Selecting a discrete portfolio

Wojciech Olszewski*and Rakesh Vohra†

July 12, 2014

Abstract

We study the problem of selecting an optimal portfolio out of a finite set of available assets. Assets are characterized by their expected returns and the covariance matrix, and investors are assumed to have a mean-variance utility, that is, their utility function is linear in the mean and variance of the portfolio they hold.

When assets are negatively correlated, or even when a slightly more general condition is satisfied, we provide an algorithm for selecting an optimal portfolio. We illustrate the usefulness of this algorithm by some comparative statics result. When assets can be positively correlated, we deliver a negative result regarding the existence of useful algorithms for selecting an optimal portfolio.

1 Introduction

Direct investment decisions take the form of choosing a portfolio out of a finite set of available assets. This basic problem in economics and finance was first studied by Reiter (1963), who suggested some simple algorithms for selecting portfolios, arguing by means of numerical examples that the values of portfolios selected by those algorithms are close to the values of optimal portfolios. While the algorithms suggested by Reiter are simple and perform well in some numerical examples, they do not in general select optimal portfolios, even portfolios whose values are close to the optimal ones. In the present paper, we are concerned with the existence of algorithms choosing an optimal portfolio in the general case.

The problem, as every discrete optimization problem, can in principle be solved by complete enumeration of all portfolios, but this would not be useful. We approach the problem of searching for optimal portfolios rigorously, adopting from discrete optimization the solution concept of *polynomial-time algorithm.* Such algorithms evaluate a number of portfolios that is polynomial in the number of available

*Department of Economics, Northwestern University, 2001 Sheridan Rd. Evanston IL 60208-2600

†Department of Economics and Department of Electrical & Systems Engineering, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104

assets, in contrast to complete enumeration, which requires the evaluation of a number of portfolios exponential in the number of assets.[1]

The choice of solution concept should be understood as a rigorous formulation of finding a meaningful solution, rather than the suggestion of approaching the problem numerically. Indeed, polynomial-time algorithms for solving discrete optimization problems are typically fairly simple, and have numerous useful features. The most striking examples of useful polynomial-time algorithms in economics are probably the algorithms used in the literature on matching, such as the celebrated Gale and Shapley (1962) algorithm. Other examples include recent research by Chade and Smith (2006) and Milgrom and Segal (2014), who suggested using greedy algorithms in some search and auction settings, respectively.

We provide a polynomial-time algorithm, which we call the *financing algorithm*, in the case when the returns of every pair of assets are negatively correlated. This algorithm is an instance of the Ford and Fulkerson (1956) algorithm for solving the project selection problems. Our contribution is therefore in the recognition that the portfolio selection is a special case of the project selection problem and in providing an algorithm for it with natural economic interpretation.[2]

We also show that the problem of finding an optimal portfolio for arbitrary correlations is NP-hard. Finally, our analysis of the optimal portfolio problem allows for comparative statics, i.e., for studying the response of optimal portfolio to changes in parameters of the problem. As an example of this kind of analysis, we will argue that the composition of the optimal portfolio is less sensitive to the changes in the cost of investing into assets which are acquired to reduce the variance of the portfolio than to the changes in the cost of investing into assets which are acquired for their high expected returns.

The rest of the paper is organized as follows. In Section 2, we describe the model. We briefly review discrete optimization concepts in Section 3. That section also contains a description of the project selection problems, and the algorithms for solving such problems. We discuss there some earlier related results, and argue that polynomial-time algorithms for portfolio selection exist in a somewhat more general case than for negatively correlated returns. In Section 4, we present our algorithm for selecting an optimal portfolio when assets are negatively correlated. Section 5 contains a comparative statics exercise, and Section 6 a negative result for the case when some correlations are positive. We conclude in Section 7.

