Abstract
In this work we present a model for the solution of the multi-period portfolio selection problem. The model is based on a time consistent dynamic risk measure. We apply $l_1$-regularization to stabilize the solution process and to obtain sparse solutions, which allow one to reduce holding costs. The core problem is a nonsmooth optimization one, with equality constraints. We present an iterative procedure based on a modified Bregman iteration, that adaptively sets the value of the regularization parameter in order to produce solutions with desired financial properties. We validate the approach showing results of tests performed on real data.

Keywords  Portfolio optimization · Time consistency · $l_1$ norm · Constrained optimization

Mathematics Subject Classification  91G10 · 90C30 · 65K05

1 Introduction

In this work we focus on dynamic portfolio selection problem in a Markowitz framework. Dynamic portfolio selection arises in medium and long-term investments, in which one allows decisions to change over time by the end of the investment, taking into account the time evolution of available information. We consider dynamic decision problems formulated in a discrete multi-stage setting, with underlying time evolving continuously. This formulation is usually referred to as multi-period portfolio selection. A set of rebalancing dates is introduced,
which split the investment period into sub-periods; decisions are assumed at the rebalancing dates, and are kept within sub-periods. In order to ensure that investors preferences remain consistent over time, dynamic time consistent models should be defined. Different definitions of time consistency can be found in literature, either related to risk measures or investment policies (Chen et al. 2017). We consider the first case, in which one focuses on the properties of the multi-period risk measure employed for modelling the investment problem; this is time consistent if, according to it, the time evolving filtration related to the evaluation of a stochastic process does not modify decisions taken using values computed previously. Roughly speaking, if we today establish that two investments have the same level of risk, then the same level of riskiness should have been estimated for them yesterday. In Chen et al. (2013) authors show that a separable expected conditional mapping, obtained by summing single-period terms, is a time consistent risk measure in Markowitz framework. We then adopt a separable formulation. In our model single-period terms are defined taking the variance as risk measure; more precisely, we consider single-period minimum variance terms and fix a final target expected wealth. Thus, we consider medium and long-term investments in which investors are not interested with positions at intermediate periods.

A common choice to estimate Markowitz model parameters is to use historical data; correlation between assets returns can lead to ill-conditioned covariance matrices. Different regularization techniques have been suggested with the aim of improving the problem conditioning. Among these, we consider \(l_1\)-regularization to stabilize the solution process, since this provides relevant financial features in solutions. Since solutions establish the amount of capital to be invested in each available security, sparsity means that money are invested in a few securities, the so-called active positions. This allows investor to reduce the number of positions to be monitored and, thus, the holding and transaction costs. The model that we present in this paper leads to a nonsmooth optimization problem, because of the \(l_1\) term, with equality constraints. We solve it by means of Bregman iteration (Bregman 1967). More precisely, we present a modified version of Bregman iteration, following the idea presented in Corsaro and De Simone (2019) for the single-period case. The method is modified to adaptively select the regularization parameter that realizes a good trade-off between the fidelity to data and the financial properties required to solution, given in terms of sparsity and number of short positions. Bregman iteration converts the constraint optimization problem in a short sequence of unconstrained ones; the presence of the \(l_1\) term makes the solution of the involved optimization sub-problem not trivial, thus we apply ad hoc methods to deal with non-smoothness (Beck and Teboulle 2009).

In Sect. 2 we present related work; in Sect. 3 we describe the regularized portfolio selection model; in Sect. 4 we describe Bregman iteration method. In Sect. 5 we show some numerical experiments performed on real data.

## 2 Related work

One of the pioneering works in the field of multi-period portfolio modelling dates back to almost 20 years ago Li and Ng (2000); in that paper authors present an extension of Markowitz mean-variance model to multi-period investments. Their model does not satisfy the time consistency property, which was recognized as a relevant one in the following years. In Basak and Chabakauri (2010) authors obtain the time-consistent solution of the dynamic mean-variance portfolio problem using dynamic programming in a continuous time setting. In Wang and Forsyth (2012) a continuous time setting is assumed as well. The optimal dynamic investment
policy is obtained as numerical solution of a nonlinear Hamilton-Jacobi-Bellman partial differential equation for different mean variance like strategies, among which the time-consistent mean variance. In Chen et al. (2013) multi-period portfolio selection is addressed. Authors introduce a time consistent mean-variance model, which extends the classical Markowitz approach to the multi-period case. The model is based on a separable expected conditional mapping, that, as already pointed out, means that it is build by summing single-period terms; authors prove that this construction provides a time consistent risk measure. This motivated our choice of a separable formulation in the model that we present in this paper.

