Nonlinear Integer Programming and its Application in Portfolio Selection

Duan Li
The Chinese University of Hong Kong

Xiaoling Sun
Shanghai University

Jun Wang
Qingdao University
1 Introduction

• Portfolio selection is to seek a best allocation of wealth among a basket of securities.

• The mean-variance formulation proposed by Markowitz (1959, 1989) provides a fundamental basis for portfolio selection.

• Analytical expression of the mean-variance efficient frontier in single-period portfolio selection was derived in Markowitz (1956) and Merton (1972).

• The Markowitz’s mean-variance model was extended to multi-period portfolio selection problems in Li and Ng (2000) and to continuous-time portfolio selection problems in Zhou and Li (2000).
Need & Significance of the Research

• Although the real trade practice possesses certain discrete features, such as cardinality, round-lots, buy-in threshold, the discontinuity of transaction costs at the origin and dependency purchasing constraints among risky assets, the current literature on portfolio selection has been primarily developed for the continuous solution which could be far away from the real integer optimum.
Most of the few suggested solution methodologies in the literature that tackle discrete features in portfolio selection are heuristic in nature, for example,

- Heuristics for cardinality constrained mean-variance model by Chang et al. (2000),
- A simulated annealing method for cardinality constrained mean-variance formulation with side constraints by Crama and Schyns (2003),
- Heuristics for the mean-variance model with cardinality and round lots constraints by Jobst et al. (2001),
- Heuristics for the mean-absolute deviation formulation with round-lots by Mansini and Speranza (1999), and
- heuristics for the mean-semi-absolute deviation formulation with round-lots by Kellerer et al. (2000).
• Exact solution methodologies include Bienstock (1996) for the cardinality constrained mean-variance model with side constraints and Syam (1998) for the mean-variance model with dependency relationships and round lots.
  
  – Bienstock (1996) suggests a relaxation for the cardinality constraint and uses branch and cut as a solution scheme.
  
  – Syam (1998) assumes independence among risky securities, which leads to a diagonal covariance matrix, and then adopts a branch and bound solution method.
• There is a need to extend the state-of-the-art of the portfolio selection theory in order to develop exact algorithms for optimal integer solutions to discrete-feature constrained portfolio selection.

• In this research, we consider optimal lot solution to the cardinality constrained mean-variance formulation for portfolio selection under concave transaction costs.
2 Problem Formulation

- We consider a market with $n$ available securities. The purchasing of the securities is confined to integer number of lots.

- An investor with initial wealth $W_0$ seeks to improve his wealth status by investing his wealth into these $n$ risky securities and into a riskless asset (e.g., a bank account).

- Since the trade practice often only allows trade of integer lots of stocks, one discrete feature in our model is to confine the decisions to be multiples of round lots.
Let $X_i$ be the random return per lot of the $i$th security ($i = 1, \ldots, n$) before deducting associated transaction costs. The mean and covariance of the returns are assumed to be known,

$$\mu_i = E(X_i), \text{ and}$$

$$\sigma_{ij} = \text{Cov}(X_i, X_j), \quad i, j = 1, \ldots, n.$$  

Let $x_i$ be the integer number of lots the investor invests in the $i$th security. We assume that no shorting is allowed for any security. Thus $x_i$ is a nonnegative integer. Denote the decision vector in portfolio selection by $x = (x_1, \ldots, x_n)^T$. Then, the random return from holding securities is $P_s(x) = \sum_{i=1}^{n} x_i X_i$. 
• The mean and variance of $P_s(x)$ are

$$s(x) = E[P_s(x)] = E\left[\sum_{i=1}^{n} x_i X_i \right] = \sum_{i=1}^{n} \mu_i x_i$$

and

$$V(x) = \text{Var}(P_s(x)) = \text{Var}\left[\sum_{i=1}^{n} x_i X_i \right] = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j \sigma_{ij} = x^T C x,$$

where $C = (\sigma_{ij})_{n \times n}$ is the covariance matrix.

• Let $r$ be the interest rate of the riskless asset. We assume that the same rate is applied when borrowing money from the riskless asset.
• Let $b_i$ be the current price of one lot of the $i$th security.

