Time and nodal decomposition with implicit non-anticipativity constraints in dynamic portfolio optimization

Diana Barro
d.barro@unive.it

Elio Canestrelli
canestre@unive.it

Dept. of Applied Mathematics
University of Venice, Dorsoduro 3825/E
30123 Venice Italy

Abstract

We propose a decomposition method for the solution of a dynamic portfolio optimization problem which fits the formulation of a multistage stochastic programming problem. The method allows to obtain time and nodal decomposition of the problem in its arborescent formulation applying a discrete version of Pontryagin Maximum Principle. The solution of the decomposed problems is coordinated through a fixed-point weighted iterative scheme. The introduction of an optimization step in the choice of the weights at each iteration allows to solve the original problem in a very efficient way.

1 Introduction

In this contribution we analyze a solution approach for a dynamic portfolio optimization problem.

The model under investigation is a sequential decision problem under uncertainty, in a discrete time framework, and fits into a multistage stochastic programming formulation. We assume a discrete probability distribution for the stochastic component of the problem and the evolution of probabilistic information is described by means of a scenario tree structure. These assumptions together with the specification of the non-anticipativity constraints, which can be added to the problem both in explicit and in implicit form, allows to recast the problem into a large-scale deterministic equivalent optimization problem.

To overcome the difficulties of exponentially increasing dimensions of the problem to be solved, in this contribution, we propose a decomposition method which can be obtained combining the main features of the stochastic programming problem and of a discrete time version of the Pontryagin Maximum Principle.

The proposed method applies to the deterministic equivalent problem written in the arborescent form and allows to obtain a time decomposition and, within each stage, a further nodal decomposition of the problem.

The stochastic programming approach to financial optimization, and to dynamic portfolio problems in particular, is well document in the literature, see for example [3][13][14] [15][22][23][46]. For a collection of stochastic programming applications in many different fields see [45].

In section 2 we analyze the description of the stochastic component of the problem and the formulation of implicit or explicit non-anticipativity constraints. In section 3 we briefly discuss solution approaches proposed in the literature. In section 4 we present the dynamic and stochastic portfolio management problem considered. In section 5 we describe in detail the time and nodal decomposition approach and the iterative solution method. Moreover in section 6 we provide some computational results comparing the proposed approach with the time decomposition in the case of explicit non-anticipativity constraints and with the direct solution of the global problem. Section 7 concludes.

2 Implicit versus explicit non-anticipativity constraints

A common method to characterize the uncertainty in a stochastic optimization problem is to introduce a set of scenarios with assigned probabilities. This can be done assuming that the probability distribution P of the random quantities in the model can be described or approximated by a discrete distribution with a finite number of values.

The discrete distribution can be represented by mean of an event tree where each node in the tree is associated with a realization of the stochastic quantities. As time passes information is revealed, this is clearly represented by the structure of the event tree where scenarios, that at early stages share information, progressively become unique. In a multistage framework information is revealed through time and at each stage the decision process can depend on the decisions made at previous stages and on the realizations of the stochastic quantities but it cannot anticipate future outcomes, that is it cannot use information which is not available yet.

The formulation of the non-anticipativity of the decision process is crucial for the formulation of multistage stochastic programming problems and is strictly linked with the solution approaches that can be applied. There are two possible ways to introduce this requirement.

In the first the nodes of the tree are associated with decision stages and by introducing a vector of decision variables for each node of the tree the property of non-anticipativity is automatically fulfilled. The non-anticipativity constraints are implicit in the formulation of the problem.

In the second approach we can split the tree considering each scenario separately. This allows to obtain S dynamic problems characterized by the same structure and where there is no more uncertainty about the future since each scenario is a unique path. The main difficulty is that each problem now includes information on the outcomes up to the end of the horizon and the non-anticipativity feature of the decision process is not guaranteed. Thus non-anticipativity constraints must be added explicitly to ensure feasibility of the decisions with respect to the set of information constraints.

In both cases the introduction of an event tree, that is of a finite number of possible outcomes for each stage t, to describe uncertainty allows to create the so-called deterministic equivalent problem which can have implicit or explicit non-anticipativity constraints (see for example [7]). The resulting problem is characterized by high dimensions and the block diagonal structure of the matrix of constraints, that is we obtain large-scale optimization problems, which requires the use of decomposition methods to be solved efficiently.

Both in the case of explicit formulation of the non-anticipativity constraints and in the case of implicit one the deterministic optimization problem obtained from our multistage portfolio optimization model can be tackled in the framework of discrete time optimal control problem with mixed constraints.

In the first case, already analyzed in [1], we relax the non-anticipativity constraints obtaining separability with respect to the scenarios. Each scenario problem is a deterministic and dynamic discrete time optimal control problem which can be solved applying a time decomposition scheme, see [11][12].

3 Solution approaches and decomposition methods

As widely pointed out in the literature (see for example [26][37][41][43]), the dimension of the deterministic equivalent problem, obtained from at least partially realistic applications, becomes soon too large to be tractable by direct solvers, even if the continuous improvement of computer capabilities allows to move farther and farther the frontier of solvable problems.

Nevertheless these problems usually present special structures which can be approached with solution methods based on decomposition.