Different regularization techniques have been suggested for the Markowitz problem in the single-period case; a review of them can be found in Carrasco and Noumon (2012). Among these, penalization techniques have been considered, both for the minimum- and the mean-variance approach. In DeMiguel et al. (2009) $l_1$ and squared $l_2$ norm constraints are proposed for the minimum-variance criterion. In Yen and Yen (2014) an algorithm for the optimal minimum-variance portfolio selection with a weighted $l_1$ and squared $l_2$ norm penalty is presented. The $l_1$-regularization for Markowitz model was introduced in Brodie et al. (2009), where a $l_1$ penalty term is added to promote sparsity in the solution. This approach is referred to as Lasso. It allows investor to reduce the number of positions to be monitored and the holding and transaction costs. Another relevant property of $l_1$-regularization is the penalization of short positions, corresponding to negative solutions. In many markets, restrictions on short sales are established, thus short-controlling is a desired feature. The $l_1$-regularized portfolio selection is successfully considered also in Corsaro and De Simone (2019), where authors show that sparsity and short-controlling can be obtained by properly tuning the regularization parameter. In this work we extend the Lasso approach to multi-period portfolio selection.

Much effort has been addressed to the application of optimization models and techniques to portfolio selection (AitSahlia et al. (2008); Pardalos (1997); Pardalos et al. (1994)). The multi-period model that we present in this paper leads to a nonsmooth optimization problem with equality constraints. We solve it by means of Bregman iteration (Bregman 1967), that was recently introduced with success in many fields formulated in term of $l_1$ minimization (image analysis Antonelli and De Simone 2018; Goldstein and Osher 2009; Yin et al. 2008, matrix rank Ma et al. 2011, portfolio selection Corsaro and De Simone 2019; Corsaro et al. 2018; Ho et al. 2015; O’Donoghue et al. 2013). In particular, in Corsaro and De Simone (2019) a modified version of Bregman iteration is presented. The method is modified to adaptively select the regularization parameter that realizes a good trade-off between the fidelity to data and the financial properties required to solution, given in terms of sparsity and number of short positions. Encouraged by the promising results presented for the single-period case, we present the modified Bregman iteration method for the multi-period case.

Bregman iteration converts the constraint optimization problem in a short sequence of unconstrained ones; the presence of the $l_1$ term makes the solution of the involved optimization sub-problem not trivial, thus we apply ad hoc methods to deal with non-smoothness (Beck and Teboulle 2009).

3 Regularized portfolio selection model

In this section we introduce an $l_1$-regularized model for multi-period portfolio selection. Let $m$ be the number of investment periods. Decisions are assumed at the rebalancing dates $j, j = 1, \ldots, m$; decision taken at time $j$ is kept in the period $[j, j + 1)$. Let $1 \leq t \leq m$ be
the evaluation date. A conditional risk mapping $\rho_{t,m}$ computes the risk of a multi-period loss process at time $t$, that is, its value is the risk estimate at time $t$ of future losses. In probabilistic terms, we have that the information is described by a filtration that evolves according to the dynamics of a certain underlying random process whose risk is to be estimated. The dynamics is supposed not to affect previous estimates under a dynamic time consistent risk measure. To simplify, given two loss processes that produce the same losses up to the evaluation date $t$, if they are estimated equally risky at time $t$, then the same is observed previously under a dynamic consistent risk measure. Time consistency has been recognized to be a relevant property in dynamic asset allocation. Indeed, much effort has been addressed to the analysis and proposal of time consistent risk measures; we address reader to Chen et al. (2017) and references therein for an overview on the existing literature. In Chen et al. (2013) a time consistent dynamic mean-variance model is presented. It is based on a separable expected conditional mapping, which authors show to be time consistent. We here report the definition given in Chen et al. (2013).

Definition 1 A conditional risk measure $\rho_{t,m}$ on the time horizon $T=[t,m]$ is separable if it can be expressed in the following way:

$$\rho_{t,m}(Y_{t,m}) = \sum_{s=t+1}^{m} E[\rho_s(Y_s|F_{s-1})|F_t]$$ (1)

where $Y_{t,m}$ is a random vector of losses estimated at the dates $\{t, t+1, \ldots, m\}$, $F_t$ the filtration at time $t$, $Y_t$ is the random loss adapted to $F_t$ and $E$ is the expectation operator.

Equation (1) states that the risk measure is decomposed into a sum of terms in such a way that each one provides a risk estimate in one period of the investment, using information available at the beginning of the period. In this paper we consider a separable conditional risk measure, taking the variance as single-period risk measure.

