

- First, the generation of a much larger mathematical programming problem which combines several variants of the deterministic version.
- Second, a heavy computational burden, since the size of the model is roughly multiplied by the number of nodes in the scenarios tree.

The real computational difficulty arises due to the number of scenarios that can often appear. Observe that we essentially need a complete set of constraints for each possible scenario.

III.2.1. Nested Decomposition ([7]) The Nested decomposition for the deterministic case is presented first in [18] and [17]. Actually their approaches represent a inner linearization. They treat all former periods as subproblems of the main problem for the current period.

The difficulty with these primal (originally) included decomposition or inner linearization methods consists in the fact that the set of inputs can be entirely different for different last periods of the realizations. Nevertheless, some results are reached in [29], as it will be described, applying inner linearization of the dual problem and this is again an outer linearization of the main problem.

The common original (primal) approach is, therefore, to use the outer linearization, built on the two-stage L-shaped method. The structure can be used to render the dynamic problems aspects, rendering the uncertainty in all parameters and securing random parameters for moderate number scenarios. The problem is decomposed into the subproblems of each scenario.

As before, you need to be clear about the sequence that applies as you move down the tree in order to formulate the problem correctly. In any MS-SLPR it is important to be clear about the order that pertains as you move through a scenario. In particular you should be clear at each stage about which stochastic elements have been realized and which unrealized.

Note here that this is a key point - *scenarios with a common history must have the same set of decisions* - and must always be taken into account when formulating a scenario based decision problem.

Louveaux is the first to present that multi-stage quadratic problem generalization. Birge extended the two-stage method for the linear case. The same approach is observed also in Pereira and Pinto [31].

The basic concept of "Nested" L-shaped method or Benders decomposition method is to place cuts on $\widehat{Q}^{i+1}\left(x^i\right)$ in (35) and add other cuts in order to reach x^i. Cuts represent successive linear approximations of $\widehat{Q}^{i+1}\left(x^i\right)$. Because of polyhedral structure of $\widehat{Q}^{i+1}\left(x^i\right)$ this is a convergence process to an optimal solution of a finite number steps. To summarize then, as we move down the scenario tree (for any scenario) we have the following order of actions:

- in the first stage a decision as to how much to produce; then
- in the second stage a realizations of the stochastic element; then
- a decision as to the values of the recourse variables; then
- in the second stage a decision as to how much to produce; then
- in the third stage a realizations of the stochastic element ;
- in the $H-$th stage a realizations of the stochastic element; and finally
- a decision as to the values of the recourse variables.

Many alternative strategies are possible in this algorithm in terms of determining the next subproblem to solve. One may alternatively enforce a move from t to $t-1$ ("fast back") or from t to $t+1$ ("fast forward") whenever it is possible. From experiments conducted by Gassmann [15], "fast-forward-fast-backward" sequencing protocol seems to work better than either of these alternatives. For very good review of possible modification in this method see chapter 7 in [7]. An advantage of Nested L-Shaped algorithm the possibility of parallel realization. For detail descriptions of the parallel procedures see [32].

III.2.2. Parallel Decision Yang and Zenios in [38] have implemented the factorization in parallel on a Connection Machine CM-5 with up to 64 processors. Due to memory efficiencies in the parallel version, they actually achieve super linear speedups in parallel compared to serial times. They also report solutions of problems with up to 18 million variables and almost 3 million constraints. These data seem to show that the special factorizations offer exceptional advantages in the framework of the stochastic programs inner point methods. The concrete factorization form can be defined by the accessibility of the parallel processing and by the possibilities to solve the necessary problems in parallel. *But it is affirmed that the decomposition methods are the best for the parallel solutions.*

III.2.3. Dynamic Generalized Network The multi-stage stochastic nonlinear programs with recourse can be presented by the generalized network formulations. This structure can be used to render the dynamic problems aspects, rendering the uncertainty in all parameters and securing random parameters for moderate number scenarios.