According to the literature, see for example the review in [5], solution approaches for multistage stochastic programming problems can be broadly classified into two main groups.

In the first we can collect general purpose algorithms which have been specialized to improve the data structures and the solution strategies according to the features, i.e. sparsity or block diagonal structure, of matrices involved in stochastic programming problems. Among these approaches we may cite [6][8][18][24].

In the second group we can gather decomposition approaches which take advantage of the stochastic program structure aiming at reducing the original problem into a collection of smaller and easier to solve sub-problems.

With a broad classification we can distinguish between methods that result in a nodal decomposition and methods that produce a scenario decomposition of the original problem.

In the first case the original problem is decomposed into a collection of subproblems each related to a node of the event tree thus the original problem results decomposed with respect to time, see for example [4][9][16][20] [38][40][42].

In the second case each subproblem corresponds to a scenario and the original problem is decomposed according to the stochastic component, see for example [27][28] [30][35] [39].

In [37] the authors propose an augmented lagrangian decomposition method which can be applied to obtain either a decomposition according to stages or a decomposition according to scenarios.

For decomposability features in the framework of large-scale linear-quadratic programming and relations with discrete time optimal control problem and stochastic programming see [33][34][36].

For a review of decomposition methods and for more extensive references on solution methods see [5] and [41].

In this contribution we propose a decomposition method for the solution of a dynamic portfolio optimization problem which fits the formulation of a multistage stochastic programming problem. The proposed method combines the main features of the stochastic programming formulation of the problem and of a discrete version of Pontryagin Maximum Principle to obtain a time and nodal decomposition of the original problem which can be solved in a very efficient way.

The proposed method allows to treat nonlinear objective functions which arise in portfolio theory due to risk-averse investors and to exploit the time-decomposability feature provided by discrete time optimal control problems.

Motivation for the development of this method arises from the portfolio management problem but the formulation is quite general and can be adapted to a broader class of problems dealing with planning under uncertainty where the dynamics are linear and the objective function is additive in time.

4 The portfolio model

We consider a dynamic and stochastic portfolio optimization problem over a finite horizon $[0, T]$.

Key features of the dynamic portfolio model are the explicit modelling of the transaction cost at the decisions dates and a risk averse utility function for the investor. The model includes also restrictions on short-selling and borrowing.

For a review of discrete time dynamic portfolio management models see [21] and [29].

We assume that the uncertainty in our problem can be modelled, or approximated, by a discrete distribution of random parameters represented by an event tree.

We assume a general structure for the event tree. We denote with $b(k)$ the ancestor of node k in the previous period and with $d(k)$, with $d(k) = 1, \ldots, D(k)$, a descendant from node k in the following period. There is a root node at time $t = 0$ denoted with $k_0 = 1$ from which the tree originates. At time t there are $K_t - K_{t-1}$ nodes denoted with $k = K_{t-1} + 1, \ldots, K_t$. At the planning horizon there are $S = K_T - K_{T-1}$ leaf nodes.

Each path connecting the root node with a leaf node is a *scenario*, i.e. a sequence of possible realizations. Therefore S is the number of scenarios which corresponds to the number of leaves of the tree.

At the initial date the prices of the risky assets are known while prices and returns at future dates are described by a discrete-time discrete-state stochastic vector process $\{p\}$.

At each trading date, conditionally to previous information, the distribution of prices and returns of risky assets is described by a finite number of realizations of the process $\{p\}$.

Each scenario has a probability of occurrence $\pi_s > 0$, with $\sum_{s=1}^{S} \pi_s = 1$. The model includes purchase and sale variables for each risky asset and a riskless asset as liquidity component of the model (see for example [10]).

The key elements of the model are

$I = \{1, \ldots, n\}$ the set of risky assets among which we can choose the composition of our portfolio;

$x_{ik};\ i \in I$ the amount of the i-th asset in node k;

$a_{ik};\ i \in I$ the amount of the i-th asset purchased in node k;

$v_{ik};\ i \in I$ the amount of the i-th asset sold in node k;

$x_{n+1\ k}$ the amount of liquidity component, or cash, held in portfolio in node k;

cta transaction cost expressed as a percentage of the amount of purchased assets;

ctv transaction cost expressed as a percentage of the amount of sold assets;

$d^+ = (1 + cta); \quad d^- = (1 - ctv)$;

r the risk-free return on the liquidity component of the portfolio, assumed constant over the horizon and across scenarios;

$p_{ik};\ i \in I$ the price of the i-th asset in node k;

$R_k = x_{n+1\,k} + \sum_{i=1}^{n} p_{ik} x_{ik}$ wealth at node k, given by the current value of the portfolio;

$U(\cdot)$ risk averse utility function.

In the following we consider the deterministic equivalent problem with implicit non-anticipativity constraints.