Let us denote the number of traded assets by $n$ and with $w_j$ the portfolio of holdings at the beginning of period $j$. Thus, for instance, $(w_j)_i$ is the amount invested in the asset $i$ at the $j$-th rebalancing date. The optimal portfolio is then defined by the vector $w = (w_1, w_2, \ldots, w_m) \in \mathbb{R}^N$, where $N = m \cdot n$ is the problem dimension. Finally, the vector $r_j$ and $C_j \in \mathbb{R}^{n \times n}$ contain respectively the expected return vector and the covariance matrix estimated at time $j$; covariance matrices are assumed to be positive definite. We aim at minimizing the risk of the strategy, estimated by

$$\Phi(w) = \sum_{j=1}^{m} w_j^T C_j w_j.$$ (2)

The objective function (2) typically leads to ill-conditioned problems, because of assets correlation. At this purpose, we apply $l_1$-regularization to stabilize the solution process; an $l_1$ penalty term is added to (2). This technique was applied in the single-period case, that is, in the classical Markowitz approach, in Corsaro and De Simone (2019). In that paper, authors show that $l_1$-regularization provides sparse solutions; since solutions establish the amount of capital to be invested in each available security, sparsity means that money are invested in a few securities. This allows investor to reduce both the number of positions to be monitored and the holding costs.

Thus, we propose the following constrained optimization problem for multi-period portfolio selection:

\[\text{Springer}\]
The objective is to minimize the expected reward subject to constraints on the portfolio weights. This can be written as:

$$
\min_w \sum_{j=1}^{m} \left[ w_j^T C_j w_j + \tau \| w_j \|_1 \right]
$$

subject to:

$$
\begin{align*}
  w_1^T 1_n &= \xi_{\text{init}} \\
  w_j^T 1_n &= (1_n + r_{j-1})^T w_{j-1}, \quad j = 2, \ldots, m \\
  w_m^T 1_n &= \xi_{\text{term}}
  
\end{align*}
$$

where $\tau > 0$ is the regularization parameter, $\xi_{\text{init}}$ is the initial wealth, $\xi_{\text{term}}$ is the target expected wealth resulting from the overall investment, and $1_n$ is a vector of ones of length $n$. The first constraint is the budget constraint. We assume that the strategy is self-financed: this is stated in constraints from 2 to $m$, where it is established that at the end of each period the wealth is given by the revaluation of the previous one. The $(m + 1)$-th constraint defines the expected final wealth.

4 Bregman iteration for portfolio selection

In this section we discuss the solution of the $l_1$-regularized constrained optimization problem (3).

A common approach to solve the generic constrained optimization problem

$$
\begin{align*}
  \min_w & \quad J(w) \\
  \text{s.t.} & \quad H(w) = 0,
\end{align*}
$$

where $J(w)$ is convex and $H(w)$ is non-negative, convex and differentiable, is to convert it into an unconstrained optimization problem. One way to do this is to use a penalty/continuation method, according to which the constrained problem (4) is approximated by problems of the form:

$$
\min_w J(w) + \lambda_k H(w)
$$

where $\{\lambda_k\}$ is an increasing sequence. In many problems it is necessary to choose very large values of $\lambda_k$ and it makes the numerical solution process an extremely difficult one. Bregman iteration can alternatively be used; it allows one to fix the value of $\lambda$.

One of the central concepts of Bregman iteration is the Bregman distance (Bregman 1967) associated with $J$ at point $v$, defined as:

$$
D^p_J(w, v) = J(w) - J(v) - \langle p, w - v \rangle,
$$

where $p \in \partial J(v)$ is a subgradient in the subdifferential of $J$ at point $v$ and $\langle . , . \rangle$ denotes the canonical vector inner product.

By using (5), Bregman iteration can be employed to reduce (4) in a short sequence of unconstrained problems, according to the following iterative scheme:

$$
\begin{align*}
  w_{k+1} &= \arg\min_w D^p_J(w_k, w) + \lambda H(w), \\
  p_{k+1} &= p_k - \lambda \nabla H(w_{k+1}) \in \partial J(w_{k+1}),
\end{align*}
$$

with $\lambda > 0$. Inspired by the results obtained in the single case Corsaro and De Simone (2019), in this section we show that the Bregman iteration is a simple and very efficient method for solving the portfolio selection problem (3), reviewed as a special case of (4), with $J$ non-smooth convex functional. In order to derive the expression of $J$ and $H$ for our financial...
problem, we introduce two block matrices. Let $N = m \cdot n$ and $C = \text{diag}(C_1, C_2, \ldots, C_m) \in \mathbb{R}^{N \times N}$ a $m \times m$ diagonal block matrix with diagonal blocks formed by covariance matrices estimated at the dates of rebalancing. $A \in \mathbb{R}^{(m+1) \times N}$ is the lower bi-diagonal equality constraints block matrix of dimension $(m+1) \times m$, with diagonal blocks $A_{i,i} = -1^n$, for $i = 1, \ldots, m$, and sub-diagonal blocks $A_{i+1,i} = (I_n + r_{i-1})^T$, for $i = 1, \ldots, m - 1$. The portfolio selection problem \((3)\) can then be expressed as \((4)\) with

\[ J(w) = w^T C w + \tau ||w||_1, \quad (7) \]

and

\[ H(w) = \frac{1}{2} ||Aw - b||_2^2, \quad (8) \]

with $b = (\xi_{\text{init}}, 0, \ldots, 0, \xi_{\text{term}})^T \in \mathbb{R}^{m+1}$.