• The balance $x_0 = W_0 - \sum_{i=1}^{n} b_i x_i$ is assumed to be deposited into the riskless asset and $rx_0$ is the corresponding return. Note that a negative $x_0$ implies a debt from the riskless asset.

• The budget constraint of the investor is given by

$$b^T x \leq W_0 + U_b$$

where $b = (b_1, \ldots, b_n)^T$ and $U_b$ is an upper borrowing limit from the riskless asset.

• Let $c(x) = \sum_{i=1}^{n} c_i(x_i)$ be the transaction cost associated with the portfolio decision $x = (x_1, \ldots, x_n)^T$. It is assumed in this paper that each $c_i(\cdot)$ is a nondecreasing concave function.
• The justification of the concavity assumption of transaction costs can be found in the literature, for example in Konno and Wijayanayake (2001).

• The total expected return of portfolio decision \( x \) can be now summarized as:

\[
R(x) = s(x) + rx_0 - \sum_{i=1}^{n} c_i(x_i) = \sum_{i=1}^{n} [(\mu_i - rb_i)x_i - c_i(x_i)] + rW_0.
\]

• We assume that \( c_i(0) = 0 \) and \( (\mu_i - rb_i)x_i \geq c_i(x_i) \) for all positive \( x_i \), \( i = 1, \ldots, n \). The above assumption and the concavity of all \( c_i \)'s guarantee that \( R(x) \) is a nondecreasing function with respect to each \( x_i \). Note that \( R(x) \) is a convex function since each \( c_i(x_i) \) is a concave function.
• Since it is always true in real world that investors do not like a portfolio which is too widely-spread, the second important discrete feature in our model is the cardinality constraint.

• We consider further a cardinality constraint in portfolio selection, $\text{supp}(x) \leq K$, where $\text{supp}(x)$ denotes the number of nonzero components in $x$ and $K$ is a given positive integer with $K \leq n$. 
Discrete-feature constrained mean-variance model:

\[(P) \quad \text{minimize} \quad V(x) = x^T C x\]

\[\text{s.t.} \quad R(x) = \sum_{i=1}^{n} [(\mu_i - rb_i) x_i - c_i(x_i)] + rW_0 \geq \varepsilon,\]

\[U(x) = b^T x \leq W_0 + U_b,\]

\[\text{supp}(x) \leq K,\]

\[x \in X = \{x \in \mathbb{Z}^n \mid 0 \leq x_i \leq u_i, \ i = 1, 2, \ldots, n\},\]

where \(u_i\) is an upper bound on the purchasing of the \(i\)th security which is either imposed by the investor or can be set as the largest integer number less than or equal to \(\frac{W_0 + U_b}{b_i}\).
• The resulting model under our consideration is a nonseparable, nonconvex nonlinear integer programming problem. No exact solution method proposed in the literature is applicable to tackle the discrete-feature constrained portfolio selection problem considered in this paper.

• We point out that the nonconvexity in our model prevents a direct application of conventional branch-and-bound solution schemes since finding a bound now becomes a global optimization problem.
• In this research we propose a convergent Lagrangian method as an exact solution scheme for the above discrete-feature constrained portfolio selection model.

• The Lagrangian dual search method has been one of the powerful techniques in developing efficient algorithms for integer programming problems.

• While the Lagrangian relaxation leads to a powerful decomposition only when the primal problem is separable, the mean-variance formulation is, by its nature, a nonseparable problem due to the correlation among different risky securities reflected in a nondiagonal covariance matrix.
• The classical Lagrangian method is not always able to find an exact optimal solution to the primal integer problem due to the frequent existence of a duality gap (see Fisher (1981), Geoffrion (1974) and Shapiro (1979)).