2 Model

An investor assembles a portfolio from a set of n assets with random, possibly correlated returns. Let $N = \{1, ..., n\}$ be the grand set of assets. Let μ_i be the expected return of asset i, and $(\sigma_{ij})_{i,j=1}^{n}$ be the

[1] The idea is somewhat similar to first- and second-order conditions in calculus. Every optimization problem can in principle be solved by comparing all values of the objective function. However, this method is typically not tractable or useful, in contrast to first- and second-order conditions.

[2] Note, in addition, that we somewhat simplified the original Ford and Fulkerson algorithm, which was possible due to the specific structure of the portfolio selection problem.

covariance matrix of the available assets. By definition, $\sigma_i^2 = \sigma_{ii}$. It costs c_i to invest in asset i.

The investor has a mean-variance utility function, as is often assumed in finance, which is a linear function $a\mu - b\sigma^2$ of the mean μ and variance σ^2 of a portfolio. The coefficients a and b are positive. Thus, the investor's utility of a portfolio $S \subset N$ is

$$U(S) = a\left(\sum_{i\in S}\mu_i - \sum_{i\in S}c_i\right) - b\sum_{i,j\in S}\sigma_{ij}. \tag{1}$$

It will be convenient to define asset i's *individual (net) return* as

$$x_i = a\left(\mu_i - c_i\right) - b\sigma_{ii},$$

asset i's *hedging value* for asset $j \neq i$ as $\alpha_{ij} = -b\sigma_{ij}$, and asset i's *hedging value* for portfolio S as

$$\alpha_{iS} = -\sum_{i\neq j\in S} b\sigma_{ij}.$$

The investor's objective is to assemble a portfolio S that maximizes the utility $U(S)$. In order to illustrate the concepts, we conclude this section with a simple example of investment problem, in which an optimal portfolio will be found by inspection.

Example 1 *Let:* $N = 5$, $a = b = 1$, $\mu_1 = 10$, $\mu_2 = \mu_3 = \mu_4 = \mu_5 = 1$, *and the covariance matrix* $(\sigma_{ij})_{i,j=1}^n$ *be given by*

$$\begin{bmatrix} 5 & -1 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, let $c_1 = 1$, $c_2 = 1/2$, $c_3 = 3/2$, $c_4 = 5/2$, *and* $c_5 = 7/2$.

Then the optimal portfolio must contain asset 1, *since its individual return* $x_1 = 4$ *is positive. The individual returns of other assets are negative, since* $1\left(1 - c_i\right) - 1 = -c_i$. *However, it pays back to include assets* 2 *and* 3 *together with asset* 1, *since their hedging value for asset* 1 *is* 1, *and the hedging value of asset* 1 *for each of them is also* 1. *This hedging value of* 2 *exceeds the cost* c_i *for* $i = 2$ *and* 3 *but not for* $i = 4$ *and* 5. *Thus, the optimal portfolio has the form of an asset generating high net returns, and two assets included for the purpose of hedging.*

In Example 1, finding an optimal portfolio was easy due to the simple structure of means, costs and covariance matrix. However, if the structure of these parameters is more complicated, finding an optimal portfolio by inspection may not be possible. So, one would like to have a more systematic method of searching for an optimal portfolio. We will provide such a method in the Section 4.

3 A review of related discrete optimization results

The problem of selecting a discrete portfolio is a special case of what might be called project or activity selection problems. Such problems are interesting only when the projects or activities interact with each other. There are three basic ways in which these interactions can take place. The first is when different sets of projects compete for common resources. Framed this way, the project selection problem is no different from a generic integer program. Positive results can be obtained only if the underlying constraint matrix has suitable structure.

The second is when the a project becomes feasible or available provided some other subset of projects has already been selected. Thus, associated with each project i is a set $P(i)$ of predecessors all of which must be selected before project i becomes available. The classic paper on this subject is Rhys (1970). The main result in that paper was to show that this version of the problem could be solved as a linear program. In other words, the linear programming relaxation of the natural integer programming formulation of the problem had integer optima. Subsequent authors showed that this linear program could be represented as a maximum flow problem and solved more rapidly using special purpose algorithms. Hochbaum (2004) surveys a range of ostensibly different problems that can be represented as instances of this version of the project selection problem. All these problems can be solved using algorithms for finding maximum flows in networks.