The network structure is used in the decision procedure. The problem is decomposed into the subproblems of each scenario.

The nonanticipativity of the constraints is an obstacle to maintain the network structure of each subproblem. The desired decomposition is obtained through the dualization of the constraints nonanticipativity.

In [26] the increased numbers of scenarios are included in the network formulation of the problem for financial planning. They use the option of variables splitting with additional observation, and many of the columns are zero, and it is not necessary to split the corresponding variables. By splitting only those variables with nonzero elements they develop a *partial splitting* model, which proves very efficient. Mulvey and Vladimiru in [25], [27], [28] present some aspects of the generalized stochastic network models. See also the review of Ziemba and Mulvey in [39], describing the model general design.

References

[1] Benders J.F., Partitioning procedures for solving mixed-variables programming problems, *Numerische Mathematik*, **4**, 1962, 238-252.

[2] Birge J.R., Solution Methods for Stochastic Dynamic Linear Programs, Ph.D. Disertation, *Stanford University and Technikal Report SOL 80-29,* Systems Optimization Laboratory, Stanford Univesyty, December 1980.

[3] Birge J.R., The relationship between the L-shaped metod and dual basis factorization for stochastic linear programming, in: Y.Ermoliev, R.Wets, eds., *Numerical Techniques for Stochastic Optimization,* Springer, Berlin, 1988, 267-272.

[4] Birge J.R., Current Trends in Stochastic Programming Computation and Applications, 1995. (*see http://users.iems.nwu.edu/simjrbirge/.*

[5] Birge J.R., M.A.H. Dempster, H.I.Gassmann, E.A.Gunn, A.J.King, S.W. Wallace, A standard input format for multistage stochastic linear programs, *COAL Newsletter*, **17**, 1987, 1-19.

[6] Birge J.R., R.M.Freund, R.J.Vanderbei, Prior reduced fill-in in the solution of equations in interior point algorithms, *Operations Research Letters*, **11**, 1992, 195-198.

[7] Birge J.R., F. Louveaux, Introduction to Stochastic Programing, Springer, 1997.

[8] Birge J.R., L.Qi, Computing block-angular Karmarkar projections with applications to stochastic programming, *Management Science*, **34**, 1988, 1472-1479.

[9] Carino D., T.Kent, D.Myers, C.Stacy, M.Sylvanus, A.Turner, K.Watanabe, W.Ziemba, The Russel-Yasuda Kasai Model: An Asset/Liability Model for a Japanese insurance company using multistage stochastic programming, *Interfaces*, **24**, 1994, 29-49.

[10] Carpenter T., I.Lustig, J.Mulvey, Formulating stochastic programs for interior point methods, *Operations Research*, **39**, 1991, 757-770.

[11] Choi I.C., C.L. Monma, D.F.Shanno, Furter development of a primal-dual interior point method, *ORSA Jouranl on Computing*, **2**, 1994, 304-311.

[12] Coudevaux -- Lauloy C., E. Fragniere, SETSTOCH: A tool for Multistage Stochastic Programming with Recourse. 06.08.1998.

[13] Dantzig G.B., P. Wolfe, The decomposition principle for linear programs, *Operations research* **8**, 1960, 101-111.

[14] Duffie D., Dynamic Asset Pricing Theory, Princeton University Press 2001.

[15] Gassman H.I., MSLiP: a computer code for the multistage stochastic linear programming problem, *Mathematical Programming*, **47**, 1990, 407-423.

[16] Geoffrion A.M., Elements of large-scale mathematical programming, *Management Science*, **16**, 1970, 652-675.

[17] Glasei C.R., Nested decomposition and multistage linear programs, *Manegement Science*, **20**, 1973, 282-292.

[18] Ho J.K., A.S. Manne, Nested decomposition for dynamic models, *Mathematikal Programing*, **6**, 1974, 121-140.

[19] Kall P., Stochastic Linear Programing, Springer-Verlag, Berlin, 1976,

[20] Kusy M., W.Ziemba, A bank asset and liability management model, *Operations Research*, **35**, 1986, 356-376.