The model in arborescent form is

$$
\begin{aligned}
\max \quad & \sum_{s=K_{T-1}+1}^{K_T} \pi_s U(R_s) && (1)\\
\text{s.t.} \quad & x_{i\,k} = x_{i\,b(k)} + a_{i\,k} - v_{i\,k} && (2)\\
& x_{i\,0} = \bar{x}_0 && (3)\\
& a_{i\,k} \geq 0 && (4)\\
& v_{i\,k} \geq 0 && (5)\\
& x_{i\,b(k)} - v_{i\,k} \geq 0 && (6)\\
& \qquad k = 1, \ldots, K_T \quad i = 1, \ldots, n &&\\
& x_{n+1\,k} = (1+r)[x_{n+1\,b(k)} - d^+ \sum_{i=1}^{n} p_{i\,k} a_{i\,k} + d^- \sum_{i=1}^{n} p_{i\,k} v_{i\,k}] && (7)\\
& x_{n+1\,0} = \bar{x}_{n+1\,0} && (8)\\
& x_{n+1\,b(k)} - d^+ \sum_{i=1}^{n} p_{i\,k} a_{i\,k} + d^- \sum_{i=1}^{n} p_{i\,k} v_{i\,k} \geq 0 && (9)\\
& \quad k = 1, \ldots, K_T. &&
\end{aligned}
$$

where R_s denotes the value of the portfolio in the leaf node corresponding to scenario s, and $x_{i\,b(1)},\ i = 1, \ldots, n+1$ denotes the initial endowment of asset i.

The objective (1) is to maximize the expected utility of final wealth. Constraints (2) and (7) are, respectively, asset inventory constraints and the cash balance equations, no borrowing and no short selling are allowed in the model.

5 Decomposition of the problem

We follow the convention that variables are determined in node k at time t according to the following scheme

- $x_{ib(k)}$ denotes the amount of asset x_i which enters node k, inherited from the ancestor node $b(k)$ and which relates the amount of asset i in the period $[t-1, t]$;
- a_{ik} and v_{ik} denote the decision variables at time t in node k;
- x_{ik} denotes the new amount for asset x_i which is available for period $[t, t+1]$. We follow the convention of assigning this to time $(t+1)$ since it takes into account also the interest that matures in the period $[t, t+1]$ due to the presence of an interest rate on the liquidity deposits.

In equation (2) and (7) we can recognize an implicit dynamics from time t to time $(t+1)$. Exploiting this feature we can write problem (1)-(9) as a discrete time optimal control problem where x_{ik} represent the state variables and a_{ik} and v_{ik} the controls.

The state variables dynamics are given by

$$x_{i\,k}(t+1) = x_{i\,b(k)}(t) + a_{i\,k}(t) - v_{i\,k}(t) \tag{10}$$

$$\begin{aligned} x_{n+1\,k}(t+1) = (1+r)[x_{n+1\,b(k)}(t) - d^{+}\sum_{i=1}^{n} p_{i\,k}(t)a_{i\,k}(t) + \\ +d^{-}\sum_{i=1}^{n} p_{i\,k}(t)v_{i\,k}(t)]. \end{aligned} \tag{11}$$

Accordingly constraints (6) and (9) and the non negativity constraints (4) and (5) can be written as

$$x_{i\,b(k)}(t) - v_{i\,k}(t) \geq 0 \tag{12}$$

$$x_{n+1\,b(k)}(t) - d^{+}\sum_{i=1}^{n} p_{i\,k}(t)a_{i\,k}(t) + d^{-}\sum_{i=1}^{n} p_{i\,k}(t)v_{i\,k}(t) \geq 0 \tag{13}$$

$$a_{i\,k}(t) \geq 0 \tag{14}$$

$$v_{i\,k}(t) \geq 0 \tag{15}$$

5.1 Time decomposition

Problem (1)-(9) can be written as a discrete time optimal control problem, with mixed constraints, where the dimensions of the state and control variables vary with time (see [34]).

Let us denote with $x(t)$ the vectors of state variables and with $u(t)$ the vector of control variables, at time t. The discrete time optimal control problem can be written as

$$\max \quad \{L_T(x(T))\} \tag{16}$$

$$x(t+1) = A(t)x(t) + B(t)u(t) + q(t) \tag{17}$$

$$x(0) = x_0 \tag{18}$$

$$G(t)x(t) + H(t)u(t) + r(t) \leq 0 \tag{19}$$

$$u(t) \geq 0 \tag{20}$$

$$t = 0, \ldots, T-1$$

where the matrices involved in the dynamics (17) and in the mixed constraints (19) are characterized by time varying dimensions and have a block structure as follows

$$A(t) = \begin{pmatrix} \begin{pmatrix} A_1(t) \\ \vdots \\ A_{D(K_{t-2}+1)}(t) \end{pmatrix}_{K_{t-2}+1} \\ \\ \vdots \\ \\ \begin{pmatrix} A_1(t) \\ \vdots \\ A_{D(K_{t-1})}(t) \end{pmatrix}_{K_{t-1}} \end{pmatrix} B(t) = \begin{pmatrix} B_{K_{t-2}+1}(t) & \dots & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & B_{K_{t-1}}(t) \end{pmatrix} \quad (21)$$

$$G(t) = \begin{pmatrix} G_{K_{t-2}+1}(t) & \dots & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & G_{K_{t-1}}(t) \end{pmatrix} \qquad H(t) = \begin{pmatrix} H_{K_{t-2}+1}(t) & \dots & \dots & 0 \\ 0 & \ddots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & H_{K_{t-1}}(t) \end{pmatrix} \quad (22)$$

$$q(t) = \underline{0} \qquad r(t) = \underline{0}. \quad (23)$$

Each sub-matrix is defined as follows

$$A_k(t) = A = \begin{pmatrix} I_n & \underline{0} \\ \underline{0} & (1+r) \end{pmatrix} B_k(t) = \begin{pmatrix} I_n & -I_n \\ -(1+r)d^+ p_k(t)' & (1+r)d^- p_k(t)' \end{pmatrix} \quad (24)$$

$$G_k(t) = G = \begin{pmatrix} -I_{n+1} \end{pmatrix} \qquad H_k(t) = \begin{pmatrix} \underline{0} & I_n \\ d^+ p_k(t)' & -d^- p_k(t)' \end{pmatrix}. \quad (25)$$