The third is when there are complementarities or correlations in the returns (or costs) of the projects. In a sense the sequencing version just described can be seen as an extreme case of this. The returns on project i are zero unless all projects in $P(i)$ are completed in which case they become positive. The problem we consider falls in this third class. When all assets are negatively correlated the problem of finding a portfolio that maximizes the payoff function given by (1) is a special case of the problem considered by Rhys (1970), i.e., a maximum flow problem. Our algorithm for selecting an optimal portfolio is a simplified instance of the Ford-Fulkerson (1956) algorithm for maximum flow that has a natural economic interpretation.

This financing algorithm as described will not run in polynomial time. One needs to define the search order for a refinancing opportunity by adapting the *breadth-first search algorithm*. With this modification our algorithm is guaranteed to run in polynomial time. This version of the Ford-Fulkerson algorithm was proposed independently by Dinic (1970), and Edmonds and Karp (1972). Because we are interested in a method of selecting an optimal portfolio, rather than its computational properties, we omit the discussion of the breadth-first search algorithm.

The assumption that all assets are negatively correlated can be relaxed slightly to the case when the covariance matrix of the assets is sign balanced. To define the notion of sign-balance, construct a graph G that has an edge between i and j if and only if $\sigma_{ij} \neq 0$. The edges which correspond to positive (resp. negative) σ_{ij} are called positive edges (resp. negative edges). This graph G is sign-balanced if it does not contain any cycle with an odd number of positive edges. To this end, one can use the algorithm from Hansen and Simeone (1986).

4 Financing algorithm for selecting optimal portfolio

The *financing algorithm* assembles a portfolio as follows:

Step 0. All assets i with positive individual returns $x_i \geq 0$ are included to the portfolio. Each asset such that $x_i < 0$ is assigned a number $d_i = -x_i$, which will be called i's *deficit*; if $x_i \geq 0$, set $d_i = 0$. If an investor decides to acquire asset i, this asset alone will generate a negative return of $-x_i$. However, it may pay to acquire asset i together with other assets for the purpose of hedging. Each pair of assets i and j is assigned a number $s_{ij} = \alpha_{ij} + \alpha_{ji}$. (By ij, we will denote the set $\{i, j\}$, that is, we assume that $ij = ji$.) This is an increment in the utility of the investor who acquires the two assets together; she obtains this increment on top of the two assets' individual returns. This increment is the result of each asset providing some hedging against the risk imposed by the other asset, and it will be called ij's *(mutual) surplus*.

While describing our algorithm, we will illustrate its particular steps by the following example, which is depicted in Figure 1.

Example 2 *Let $N = \{1, 2, 3\}$, $a = b = 1$, $\mu_1 = 4$, $\mu_2 = 1$, $\mu_3 = 0$, and the covariance matrix $(\sigma_{ij})_{i,j=1}^n$ be given by*

$$\begin{bmatrix} 2 & -1/2 & -1/2 \\ -1/2 & 2 & -1 \\ -1/2 & -1 & 2 \end{bmatrix}.$$

Finally, let $c_1 = c_2 = c_3 = 1$.

In Example 2, we have $x_1 = 1$, $d_2 = -x_2 = 2$, $d_3 = -x_3 = 3$, $s_{12} = 1$, $s_{23} = 2$, and $s_{13} = 1$. That is, in step 0, asset 1 (and only asset 1) is included to the portfolio. This is depicted in the upper diagram of Figure 1.