• To achieve a strong duality, nonlinear Lagrangian methods were proposed (Li and White (2000), Li and Sun (2000), Sun and Li (2000)) (but not computationally implementable).
Recent prominent progress in convergent Lagrangian methods for nonlinear separable integer programming (Li and Sun (2005)) – To achieve strong duality with computational efficiency

- Objective level cut method for nonlinear separable integer programming
- Domain cut method for nonlinear integer knapsack problems
- Contour cut method for quadratic separable integer programming
- This paper exploits special features of the mean-variance formulation and develops a convergent Lagrangian and contour-domain cut method, thus leading to an efficient exact solution algorithm to identify an optimal lot solution to the cardinality constrained mean-variance formulation for portfolio selection.
HOW TO ACHIEVE A SEPARABILITY?

- The primal mean-variance problem \((P)\) is of a nonseparable structure as the objective function \(V(x) = x^T C x\) is non-separable.

- This nonseparability prevents a direct use of the Lagrangian dual method from being an efficient lower bounding procedure for \((P)\). To overcome this difficulty, we approximate \(V(x)\) by a separable quadratic function.
Consider the following auxiliary problem:

\[(P_A) \quad \text{minimize} \quad V_A(x) = \lambda_{\text{min}} x^T x \]

\[\text{s.t.} \quad R(x) = \sum_{i=1}^{n} [(\mu_i - r b_i) x_i - c_i(x_i)] + r W_0 \geq \varepsilon, \]

\[U(x) = b^T x \leq W_0 + U_b, \]

\[\text{supp}(x) \leq K, \]

\[x \in X = \{x \in \mathbb{Z}^n \mid 0 \leq x_i \leq u_i, \ i = 1, 2, \ldots, n\}, \]

where \(\lambda_{\text{min}}\) is the minimum eigenvalue of \(C\).
The primal problem \((P)\) and the auxiliary problem \((P_A)\) only differ in their objectives. Let \(v(Q)\) be the optimal value of problem \(Q\). It is easy to see that \(v(P_A) \leq v(P)\). In other words, the auxiliary formulation provides a lower bound for the primal problem,

\[
\min_{x \in S} \lambda_{\min} x^T x \leq \min_{x \in S} x^T Cx.
\]

Since both \(V_A(x)\) and \(R(x)\) are increasing functions with respect to each \(x_i\), the optimal solution of \((P_A)\) should locate as close to the boundary of constraint \(R(x) \geq \varepsilon\) from above as possible.
• Dualizing constraint \( R(x) \geq \varepsilon \) and omitting constraints 
\( U(x) \leq W_0 + U_b \) and \( \text{supp}(x) \leq K \) give rise to the following Lagrangian relaxation:

\[
d(\lambda) = \min_{x \in X} L(x, \lambda)
\]

\[
:= \lambda_{\min} x^T x - \lambda \left\{ \sum_{i=1}^{n} \left[ (\mu_i - r b_i) x_i - c_i(x_i) \right] + r W_0 - \varepsilon \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ \min_{x_i \in X_i} [\lambda_{\min} x_i^2 - \lambda (\mu_i - r b_i) x_i + \lambda c_i(x_i)] \right\}
\]

\[
+ \lambda (\varepsilon - r W_0),
\]

(1)

where \( X_i = \{x_i \mid 0 \leq x_i \leq u_i, x_i \text{ integer}\} \).
• The corresponding dual problem is

$$(D) \quad \max_{\lambda \geq 0} d(\lambda).$$

• We have the following relation:

$$v(D) \leq v(P_A) \leq v(P).$$

Thus, the optimal dual value $v(D)$ provides a lower bound for problem $(P)$. The dual problem $(D)$ can be solved by some efficient numerical procedure.
How to eliminating duality gap?

- Denote a hyper-rectangle in $\mathbb{R}^n$ and in $\mathbb{Z}^n$ by

  $$[\alpha, \beta] = \{x \mid \alpha_i \leq x_i \leq \beta_i, i = 1, \ldots, n\}, \quad \alpha, \beta \in \mathbb{R}^n,$$

  $$\langle \alpha, \beta \rangle = \prod_{i=1}^{n} \langle \alpha_i, \beta_i \rangle = \langle \alpha_1, \beta_1 \rangle \times \langle \alpha_2, \beta_2 \rangle \cdots \times \langle \alpha_n, \beta_n \rangle,$$

  $$\alpha, \beta \in \mathbb{Z}^n.$$

- The set $\langle \alpha, \beta \rangle$ is called an integer box (subbox).