[21] Lustig I.J., R.E.Marsten, D.F.Shanno, Computational experience with a primal-dual interior point method for linear programming, *Linear Algebra and Its Aplication*, **152**, 1991, 191-222.

[22] Lustig I.J., R.E.Marsten, D.F.Shanno, Interior point methods for linear programming: Computational state of the art, *ORSA Journal on Computing*, **6**, 1994, 1-14.

[23] Merton R.C., Continuous Time Finance, Blackwell, 1990.

[24] Mulvey J.M., Introduction to financial optimization: Mathematical programming special issue, Math. Program., **B89** 2001, 205-216.

[25] Mulvey J.M., H. Vladimiru, Stochastic network optimization models for investment planning , *Annals of Operations Research*, **20**, 1989, 187-217.

[26] Mulvey J.M., H.Vladimirou, Applying the progressive hedging algorithm to stochastic generalized networks, *Annals of Operations Research*, **31**, 1991, 399-424.

[27] Mulvey J.M., H. Vladimiru, Solving multistage stochastic networks an aplikation of scenario aggregation, *Networks* , **21**, 1991, 619-643.

[28] Mulvey J.M., H.Vladimiru, Stochastic network programming for financial planning problems, Management Science, **38**, 1992, 1642-1664.

[29] Noel M.C., Y. Smeers, Nested decomposition of multistage nonlinear programs with recourse, *Mathematical Programming*, **37**, 1987, 131-152.

[30] Oksendal B., Stochastic Differential Equations, Springer, 1998.

[31] Pereira M.V.F., L.M.V.G. Pinto, Stochastic optimization of a multireservoir hydroelectric system – A decomposition approach, *Water Recourse Research*, **21**, 1985, 779-792.

[32] Ruszczynski A., “Parallel decomposition of multistage stochastic programming problems”, *Mathematical Programming*, **58**, 1993, 201-228.

[33] S t r a z i c k y B ., Some results concerning an algorithm for the discrete recourse problem, in: M.A.H. Dempster, ed., Stochastik Programing, Academik Press, New York, 1980.

[34] T o k a t Y . , S . R a c h e v , E . S c h w a r t z, Asset Liability Management: A Review and Some New Results in the Presence of Heavy Tails, in *Handbook of Heavy Tailed Distributions in Finance,* North Holland Handbooks of Finance (Series Editor W. T. Ziemba), 2002.

[35] T o k a t Y . , S . R a c h e v , E . S c h w a r t z, The Stable non-Gaussian Asset Allocation: A Comparison with the Classical Gaussian Approach, *Journal of Economic Dynamics and Control*, **27**, 2003, 937-969.

[36] V a n S l y k e R . , R . J - B W e t s, L-shaped linear programs with application to optimal control and stochastic programming, *SIAM Journal on Applied Mathematics*, **17**, 1969, 638-663.

[37] V a n d e r b e i R . J ., LOQO user`s manual, *Technical Report SOR 92-5*, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1992.

[38] Y a n g D . , S . A . Z e n i o s, A scalable parallel interior point algorithm for stochastic linear programming and robust optimization, *Report 95-07,* Department of Public and Busines Administration, University of Cyprus, Nicosia, Cyprus, February 1995.

[39] Z i e m b a W . T . , J . M . M u l v e y, Worldwide Asset and Liability Modeling, Cambridge University Press, 1998.

[40] *http://www.ms.ic.ac.uk/jeb/or/sp.html*

Received 31.01.2005

[1] *Krastyu Gumnerov*
Institute of Information Technologies
Bulgarian Academy of Sciences
Sofia, BULGARIA
e-mail: gumnerov@iinf.bas.bg

[2] *Racho Denchev*
Faculty of Mathematics and Informatics at the
University of Sofia "St. Kliment Ohridski"
Sofia, BULGARIA
e-mail: rachodenchev@yahoo.com