Consecutive steps (1, 2, etc.), financing opportunity. In each consecutive step of our algorithm, we first check for what we call a "financing opportunity", i.e., whether there are pairs i and j with $x_i < 0$ and $s_{ij} > 0$. If we find such i and j,[3] we "transfer" the surplus from ij to reduce the deficit of i up to the maximum possible amount; this amount is simply the larger of the two numbers: $-x_i$ and $\alpha_{ij} + \alpha_{ji}$. More precisely, both d_i and s_{ij} take new values. If $s_{ij} \leq d_i$, then d_i becomes $d_i - s_{ij}$, and s_{ij} becomes 0; and if $s_{ij} > d_i$, then s_{ij} becomes $s_{ij} - d_i$, and d_i becomes 0.

So, due to the transfers, deficits are reduced or "financed". It will be important that the algorithm remembers from which surpluses each deficit was financed.

[3]For now, take an arbitrary pair with the required property if multiplicity arises. We will return at the end of our description of financing algorithm to the issue which pair to choose if multiplicity arises.

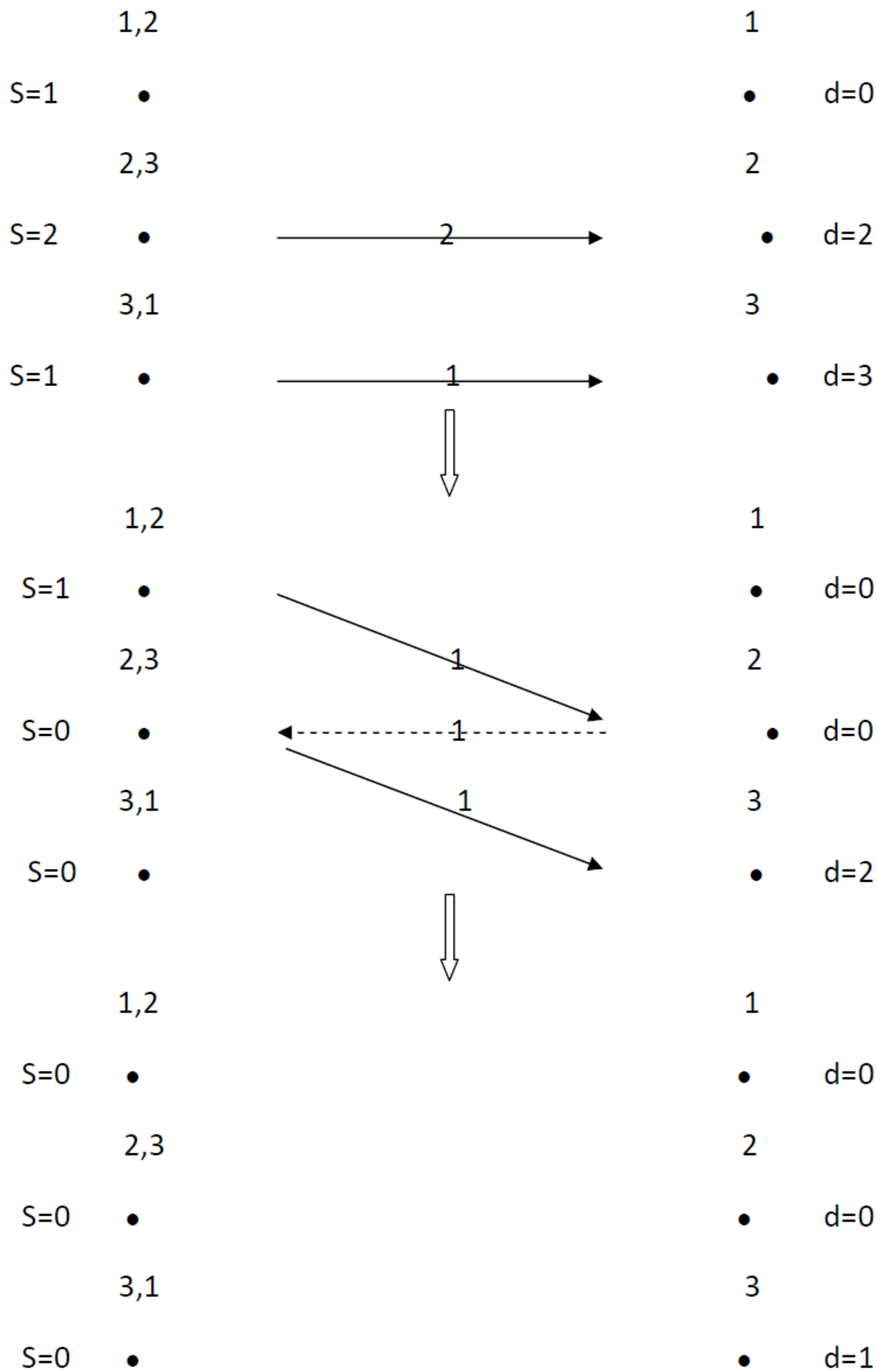

Figure 1. Pairs of assets are represented by dots on the left-hand side, and individual assets are represented by dots on the right-hand side, with the corresponding surpluses and deficits marked next to the dots. Arrows indicate transfers with their size marked on the arrows.

In Example 2, pair 23 has the property that $x_2 < 0$ and $s_{23} > 0$, so there is a financing opportunity. In step 1, both values d_2 and s_{23} become 0. This is depicted in the upper diagram of Figure 1 by the arrow from 23 to 2. In Step 2, we first check for a financing opportunity, exactly as in step 1. In Example 2, there is another financing opportunity, depicted by the arrow from 13 to 3 in Figure 1. The new values of surpluses and deficits obtained by making the two transfers (form 23 to 2 and from 13 to 3) are marked next to the dots in the middle diagram of Figure 1. As we have said, the algorithm remembers that 2's deficit was financed from 23's surplus, and that 3's deficit was financed from 13's surplus.