The choice of the regularization parameter in \((7)\) plays a key role to obtain optimal portfolios that meet certain financial requirements and fidelity to data. In the single case, where the weights normalization is assumed, the regularization parameter controls both the sparsity and the number of short positions; indeed it can be easily shown that the $l_1$ term permits to penalize the short positions (Brodie et al. 2009):

\[ \tau ||w||_1 = 2\tau \sum_{i:w_i < 0} |w_i| + \tau. \]

Unlike the single case, due to self-financial constraint, the regularization parameter $\tau$ cannot drive the number of short positions; then it controls only the sparsity. However, as shown in the following, we observe that the number of short positions decreases with respect to $\tau$. For this reason we extend the algorithm proposed in Corsaro and De Simone (2019) to the multi-period case; the algorithm is based on a modified Bregman iteration to automatically select the regularization parameter so to satisfy desired financial requirements. The basic idea is to generate an increasing sequence of parameter values $\tau_k \in [0, \tau_{\text{max}}]$ that tries to produce solutions satisfying a fixed financial target, defined in terms of sparsity or short-controlling. Given $(w_{k+1}, p_{k+1})$ provided by \((6)\) applied to

\[ J_k(w) = w^T C w + \tau_k ||w||_1, \]

if $w_{k+1}$ does not satisfy the financial requirement, $\tau_{k+1} > \tau_k$ is defined. Consequently, $p_{k+1}$ must be changed to guarantee the well-definiteness of Bregman iteration. Following Corsaro and De Simone (2019), it is possible to show that the vector $\tilde{p}_{k+1} \in \partial J_{k+1}(w_{k+1})$, with

\[ \tilde{p}_{k+1} = \frac{\tau_{k+1}}{\tau_k} p_{k+1} + 2 \left( 1 - \frac{\tau_{k+1}}{\tau_k} \right) C w_{k+1}. \]

The proposed multi-period algorithm is summarized in Algorithm 1. Under suitable hypotheses the convergence of the sequence $(w_k)$ to a solution of the constrained problem \((4)\) is guaranteed in a finite number of steps (Osher et al. 2005). Note that the convergence results for Bregman method guarantee the monotonic decrease of $||Aw_k - b||_2^2$, thus for large $k$ the constraint conditions are satisfied to an arbitrary high degree of accuracy. This yields a natural stopping criterion according to a discrepancy principle.

Since there is generally no explicit expression for the solution of the sub-minimization problem involved in \((6)\), at each iteration the solution is computed inexactly using an iterative solver. At this purpose, we focus on first order methods, which are gradient-based that converge rather slowly; however, for large problem dimensions, usually a fast lower-precision

\[ Springer \]
Algorithm 1 Modified Bregman Iteration for portfolio selection

Given $\tau_0 > 0$, $\tau_{\text{max}}$, $\lambda$, $\theta > 1$ % Model parameters

Given $n_{\text{short}}$, $n_{\text{act}}$ % Financial target parameters

$k := 0$

$w_0 := 0, p_0 := 0, \tau_{-1} := \tau_0$,

while “stopping rule not satisfied” do

$p_k = \frac{\tau_k}{\tau_{k-1}} p_k + \left( 1 - \frac{\tau_k}{\tau_{k-1}} \right) C w_k$

$w_{k+1} = \text{argmin}_w D(p_k, w, w_k) + \frac{\lambda}{2} \| A w - b \|_2^2$

$p_{k+1} = p_k - \lambda A^T (A w_{k+1} - b)$

$W_{k+}^ω = [i : (w_{k+1})_i < 0]$

$W_{k+1}^ω = [i : (w_{k+1})_i \neq 0]$

if $|W_{k+1}^ω| > n_{\text{short}}$ or $|W_{k+1}^ω^a| > n_{\text{act}}$ then

$\eta_{k+1} = 0$

else

$\eta_{k+1} = 1$

end if

$\tau_{k+1} = \min\{\eta_{k+1} \tau_k, \tau_{\text{max}}\}$

$k := k + 1$

end while

$\tau_f := \tau_{k+1}$

solution is favoured. In particular, we use the Fast Proximal Gradient method with backtracking stepsize rule (FPG) (Beck and Teboulle 2009), an accelerated variant of Forward Backward algorithm, suitable for minimizing convex objective functions given by summation of smooth and non-smooth terms. Using (7), the objective function in the sub-minimization problem involved in (6),