Let us denote with $\psi(t+1)$ the lagrangian multipliers associated with the dynamics of the state variables at each node and with $\lambda(t)$ the multipliers associated to the mixed constraints. Tacking into account (23) the generalized Hamiltonian and the Lagrangian of the problem are given respectively by

$$\begin{aligned} \tilde{H}(x(t), u(t), \psi(t+1), \lambda(t)) = \psi(t+1)'[A(t)x(t) + B(t)u(t)] + \\ \lambda(t)'[G(t)x(t) + H(t)u(t)] \end{aligned} \quad (26)$$

$$\begin{aligned} \mathcal{L}(x, u, \psi, \lambda) &= \sum_{s=1}^{S} \pi_s U(R_s(T)) + \sum_{t=0}^{T-1} \psi(t+1)'[A(t)x(t) + B(t)u(t)] + \\ &\quad + \psi(0)'(x(0) - \hat{x}) + \sum_{t=0}^{T-1} \lambda(t)'[G(t)x(t) + H(t)u(t)]. \end{aligned} \quad (27)$$

Applying a discrete version of Pontryagin Maximum Principle [31], we obtain for any time t the necessary, and in this case also sufficient, optimality conditions which can be written as the optimality conditions for a saddle point of the generalized Hamiltonian as follows

$$
\begin{aligned}
&\max_{u(t)\geq 0}\ \min_{\lambda(t)\geq 0}\ \tilde{H}(x(t),u(t),\psi(t+1),\lambda(t)) && (28)\\
&x(t+1)=A(t)x(t)+B(t)u(t) && (29)\\
&x(0)=\hat{x} && (30)\\
&\psi(t)=A(t)'\psi(t+1)-G(t)'\lambda(t) && (31)\\
&\psi(T)=\frac{\partial U(R(T))}{\partial x(T)} && (32)\\
&\qquad t=0,\dots,T-1.
\end{aligned}
$$

To solve this problem we apply the time decomposition proposed by [11][12] and already applied in the case of explicit non-anticipativity constraints [1].

The resulting decomposed problems are given by

$$
\begin{aligned}
&x(t+1)=A(t)x(t)+B(t)u(t) && (33)\\
&x(0)=\bar{x} && (34)
\end{aligned}
$$

$$
\begin{aligned}
&\psi(t)=A(t)'\psi(t+1)-G(t)'\lambda(t) && (35)\\
&\psi(T)=\frac{\partial U(R(T))}{\partial x(T)} && (36)
\end{aligned}
$$

$$
\begin{aligned}
\max_{u(t)}\quad &\{\psi(t+1)'B(t)u(t)\} && (37)\\
&H(t)u(t)\leq -G(t)x(t) && (38)\\
&u(t)\geq 0 && (39)
\end{aligned}
$$

$$
\begin{aligned}
\min_{\lambda(t)}\quad &\{-[G(t)x(t)]'\lambda(t)\} && (40)\\
&H(t)'\lambda(t)\geq B(t)'\psi(t+1) && (41)\\
&\lambda(t)\geq 0 && (42)\\
&\qquad t=0,\dots,T-1
\end{aligned}
$$

5.2 Nodal decomposition

In the following we point out a further decomposition feature of the problem. Let us denote with $x_k=(x_{i\,k},\dots,x_{n\,k},x_{n+1\,k})$ the vector of decision variables at time t in node k; and with $u_k=(a_{i\,k},\dots,a_{n\,k},v_{i\,k},\dots,v_{n\,k})$ the vector of controls at time t in node k.

For each t conditions (33)-(42) can be decomposed with respect to the nodes of the event tree at time t. The resulting problems are given by

$$x_k(t+1) = A_{b(k)}(t)x_{b(k)}(t) + B_k(t)u_k(t) \tag{43}$$
$$x_k(0) = \bar{x} \tag{44}$$

$$\psi_k(t) = \sum_{j=1}^{D(k)} A_k(t)'\psi_j(t+1) - \sum_{j=1}^{D(k)} G_k(t)'\lambda_j(t) \tag{45}$$
$$\psi_k(T) = \frac{\partial U(R_k(T))}{\partial x(T)} \tag{46}$$

$$\max_{u_k(t)} \quad \{\psi_k(t+1)'B_k(t)u_k(t)\} \tag{47}$$
$$H_k(t)u_k(t) \leq -G_{b(k)}(t)x_{b(k)}(t) \tag{48}$$
$$u_k(t) \geq 0 \tag{49}$$

$$\min_{\lambda_k(t)} \quad \{-[G_{b(k)}(t)x_{b(k)}(t)]'\lambda_k(t)\} \tag{50}$$
$$H_k(t)'\lambda_k(t) \geq B_k(t)'\psi_k(t+1) \tag{51}$$
$$\lambda_k(t) \geq 0 \tag{52}$$
$$k = 1, \ldots, K_T \quad t = 0, \ldots, T-1$$

Conditions (43)-(52) can be solved separately in the framework of an iterative scheme.