- For $x \in \mathbb{R}$, $\lceil x \rceil$ denotes the minimum integer that is greater than or equal to $x$ and $\lfloor x \rfloor$ denotes the maximum integer that is smaller than or equal to $x$. 

0-23
The contour of $x^T C x = v$, where $v$ is a constant, is an ellipse. We denote by $E(v)$ the ellipse defined by $x^T C x = v$. 

Figure 1: Ellipse contour and associated hyper-rectangle boxes.
Let $\lambda^*$ be the optimal solution to the dual problem, and $x^*$ and $y^*$ be two optimal solutions to the corresponding Lagrangian relaxation problem, one being feasible to the Lagrangian relaxation problem and one being infeasible.

- If $\lambda^* \neq 0$, then the integer box $B^I_v = \langle -\lfloor a^I \rfloor, \lceil a^I \rceil \rangle \cap \langle \alpha, \beta \rangle$ contains no feasible solution to $(P)$ based on the weak duality, where $a^I$ can be calculated analytically from the dual value.

- If $x^*$ is feasible to $(P)$, then there is no feasible solution better than $x^*$ outside $B^O_v \cap \langle \alpha, \beta \rangle$, where $a^O$ can be calculated analytically.

- Moreover, there is no feasible solution better than $x^*$ in the integer subbox $B_{x^*} = \langle a^M, b^M \rangle$, where $a^M$ and $b^M$ are on the boundary of $\langle \alpha, \beta \rangle$. 
• If $x^*$ is infeasible to $(P)$, either violating the budget constraint or the cardinality constraint, then $\langle x^*, \beta \rangle$ contains no feasible solution to $(P)$ based on the monotonicity.

• The integer subbox $\langle \alpha, y^* \rangle$ contains no feasible solution to $(P)$.

• A key component method is to fathom those non-promising integer subboxes in the solution process, thus resulting in non-rectangular integer domain.
Figure 2: Partition of $A \setminus B$. 
Lemma 1 Let $A = \langle \alpha, \beta \rangle$ and $B = \langle \gamma, \delta \rangle$, where $\alpha, \beta, \gamma, \delta \in \mathbb{Z}^n$ and $\alpha \leq \gamma \leq \delta \leq \beta$. Then $A \setminus B$ can be partitioned into at most $2n$ integer boxes.

\[ A \setminus B = \left\{ \bigcup_{j=1}^{n} \left( \Pi_{i=1}^{j-1} \langle \alpha_i, \delta_i \rangle \times \langle \delta_j + 1, \beta_j \rangle \times \Pi_{i=j+1}^{n} \langle \alpha_i, \beta_i \rangle \right) \right\} \]

• Suppose that $\tilde{X}$ can be expressed as a union of integer subboxes $\tilde{X} = \bigcup_{i=1}^{k} \tilde{X}^i$. Applying the dual search procedure separately on each integer subbox yields a better lower bound than performing a dual search on $\tilde{X}$ as a whole.
• The proposed convergent Lagrangian method: procedure.
  
  – Obtain initially a lower bound \( d(\lambda^0) \), a solution \( x^0 \) with \( R(x^0) \geq \varepsilon \) and a solution \( y^0 \) with \( R(y^0) < \varepsilon \), by solving a dual problem.
  
  – At each iteration, the algorithm generates a set of new integer subboxes by using a cutting procedure and the partition formula.
  
  – Apply the dual search procedure to each newly generated integer subbox to obtain a lower bound and two optimal solutions to the corresponding Lagrangian relaxation.
  
  – Fathom a new integer subbox if its dual value is greater than or equal to the incumbent value.
  
  – The algorithm terminates when the set of unsolved integer subboxes is empty.
3 Computational Experiment

• We consider 30 blue chips in Hang Seng Index.
  – The monthly return for each stock is calculated by the difference between the price per lot at the end of the month and the price per lot at the beginning of the month.

• The covariance matrix \( C \) is derived from the data of the monthly returns, while the expected return vector \( \mu = (\mu_1, \ldots, \mu_n)^T \) is generated by making an adjustment on the average monthly return from the three years’ data.