$$\Phi_k(w) = w^T C w - < p_k, w > + \tau_k ||w||_1 + \frac{\lambda}{2} \| A w - b \|_2^2,$$

can be split into the sum of a smooth ($f_k$) and a non-smooth ($g_k$) term with

$$f_k(w) = w^T C w - < p_k, w > + \frac{\lambda}{2} \| A w - b \|_2^2, \quad g_k(w) = \tau_k ||w||_1.$$

FPG produces a new approximation according to:

$$w_{k+1} = \text{prox}_{\beta_k \beta_k} (w_k - \beta_k \nabla f_k(w_k)),$$

with a suitable $\beta_k$ (Beck and Teboulle 2009). The proximal operator of $g_k$ can be decomposed into a sum of Soft threshold operator, defined as

$$\text{Prox}_{\tau_k ||w||_1}(w_i) = \text{sgn}(w_i) (|w_i| - \min\{|w_i|, \tau_k\}).$$

5 Numerical results

In this section, we present the results of some tests to show the effectiveness of Algorithm 1 for solving the multi-period portfolio optimization problem (3).

In Algorithm 1 we set $\lambda = 1$, $\tau_0 = 10^{-5}$, $\tau_{\text{max}} = 0.5$ and $\theta = 1.5$. Iterations are stopped as soon as $\| A w_k - b \|_2 \leq \text{Tol}$ with $\text{Tol} = 10^{-4}$ that, from the financial point of the view, guarantees constraints at a sufficient accuracy. The maximum number of allowed Bregman iterations is set to 100. Inner iterations are stopped when the relative difference in Euclidean norm between two successive iterates is less than $\text{Tol}_{\text{inn}} = 10^{-5}$. © Springer
The tests have been performed in Matlab R2018a environment, on a PC with Intel Core i5-8250U processors, running Windows 10 Pro 64 bit.

We compare our investment strategy with the so-called $1/n$ strategy, where one invests the same amount of money in all available assets. The portfolio built following this strategy is referred to as the \textit{naive} portfolio. It is a common choice to take it as benchmark since investors often apply this heuristic as well as simple rule to allocate their wealth across assets, perceiving it as a diversification strategy that allows one to reduce risk (DeMiguel et al. 2009).

We assume that the investor has one unit of wealth at the beginning of the planning horizon, that is, $\xi_{\text{init}} = 1$. In order to compare optimal portfolio with the naive one, we set as expected final wealth the expected wealth of the naive one, that is, $\xi_{\text{term}} = \xi_{\text{naive}}$. The expected wealth of the naive portfolio is obtained by applying recursively the $1/n$ allocation rule. At each rebalancing date the wealth is evenly invested among the available securities, that is:

$$
\xi_{\text{naive}} = \frac{1}{n} \left( \cdots \left( \frac{1}{n} \left( \xi_{\text{init}} + r_1^T 1_n \right)^T (1_n + r_2^T 1_n) \cdots \right)^T (1_n + r_m^T 1_n) \right)
$$

We show results obtained using three real data sets, described below. The first and the second data sets come from Fama and French database.\footnote{Data available at \url{http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#BookEquity}.} The third case study refers to the EURO STOXX50 Index.

- **Test 1: FF48** The first database—denoted with FF48—contains monthly returns of 48 industry sector portfolios from July 1926 to December 2015. We simulate investment strategies of length 10, 20 and 30 years, with annual rebalancing.

- **Test 2: FF100** The second database—denoted with FF100—contains data of 100 portfolios which are the intersections of 10 portfolios formed on size and 10 portfolios formed on the ratio of book equity to market equity. Also FF100 contains monthly returns from July 1926 to December 2015. We consider 96 portfolios of the 100 available, selected with a preprocessing procedure which eliminates elements with highest volatilities. Also in this case we simulate investment strategies of length 10, 20 and 30 years, with annual rebalancing.

- **Test 3: EURO STOXX50** The third database—denoted with ES50—contains the daily returns of stocks included in the EURO STOXX 50 Index Europe’s leading blue-chip index for the Eurozone. The index covers 50 stocks from 11 Eurozone countries: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain. The dataset contains daily returns for each stock in the index from January 2008 to December 2013. For this test case we consider both annual and quarterly rebalancing.