As we have shown applying a discrete version of Maximum Principle to the arborescent formulation of the multistage stochastic programming problem we have obtained a time and nodal decomposition of the dynamic portfolio problem.

The main advantage obtained is that the deterministic equivalent problem can be tackled solving a number of smaller and easier subproblems linked together.

The proposed time decomposition applies both to the problem in the case of *implicit* non-anticipativity constraints and to the problem with *explicit* non-anticipativity constraints. In the first case it is self contained and allows to obtain a further nodal decomposition of the problem, in the second case it must be jointly applied with a solution approach that supplies scenario decomposition. We refer to [1] for the development of this time decomposition in the case of explicit non-anticipativity constraints in conjunction with the Progressive Hedging Algorithm [35], and to [2] for an application of the method to a dynamic tracking error portfolio problem. The resulting solution algorithm, in what follows denoted with PMPTD (Portfolio Maximum Principle Time Decomposition), is used as a term of comparison to test the computational efficiency of the time and nodal decomposition proposed in this contribution and denoted with ICMP (Implicit Constraints Maximum Principle). For a different time decomposition within the Progressive Hedging Algorithm see also [19][44].

6 The iterative solution scheme

To obtain the optimal solution of the global problem (16)-(20) we apply an iterative scheme in which, at each iteration, firstly we solve conditions (28)-(32), separately for each time t and each node k, by means of the following subproblems

- I subproblem: conditions (43)-(44)
- II subproblem: conditions (45)-(46)
- III subproblem: conditions (47)-(49)
- IV subproblem: conditions (50)-(52).

and secondly we apply an iterative fixed-point scheme defined by

$$y^{\nu+1} = F(y^{\nu}) \tag{53}$$

where F is the transformation defined by conditions (33)-(42). We set $y^{\nu} = \{x_k(t), k = 1, \ldots, K_T;\ t = 0, \ldots, T\}$; and for each ν the next value $y^{\nu+1}$ is obtained solving the four subproblems for each t and each k. At the first step an initial admissible solution y^0 is obtained fixing $u_k(t) \equiv 0\ \forall k, t$ in (43)-(44). The values obtained for $x_k(t)\ \forall k, t$ together with $u_k(t) \equiv 0\ \forall k, t$ are then used to obtain initial values for $\lambda_k(t)\ \forall k, t$ and $\psi_k(t)\ \forall k, t$ through conditions (50)-(52) and (45)-(46), respectively.

The iterative scheme is applied to (53) according to the mean value iteration method of Mann (see [25] [17] [32]) which, at each step of the algorithm considers a weighted average of the admissible solutions found in previous steps. In the following we denote with z^{ν} the weighted average of optimal solutions up to iteration ν and $z^0 = y^0$. The mean value iteration scheme applied is defined as follows

$$y^{\nu+1} = F\left(z^{\nu}\right) \tag{54}$$

$$z^{\nu} = \sum_{i=1}^{\nu} \delta_{\nu\, i} y^i \tag{55}$$

where $\delta_{\nu\, i}$ denotes the elements of the ν-th row of an infinite triangular matrix Δ with the following properties

$$\delta_{\nu\, i} \geq 0 \quad \forall \nu, i \tag{56}$$

$$\delta_{\nu\, i} = 0 \quad \forall i > \nu \tag{57}$$

$$\sum_{j=1}^{i} \delta_{\nu\, j} = 1 \quad \forall i. \tag{58}$$

Different matrices can be applied, among them the Cesáro matrix (see [25]) given by

$$\Delta = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \dots & \dots \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{1}{i} & \frac{1}{i} & \frac{1}{i} & \dots & \frac{1}{i} & 0 & \dots \end{pmatrix}. \tag{59}$$

Equation (55) represents a weighted average of solutions obtained in the previous steps of the iteration scheme where the weights, $\delta = (\delta_1, \dots, \delta_\nu)$, satisfy the following properties:

$$\sum_{i=1}^{\nu} \delta_i = 1 \tag{60}$$

$$\delta_i \geq 0 \quad \forall i = 1, \dots, \nu. \tag{61}$$

This iterative scheme allows to reach the convergence and has been applied also in the solution approach to the problem with explicit non-anticipativity constraints (see [1]).

In the present contribution we analyze in more detail this step of the algorithm. If we apply the Cesáro matrix, which yields the arithmetic mean, the speed of the convergence of the iterative scheme is slow. In order to improve the speed of convergence we introduce an optimization step which allows us to choose the weights in an optimal way with respect to the objective function of the original problem. We tested two different methods.