• The current price vector \( b = (b_1, \ldots, b_n)^T \) is set at the price on January 3, 2005.
• In our experiment, the interest rate of the riskless asset, \( r \), is set at 1%.

• The concave transaction cost function is of a form \( \alpha_i \ln(1 + x_i) \), while for different stocks, \( \alpha_i \) is randomly chosen from \([0.001, 0.004]\).

• We set the initial wealth at \( W_0 = 1,000,000 \), the upper borrowing limit \( U_b \) at \$4,000,000 and the maximum purchasing lot of each stock, \( u_i \), at 10, i.e., \( u = (10, \ldots, 10)^T \).

• The minimum return level \( \varepsilon \) varies according to \( \text{ratio} \times R(u) \), where \( \text{ratio} \) ranges from 0.70 to 0.90.

• Algorithm 1 was programmed by Fortran 90 and run on a Pentium IV PC with RAM of 256M.
Figure 3: The efficient frontiers of integer portfolio selection with \( n = K = 20, 25, 30 \).

- The efficient frontier with a larger \( n \) dominates the one with a smaller \( n \).
Figure 4: The efficient frontiers of integer portfolio selection with $n = 20$ and $K = 15, 17$.

- The larger the $K$, the more choices and the better outcome in the mean-variance space.
Table 1: Numerical results for problems with $n = K = 20, 25, 30$

<table>
<thead>
<tr>
<th>Ratio</th>
<th>Number of Iterations</th>
<th>Number of Subboxes</th>
<th>CPU Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n = 20$</td>
<td>$n = 25$</td>
<td>$n = 30$</td>
</tr>
<tr>
<td>0.70</td>
<td>276</td>
<td>66</td>
<td>809</td>
</tr>
<tr>
<td>0.73</td>
<td>138</td>
<td>265</td>
<td>135</td>
</tr>
<tr>
<td>0.75</td>
<td>60</td>
<td>252</td>
<td>179</td>
</tr>
<tr>
<td>0.77</td>
<td>51</td>
<td>177</td>
<td>284</td>
</tr>
<tr>
<td>0.80</td>
<td>42</td>
<td>81</td>
<td>166</td>
</tr>
<tr>
<td>0.83</td>
<td>96</td>
<td>50</td>
<td>83</td>
</tr>
<tr>
<td>0.85</td>
<td>48</td>
<td>45</td>
<td>65</td>
</tr>
<tr>
<td>0.87</td>
<td>55</td>
<td>66</td>
<td>90</td>
</tr>
<tr>
<td>0.90</td>
<td>42</td>
<td>50</td>
<td>60</td>
</tr>
</tbody>
</table>
Table 1 demonstrates a tendency that the number of iterations, the number of subboxes and the CPU time decrease as the minimum return level \( \varepsilon \) increases. This could be due to the fact that when \( \varepsilon \) increases, the feasible region decreases, thus a fast reduction of the duality gap and a rapid convergence of the algorithm.
Table 2: Numerical results for problems with $K = 17$ and $K = 15$ ($n = 20$)

| Ratio | Number of Iterations | | Number of Subboxes | | CPU Seconds |
|---|---|---|---|---|
| 0.80 | $K = 17$ | $K = 15$ | $K = 17$ | $K = 15$ | $K = 17$ | $K = 15$ |
| 0.83 | 40 | 79 | 471 | 771 | 44.1 | 66.9 |
| 0.84 | 62 | 49 | 616 | 576 | 38.7 | 45.5 |
| 0.85 | 41 | 240 | 583 | 1890 | 41.3 | 113.6 |
| 0.86 | 35 | 1191 | 328 | 6797 | 21.2 | 602.6 |
| 0.87 | 27 | 2091 | 279 | 11617 | 18.6 | 857.2 |
| 0.88 | 40 | 2993 | 405 | 16146 | 28.8 | 1575.8 |
| 0.88 | 54 | 4141 | 559 | 25383 | 37.8 | 2279.6 |
| 0.90 | 592 | 3940 | 5299 | 21144 | 457.6 | 1971.4 |
| 0.95 | 2190 | 5120 | 14231 | 40018 | 845.5 | 2918.5 |
• From Table 2, we observe that the problem with $K = 15$ is much harder than the one with $K = 17$. This is understandable because of finding a solution to a cardinality constrained problem is to implicitly search $\binom{n-K}{K} = \frac{n!}{K!(n-K)!}$ $K$-dimensional subspaces.