In Table 1 we show results of tests in which the investor requires that the number of active positions is at most the 30\% of the available positions during the investment period, that is $n_{\text{act}} = 0.3N$. It is equivalent to require at least the 70\% of sparsity in the solution. In this case there is no limit on the number of short sells and $n_{\text{short}} = N$ in Algorithm 1.

Optimal and naive portfolios are compared in terms of risk, given by (2). In particular the ratio between the estimated risk of the $1/n$ strategy and the optimal strategy is reported. For each test case we report the period of the investment. Note that in all cases the financial Crisis period is included in the simulation. For optimal portfolios we also report the output value $\tau_f$ of the regularization parameter, the number of Bregman iterations, the percentage of short positions and sparsity. Results show that the required level of sparsity is achieved
Table 1 Results for $n_{\text{act}} = 0.3N$, $n_{\text{short}} = N$

| Test  | Period           | $\tau_f$   | It   | Shorts (%) | Sparsity(%) | Ratio   |
|-------|------------------|------------|------|------------|-------------|---------|
| FF48  | July 2005–June 2015 | 8.65E−04  | 12   | 9          | 72          | 5.41    |
| FF48  | July 1995–June 2015 | 1.30E−03  | 14   | 8          | 70          | 7.72    |
| FF48  | July 1985–June 2015 | 4.38E−03  | 17   | 6          | 76          | 8.83    |
| FF100 | July 2005–June 2015 | 5.77E−04  | 11   | 8          | 80          | 5.99    |
| FF100 | July 1995–June 2015 | 8.65E−04  | 12   | 8          | 77          | 15.35   |
| FF100 | July 1985–June 2015 | 4.38E−03  | 17   | 4          | 85          | 16.84   |
| ES50  | Jan. 2008–Dec. 2013 (quarterly) | 1.71E−04  | 8    | 1          | 79          | 3.25    |
| ES50  | Jan. 2008–Dec. 2013 (annual)   | 1.71E−04  | 8    | 1          | 82          | 3.45    |

The first column contains the label identifying the test case. The columns 2–7 contain respectively the investment period, the final value of the regularization parameter, the number of Bregman iterations, the percentage of short positions, the level of sparsity of solutions, the ratio between the risk of the $1/n$ strategy and the optimal one.

Table 2 Number of short positions produced with fixed values of $\tau$

| Period             | $\tau$ | $10^{-5}$ | $10^{-3}$ | $10^{-1}$ |
|--------------------|--------|-----------|-----------|-----------|
| FF48 July 2005–June 2015 | 202    | 35        | 0         |
| FF48 July 1995–June 2015 | 426    | 87        | 0         |
| FF48 July 1985–June 2015 | 670    | 225       | 1         |

in all cases, with a modest number of Bregman iterations. The $1/n$ strategy is at least 3.25 times riskier than the optimal one, for FF100 the ratio is close to 17.

The percentage of shorting varies from the 1% of $N$ for test SP50 to the 9% of $N$ for test FF48 with 10-years simulation.

We note that for the longest investment periods we have the highest values of $\tau_f$; consequently, also the number of required Bregman iterations and the sparsity level are slightly higher.

As already pointed out, we observe that the number of short positions decreases with respect to $\tau$. In order to show this behaviour, in Table 2 we report the number of short positions in optimal portfolios produced with fixed values of the regularization parameter, obtained by forcing $\theta = 1$ in Algorithm 1 for FF48 dataset. Same results are obtained for the other datasets. As explained in the previous section, this motivates the introduction of short-selling control in Algorithm 1. Table 3 shows how the proposed adaptive rule for the selection of the regularization parameter works when short positions are not allowed. This requirement is imposed by setting $n_{\text{short}} = 0$, $n_{\text{act}} = N$. We note that the algorithm is able to find a solution with the desirable financial features in almost all the test. One short position is produced for FF48 in a 30-years investment. This result is consistent with the ones reported in Table 2, where one short position is produced for $\tau = 0.1$, since the value of $\tau_f$ reaches the maximum allowed $\tau_{\text{max}}$, and $\tau_{\text{max}}$ is $O(10^{-1})$. We observe that, in general, no-short selling requires quite large values of $\tau$. This induces high levels of sparsity with a slight increase of risk, which is anyway lower then the risk of the $1/n$ strategy. This suggests that the choice of a suitable $\tau_0$ depends on the financial target, that is, it could be convenient to set a slightly larger starting value to reduce the iterations if we require no-short controlling.
Table 3 Results for $n_{short} = 0$, $n_{act} = N$