In the first we choose the best new point $y^{\nu+1}$ performing an optimal line search step between z^ν and $F(z^\nu)$ and we set $z^{\nu+1} = \beta^* z^\nu + (1 - \beta^*) y^{\nu+1}$ where $y^{\nu+1}$ is given by (54) and the weight β^* is determined as the solution to the following (one-dimensional) optimization problem

$$\max_{\beta} \quad f(\beta) \tag{62}$$

$$0 \leq \beta \leq 1 \tag{63}$$

where f denotes the objective function of the original problem in (1) expressed as a function of β. This approach, which is faster than the Cesáro method, still has a rather slow convergence.

In the second method we apply the same optimization idea considering, instead of the segment generated by the last two iterations, the whole region obtained as convex combination of y^ν obtained in all previous iterations.

At each step of the iterative scheme we do not fix a priori the weights, β_i, as in the Cesàro matrix, but we look for the best choice of the coefficients solving the following optimization problem

$$\max_{\beta} \quad f(\beta) \tag{64}$$

$$\sum_{i=1}^{\nu} \beta_i = 1 \tag{65}$$

$$\beta_i \geq 0 \quad \forall i = 1, \dots, \nu. \tag{66}$$

We note that at each step of the iterative scheme the feasibility of the proposed solution, z^ν, is guaranteed by the constraints imposed on β, since the feasible region of the original problem is convex.

For example, if f is a quadratic utility function of the form $f(R) = R - aR^2$ problem (64)-(66) is a quadratic optimization problem rather easy to solve. In general, if f is nonlinear it is possible either to directly solve the resulting nonlinear optimization problem or to consider a linear-quadratic approximation to $f(\beta)$ in (64) and solve the resulting quadratic optimization problem. It is important to note that the weights must satisfy constraints (65) and (66) and that choosing the weights in an optimal way improve considerably the convergence speed of the iterative scheme.

The number of variables of the optimization problem (64)-(66) increases linearly with the number of iterations of the fixed-point scheme, which ultimately depends on the precision required for the optimal solutions.

The convergence of the iterative scheme is monitored through two different stopping criteria. The first relates the objective function while the second applies to the sequence of the proposed solutions. In more detail if we denote with ϵ_1 and ϵ_2 the parameters for the precisions, we require that $||f^\nu - f^{\nu-1}|| \leq \epsilon_1$ and/or $||z^\nu - z^{\nu-1}||_\infty \leq \epsilon_2$.

This improvement in the iterative scheme has benefits also in the case of PMPTD (see [1]). But it is not so noticeable because each scenario problem is fully deterministic and the fixed-point iterative scheme is not so crucial. In that case the major concern for the convergence is the outer iterative procedure governed by the Progressive Hedging Algorithm. The main drawback of PHA, which is widely documented in the literature, is the sensitivity of the convergence speed and solution accuracy to the penalty parameter involved in the augmented objective function.

6.1 Computational results

In the following we compare three different solution approaches for problem (1)-(9). The first method, referred as PMPTD (Portfolio Maximum Principle Time Decomposition), applies the Progressive Hedging Algorithm obtaining a scenario decomposition and solves each scenario problem applying the Maximum Principle that brought a time decomposition (see [1]). The second method, referred as ICGLOBAL, solves the global deterministic-equivalent optimization problem with a general purpose routine without exploiting the structure. The third method is the ICMP (Implicit constraints Maximum Principle), described in the previous sections, which applies the Maximum Principle to the deterministic equivalent problem written in the arborescent form.

In our tests we consider a quadratic utility function in (1). The objective function (64) is quadratic in the vector of weights β, too. We denote with w the vector whose elements are the final wealths in each scenario, that is the value of the portfolio, corresponding to the vectors of the amounts y and observed prices p. W is the matrix whose columns are given by the vectors w^i, $i = 1, \dots, \nu$ obtained in the first ν iterations of the fixed-point scheme. Moreover using the vector of probabilities assigned to each scenario, π, and denoting $\Pi = diag(\pi)$ a diagonal matrix that has the elements of π as diagonal elements we obtain for (64) the following expression

		PMPTD		ICGLOBAL		ICMP	
n	S	iter.	time(sec.)	iter.	time(sec.)	iter.	time(sec.)
4	64	1196	141	392	0.3	10	0.2
6	512	837	379	2186	78	9	0.5
7	2744	1266	2416	9218	7707	40	3
8	2744	657	1395	i.m.	-	38	3
10	2197	959	2130	9357	7853	51	4
11	2197	1451	3666	i.m.	-	40	3
13	1728	2951	4429	10407	8611	41	3
14	1728	7612	11363	i.m.	-	28	2
17	1331	6042	8389	10401	9714	31	3
18	1331	9931	18261	i.m	-	37	4
22	1000	9389	16370	9799	7678	30	3
23	1000	13790	17313	i.m.	-	29	3
30	27000	-	t.l.	i.m.	-	279	319
30	91125	-	t.l.	i.m.	-	453	2218
30	140608	-	t.l.	i.m.	-	551	3774

Table 1: Comparison among PMPTD, ICGLOBAL and ICMP solution approaches for a set of problems with increasing number of scenarios (S), and risky assets (n); i.m. = insufficient memory and t.l. = time limit exeeded.

$$f(\beta) = \pi' W\beta + a\beta' W'\Pi W\beta \tag{67}$$

We consider a set of test problems with increasing number of scenarios and risky assets. To generate the scenario trees we apply an historical simulation approach using the data from the last three years of the Italian stock market.