• Furthermore, we notice in Table 2 an interesting phenomenon which is opposite to what we conclude for situations without a cardinality constraint. For a fixed $K$, the number of iterations, the number of subboxes and the CPU time all increase when the minimum return level $\varepsilon$ increases. A possible reason could be that an increase of the value $\varepsilon$ in the constraint $R(x) \geq \varepsilon$ conflicts a satisfaction of the cardinality constraint $\text{supp}(x) \leq K$, thus making the search of a better feasible point more difficult.
4 Conclusions

- We have proposed in this paper an exact solution algorithm – the convergent Lagrangian and contour-domain cut method for obtaining an optimal lot solution to the cardinality constrained mean-variance formulation for portfolio selection problems under concave transaction costs.

- The inapplicability of the conventional Lagrangian dual method in providing a lower bound caused by an inherent non-separability in the mean-variance formulation is alleviated by constructing an auxiliary separable problem.
• Furthermore, this algorithm takes advantages of special geometric structures of the quadratic contour and the monotone property of the expected return function, the budget function and the cardinality constraint.

• A contour-domain cutting procedure is devised accordingly in the algorithm to cut non-promising integer sub-boxes from further consideration, thus reducing the duality gap of the Lagrangian dual search consecutively in an iteration process. The convergence of the Lagrangian dual method is assured due to a continuous shrinking of the solution domain. The efficiency of the proposed algorithm has been witnessed in numerical testing.
• One future work is to extend our current solution framework from the mean-variance formulation to general mean-risk formulations by considering alternative risk measures.

• Construction of a suitable risk measure plays an essential role in portfolio selection. After the ice-breaking quantitative risk measure using the variance, many other risk measures have been proposed in the literature to capture the essence of the underlying risk in portfolio selection.
• Notice that neither non-linearity nor non-convexity in general risk measures generates obstacles to block an application of our solution concepts. The real difficulty comes from non-separability which is inherent in risk measures due to correlations among risky securities. Suitable decomposition schemes need to be identified for different risk measures.

• Adopting a “separable” structure in some non-variance risk measures (such as the absolute derivation) may lead to a more efficient convergent Lagrangian method for solving large-scale portfolio selection problems.
Another extension is to consider dependency purchasing constraints among risky assets in the model. More specifically, the solution framework developed in this paper can be readily extended to solve the following model:

\[
\begin{align*}
\text{minimize} & \quad V(x) = x^T C x \\
\text{s.t.} & \quad R(x) = \sum_{i=1}^{n} [(\mu_i - r b_i) x_i - c_i(x_i)] + r W_0 \geq \varepsilon, \\
& \quad A x \leq \gamma \\
& \quad \text{supp}(x) \leq K, \\
& \quad x \in X = \{x \in \mathbb{Z}^n \mid 0 \leq x_i \leq u_i, \ i = 1, 2, \ldots, n\},
\end{align*}
\]

where \(A\) is an \(m \times n\) constant matrix and \(\gamma\) is a given \(m\)-dimensional vector.
• This newly added $m$-dimensional constraint can include not only the budget constraint existed in our current model, but also some other dependency purchasing constraints among risky assets, for example,

$$\sum_{i \in I_1} b_i x_i \leq \sum_{i \in I_2} b_i x_i.$$ 

• There exist efficient dual search methods, for example, the outer approximation method (see Parker and Rardin (1988) and Li and Sun (2005)), to deal with multiply constrained situations. Extra efforts need, however, to be devoted to explore valid cutting schemes to handle constraints which are not uniformly monotone (increasing with respect to some variables while decreasing with respect to other variables).
References


D. Li and W. L. Ng. Optimal dynamic portfolio selection: Multi-