| TEST      | Period               | $\tau_f$ | It | nshorts | Sparsity (%) | Ratio |
|-----------|----------------------|-----------|----|---------|--------------|-------|
| FF48      | July 2005–June 2015  | $6.57E-03$| 18 | 0       | 88           | 2.83  |
| FF48      | July 1995–June 2015  | $2.22E-02$| 22 | 0       | 89           | 2.36  |
| FF48      | July 1985–June 2015  | $5.00E-01$| 47 | 1       | 97           | 1.46  |
| FF100     | July 2005–June 2015  | $6.57E-03$| 18 | 0       | 93           | 1.78  |
| FF100     | July 1995–June 2015  | $3.35E-02$| 22 | 0       | 96           | 1.71  |
| FF100     | July 1985–June 2015  | $5.00E-01$| 47 | 1       | 98           | 1.22  |
| ES50      | Jan. 2008–Dec. 2013  | $8.65E-04$| 12 | 0       | 82           | 2.82  |
| ES50      | Jan. 2008–Dec. 2013  | $5.77E-04$| 11 | 0       | 77           | 2.93  |

The first column contains the label identifying the test case. The columns 2–7 contain respectively the investment period, the final value of the regularization parameter, the number of Bregman iterations, the number of short positions, the level of sparsity of solutions, the ratio between the risk of the $1/n$ strategy and the optimal one.

Our model does not explicitly take into account transaction costs; however, producing sparse solutions has an impact on transaction costs as well, allowing one to reduce them. To see this, we count the transactions. According to formulation (3) of the financial problem, if $(w_j)_i \neq (w_{j+1})_i$ we assume that security $i$ has been bought or sold in the period $[j, j+1]$. Note that this is a pessimistic estimate of transaction costs because weights could change also for effect of revaluation. Let us introduce the matrix $G \in \mathbb{R}^{n \times m}$, with:

\[
\begin{cases}
G_{i,j} = 1 & \text{if } (w_j)_i \neq (w_{j+1})_i \\
G_{i,j} = 0 & \text{otherwise}
\end{cases}
\]

for $i = 1, \ldots, n$ and $j = 1, \ldots, m$. The number of transactions associated with the optimal strategy is then given by:

\[
T = \sum_{i=1}^{n} \sum_{j=1}^{m} G_{i,j}.
\]

In Table 4 we report the number of transactions $T$ for all the tests. We denote with $T_{naive}$ the number of transactions of the $1/n$ strategy. $T_{nact}$ and $T_{nshort}$ are, respectively, the number of transactions of the optimal portfolios when either sparsity or no-shorting are required. From the table it is evident that the reduction of transaction costs is significant: the naive portfolio exhibits a number of transactions that is at least three times the corresponding value of the optimal portfolios. In general costs associated to the no-short strategy are lower due to the higher sparsity levels observed before.

Finally, in Fig. 1 we represent the optimal portfolio weights over time for two tests; for the sake of readability we consider FF48, 10-years investment, and ES50 with annual rebalancing, setting $n_{short} = 0$ and $n_{act} = N$ for both. The number of colored areas at each date represents the number of assets among which the wealth is allocated, while the height of each colored area at each date represents the amount of wealth allocated in that asset. We note that only a few assets are involved at each rebalancing date, but the height of the areas varies over time for most of assets. This suggests that a further reduction of transaction costs could be obtained, for instance, modifying the model in such a way to produce sparse solutions that are kept fixed as much as possible.
Table 4  Number of transactions for all the tests and for the naive and the optimal portfolios when either
sparsity or no-shorting are required

| TEST  | Period                     | $T_{\text{naive}}$ | $T_{\text{nact}}$ | $T_{\text{rshort}}$ |
|-------|----------------------------|--------------------|-------------------|---------------------|
| FF48  | July 2005–June 2015        | 480                | 160               | 69                  |
| FF48  | July 1995–June 2015        | 960                | 352               | 137                 |
| FF48  | July 1985–June 2015        | 1440               | 401               | 70                  |
| FF100 | July 2005–June 2015        | 960                | 242               | 93                  |
| FF100 | July 1995–June 2015        | 1920               | 584               | 27                  |
| FF100 | July 1985–June 2015        | 2880               | 553               | 62                  |
| ES50  | Jan. 2008–Dec. 2013 (quarterly) | 1100           | 279               | 234                 |
| ES50  | Jan. 2008–Dec. 2013 (annual)| 300                | 76                | 91                  |

Fig. 1  Asset weights over time. Left: FF48, 10-years investment; right: ES50 with annual rebalancing. $n_{\text{short}} = 0$ and $n_{\text{act}} = N$ have been set for both

6 Conclusion and future work

In this work we present a model and the related solution procedure, based on a modified Bregman iteration, for the multi-period portfolio selection problem. The model is corrected with a $l_1$-regularization term to improve conditioning and obtain sparse solutions. This has an impact on holding and transaction costs. A fundamental point is the choice of the regularization parameter that realizes a good trade-off between sparsity and fidelity to data. We extend the adaptive rule proposed in Corsaro and De Simone (2019) to the multi-period case. Numerical results validate our procedure; moreover, we show that the adaptive selection rule in general can control also the number of short positions. However, in some case the procedure is not able to produce no short solutions, even for quite large values of the regularization parameter.