In Table 1 we present the number of iterations and the time (in seconds) required by each method. The computational experiments have been carried out on a personal computer with Pentium 4, 3.2 Mhz CPU and 1 GB RAM. The algorithm has been coded using Gauss (Aptech Systems, Inc.) and its quadratic optimization routine.

For PMPTD and ICPM we set the following tolerance parameters $\epsilon_1 = 0.5 \cdot 10^{-5}$ and $\epsilon_2 = 10^{-3}$, while in the case of ICGLOBAL we accept the default tolerance parameter of Gauss optimization routine. Moreover in the case of PMPTD we need to choose a penalty parameter ρ which is crucial in the trade-off between solution accuracy and convercenge speed. A good range of values, in the analized cases where we set the initial wealth $R_0 = 100$ and the utility parameter $a \in [-0.3/R_0, -0.1/R_0]$, proved to be $\rho \in [0.01, 0.1]$.

We observe that in the case of ICGLOBAL the algorithm reaches the insufficient memory limit (i.m.) very soon, while the PMPTD requires a great amount of iterations which results in a time limit exceeded (t.l.), this means that the computational time exceeded the 240 000 seconds.

7 Concluding remarks

We consider a solution approach for a dynamic portfolio optimization method written as a multistage stochastic programming problem. We consider the deterministic equivalent problem in the arborescent form, that is with implicit non-anticipativity constraints.

The problem can be rewritten as a discrete time optimal control problem where the equations of the dynamics of the state variables connect a node with its descendants (forward from time t to time $t+1$), while the dynamics of the adjoint variables connect a node with its (unique) ancestor (backward from time $t+1$ to time t). The mixed constraints represent the feasibility constraints with respect to the information structure of the problem, that is they relate the optimal decision in the node with the endowment received from previous period.

Applying a discrete version of the Maximum Principle to the arborescent formulation of the problem allows to obtain a time and a further nodal decomposition, of the original problem, into smaller subproblems.

To obtain a global solution we apply an iterative scheme in which, at each iteration we solve, for each time step and for each node in the event tree, the four subproblems. The solution obtained at each iteration is certainly feasible but not necessarily optimal. To obtain optimality we propose to apply the iterative mean value method of Mann jointly with an optimization step which allows to optimally choose the weights. This method allows both an efficient decomposition of the deterministic equivalent problem and a convergence towards the optimal solution for the global problem with a limited number of iterations.

The comparison with other solution approaches, such as the direct solution of the global deterministic equivalent problem and the decomposition according to scenarios, shows that the proposed method allows to efficiently solve higher dimensional problems with reduced iterations and very competitive computational times.

The main drawback of the method is the convergence scheme which is usually slow. The use of implicit non-anticipativity constraints allows to consider only one iterative procedure instead of the two nested iterative procedures obtained in the case of explicit non-anticipativity constraints, see [1], avoiding in this way also the problem of choosing an adequate value for the penalty parameter involved in the optimization problem in that case. This avoid the problem of jointly controlling the convergence criteria but requires the introduction of a method to accelerate convergence. Our proposal is to choose the weights in the mean value iterative scheme in an optimal way.

References

[1] D. Barro, E. Canestrelli, Dynamic portfolio optimization: Time decomposition using the Maximum principle with a scenario approach, European Journal of Operational Research, 163, 217-229, 2005.

[2] D. Barro, E. Canestrelli, Tracking error: a multistage portfolio model, presented at the XXXIV EWGFM and XI FFM, Paris 2004. To appear.

[3] A.J. Berger, J.M. Mulvey, The Home Account Advisor, asset and liability management for private investors, in [46], pp. 634-665.

[4] Birge, J.R., Decomposition and partitioning methods for multistage stochastic programs, *Operation Research* , 33, (1985), 989-1007.

[5] J.R. Birge, Stochastic Programming Computation and Applications, INFORMS Journal on Computing, 9(1997)111-133.

[6] J.R. Birge, D.F. Holmes, Efficient solution of two-stage stochastic linear programs using interior point methods, Computational Optimization and Applications, 1(1992)245-276.

[7] J.R. Birge, F. Louveaux, Introduction to stochastic programming, Springer, 1997.

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

[9] J.R. Birge, C.H. Rosa, Parallel decomposition of Large-scale stochastic nonlinear programs, Annals of Operations Research, 64(1996)39-65.

[10] S.P. Bradley, D.B. Crane, A dynamic model for bond portfolio management, Management Science, 19(2)(1972)139-151.

[11] E. Canestrelli, Sulla ricerca di soluzioni ottimali per una classe di problemi di controllo ottimo nel discreto, Atti del IV Convegno A.M.A.S.E.S. Grottaferrata, Roma, 1981.

[12] E. Canestrelli, Soluzione numerica di problemi di controllo ottimo nel discreto a vincoli lineari, Atti del IX Convegno A.M.A.S.E.S. Levico Terme, Trento, 1985.

[13] D.R. Cariño, T. Kent, D.H. Myers, C. Stacy, M. Sylvanus, A.L. Turner, K. Watanabe, W.T. Ziemba, The Russell-Yasuda Kasai model: an asset/liability model for a Japanese insurance company using multistage stochastic programming, Interfaces, 24(1)(1994)29-49.

[14] G. Consigli, M.A.H. Dempster, Dynamic stochastic programming for asset-liability management, Annals of Operations Research, 81(1998)131-161.

[15] G.B. Dantzig, G. Infanger, Multi-stage stochastic linear programs for portfolio optimization, Annals of Operations Research, 45(1993)59-76.

[16] G.B. Dantzig, P. Glynn, Parallel processors for planning under uncertainty, Annals of Operations Research, 22(1990)1-21.

[17] R.L. Franks, R.P. Marzec, A theorem on mean-value iterations, Proceedings of the American Mathematical Society, 30(1971)324-326

[18] J. Gondzio, A. Ruszczyński, A sensitivity method for basis inverse representation in multistage stochastic linear programming problems. Journal of Optimization Theory and Applications, 74(1992)221-242.

[19] T. Helgason, S.W. Wallace, Approximate scenario solutions in the Progressive Hedging Algorithm. A numerical study with an application to fisheries management. Annals of Operations Research, 31(1991)425-444.

[20] J. Higle, S. Sen, Stochastic decomposition: an algorithm for two-stage stochastic programs with recourse. Mathematics of Operations Research, 16(1991)650-669.

[21] R.A. Jarrow, V. Maksimovic, W.T. Ziemba (Eds.), Finance, Handbooks in Operations Research and Management Science vol.9, North-Holland, 1995.

[22] P. Klaassen, Solving stochastic programming models for asset/liability management using iterative disaggregation, in [46], pp. 427-463.

[23] M.I Kusy, W.T. Ziemba, A bank asset and liability management model. Operations Research, 34(3)(1986)356-376.

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

[25] W.R. Mann, Mean value methods in iteration, Proceedings of the American Mathematical Society, 4(1953)506-510.

[26] J.M. Mulvey, B. Shetty, Financial planning via multi-stage stochastic optimization, Computer & Operations Research, 31(2004)1-20.

[27] J.M. Mulvey, A. Ruszczyński, A digonal quadratic approximation method for large scale linear programming, Operations Research Letters, 12(1992)205-215.

[28] J.M. Mulvey, A. Ruszczyński, A new scenario decomposition method for large scale stochastic optimization, Operations Research, 43(1995)477-490.

[29] J.M. Mulvey, W.T. Ziemba, Asset and liability allocation in a global environment, in [21], pp. 435-464.

[30] S.S. Nielsen, S.A. Zenios, A massively parallel algorithm for nonlinear stochastic network problems, Operations Research, 41(1993)319-337.

[31] A.J. Ortega, R.S. Leake, Discrete Maximum principle with state constrained control, SIAM J. on Control and Optimization, 15(1977)119-147.

[32] B.E. Rhoades, Fixed point iterations using infinite matrices, Transactions of the American Mathematical Society, 196(1974)161-176

[33] R.T. Rockafellar, Linear quadratic programming and optimal control, SIAM J. on Control and Optimization, 25(1987)781-814.

[34] R.T. Rockafellar, R.B.J. Wets, Generalized linear quadratic problems of deterministic and stochastic optimal control in discrete time, SIAM J. on Control and Optimization, 28(1990)810-822.

[35] R.T. Rockafellar, R.B.J. Wets, Scenario and policy aggregation in optimization under uncertainty, Mathematics of Operations Research, 16(1991)119-147.

[36] R.T. Rockafellar, C.Y. Zhu, Primal-dual projected gradient algorithms for extended linear quadratic programming, SIAM Journal on Optimization, 3(1993)751-783.

[37] C. Rosa, A. Ruszczynski, On Augmented Lagrangian decomposition methods for multistage stochastic programs, Annals of Operations Research 64(1996) 289-309.

[38] A. Ruszczynski, A regularized decomposition for minimizing a sum of polyhedral functions, Mathematical programming 35(1986) 309-333.

[39] A. Ruszczynski, An augmented lagrangian decomposition method for block diagonal linear programming problems, Operations Research Letters 8(1989) 287-294.

[40] A. Ruszczynski, Parallel decomposition of multistage stochastic programming problems, Mathematical programming 58(1993) 201-228.

[41] A. Ruszczynski, Some advances in decomposition methods for stochastic linear programming, Annals of Operations Research 85(1999) 153-172.

[42] Van Slyke R., Wets R.B.J., L-Shaped linear programs with applications to optimal control andd stochastic linear programs, *SIAM J. Appl. Math.*, 17, (1969), 638-663.

[43] H. Vladimirou, Computational assessment of distributed decomposition methods for stochastic linear programs, European Journal of Operational Research, 108, (1998), 653-670.

[44] S.W. Wallace, T. Helgason, Structural properties of the Progressive Hedging Algorithm, Annals of Operations Research, 31(1991)445-456.

[45] S.W. Wallace, W.T. Ziemba (Eds.), Applications of Stochastic Programming, SIAM Mathematical Programming Series on Optimization, 2005.

[46] W.T. Ziemba, J.M. Mulvey (Eds.), World Wide Asset and Liability Modeling. Cambridge University Press, 1999.