We show that transaction costs are reduced, even if they are not taken into account in the model. Future work could concern the introduction of a term that explicitly penalizes transactions in the model.

Acknowledgements  This work was partially supported by the Research Grant of University of Naples “Parthenope”, DR no. 953, November 28th, 2016, and by INdAM-GNCS, under Project 2019.
References

AitSahlia, F., Sheu, Y., & Pardalos, P. M. (2008). Optimal execution of time-constrained portfolio transactions (pp. 95–102). Berlin: Springer.

Antonelli, L., & De Simone, V. (2018). Comparison of minimization methods for nonsmooth image segmentation. Communications in Applied and Industrial Mathematics, 9, 68–96. https://doi.org/10.1017/CBO9780511804458.003.

Basak, S., & Chabakauri, G. (2010). Dynamic mean-variance asset allocation. The Review of Financial Studies, 23(8), 2970–3016.

Beck, A., & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Science, 2, 183–202.

Bregman, L. (1967). The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. USSR Computational Mathematics and Mathematical Physics, 7, 200–217.

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. Proceedings of the National Academy of Sciences, 106(30), 12267–12272.

Carrasco, M., & Noumon, N. (2012). Optimal portfolio selection using regularization, Working paper, University of Montreal.

Chen, Z., Consiglio, G., Liu, J., Li, G., Fu, T., & Hu, Q. (2017). Multi-period risk measures and optimal investment policies (pp. 1–34). Cham: Springer.

Chen, Z., Li, G., & Guo, J. (2013). Optimal investment policy in the time consistent mean-variance formulation. Insurance: Mathematics and Economics, 52(2), 145–156.

Corsaro, S., & De Simone, V. (2019). Adaptive l1-regularization for short-selling control in portfolio selection. Computational Optimization and Applications, 72(2), 457–478.

Corsaro, S., De Simone, V., Marino, Z., & Perla, F. (2018). Numerical solution of the regularized portfolio selection problem (pp. 249–252). Berlin: Springer. https://doi.org/10.1007/978-3-319-89824-7_45.

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. Management Science, 55(5), 798–812.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1-n portfolio strategy? The Review of Financial Studies, 22(5), 1915–1953.

Goldstein, T., & Osher, S. (2009). The split bregman for l1-regularization problems. SIAM Journal of Imaging Science, 2(2), 323–343.

Ho, M., Sun, Z., & Xin, J. (2015). Weighted elastic net penalized mean-variance portfolio design and computation. SIAM Journal on Financial Mathematics, 6(1), 1220–1244.

Li, D., & Ng, W. (2000). Optimal dynamic portfolio selection: Multiperiod mean-variance formulation. Mathematical Finance, 10(3), 387–406.

Ma, S., Goldfarb, D., & Chen, L. (2011). Fixed point and bregman iterative methods for matrix rank minimization. Mathematical Programming, 128(1), 321–353.

O’Donoghue, B., Stathopoulos, G., & Boyd, S. (2013). A splitting method for optimal control. IEEE Transactions on Control Systems Technology, 21(6), 2432–2442.

Osher, S., Burger, M., Goldfarb, D., Xu, J., & Yin, W. (2005). An iterative regularization method for total variation-based image restoration. SIAM Journal on Multiscale Modeling and Simulation, 4(2), 460–489.

Pardalos, P. M. (1997). Optimization techniques for portfolio selection. In C. Zopounidis (Ed.), New operational approaches for financial modelling (pp. 19–33). Heidelberg: Physica-Verlag HD.

Pardalos, P. M., Sandström, M., & Zopounidis, C. (1994). On the use of optimization models for portfolio selection: A review and some computational results. Computational Economics, 7(4), 227–244.

Wang, J., & Forsyth, P. A. (2012). Comparison of mean variance like strategies for optimal asset allocation problems. International Journal of Theoretical and Applied Finance (IJTAF), 15(02), 1–32.

Yen, Y., & Yen, T. J. (2014). Solving norm constrained portfolio optimization via coordinate-wise descent algorithms. Computational Statistics and Data Analysis, 76, 737–759.

Yin, W., Osher, S., Goldfarb, D., & Darbon, J. (2008). Bregman iterative algorithms for l1-minimization with applications to compressed sensing. SIAM Journal of Imaging Science, 1(1), 143–168.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.