Consecutive steps (1, 2, etc.), the opportunity of refinancing. If we exhaust all financing opportunities, as now happens in Example 2, then we check for what we will call an "opportunity of refinancing". That is, we check whether there exists i and j, and a sequence of assets $i, k_0, k_1, ..., , k_m, j$ such that $s_{ik_0} > 0$, $d_j > 0$, $d_i = 0$ and $d_{k_l} = 0$ for $l = 1, ..., m$, $s_{k_m j} = 0$, $s_{ik_1} = 0$ and $s_{k_l k_{l+1}} = 0$ for $l = 1, ..., m-1$; moreover, we check if the original deficit of i is (entirely or partially) financed from ik_1's surplus, the original deficit of k_l is financed from $k_l k_{l+1}$'s surplus for $l = 1, ..., m-1$, and the original deficit of k_m is financed from $k_m j$'s surplus.

If such a refinancing opportunity arises, we transfer surplus from i back to ik_1, replacing it with a the surplus returned to i from ik_0; we transfer surplus from k_l back to $k_l k_{l+1}$ replacing it with surplus returned to k_l from $k_{l-1}k_l$ for $l = 1, ..., m-1$, and we transfer surplus from k_m back to $k_m j$ replacing it with surplus returned from $k_{m-1}k_m$. Finally, we transfer surplus from $k_m j$ to j. The transfers are made up to the maximum possible amount. We change the values of deficits and surpluses according to the described transfers. An opportunity of refinancing opportunity is illustrated in Figure 2.

In Example 2, there is a possibility of refinancing. See the middle diagram of Figure 1. Namely, a surplus of 1 can be transferred to 2 from 12, which allows for transferring a surplus of 1 back to 23, and this surplus can be transferred to 3. We reach the bottom diagram in Figure 1. As it is clear from the figure, this exhausts all financing possibilities and the possibilities of refinancing.

Final step. When all financing possibilities and the possibilities of refinancing are exhausted,[4] we remove some assets from the grand set N. The set of assets which remain will turn out to maximize the payoff function given by (1).

[4]Clearly, we must reach such a stage, since in every step of the procedure the sum of all surpluses and the sum of all deficits become lower.

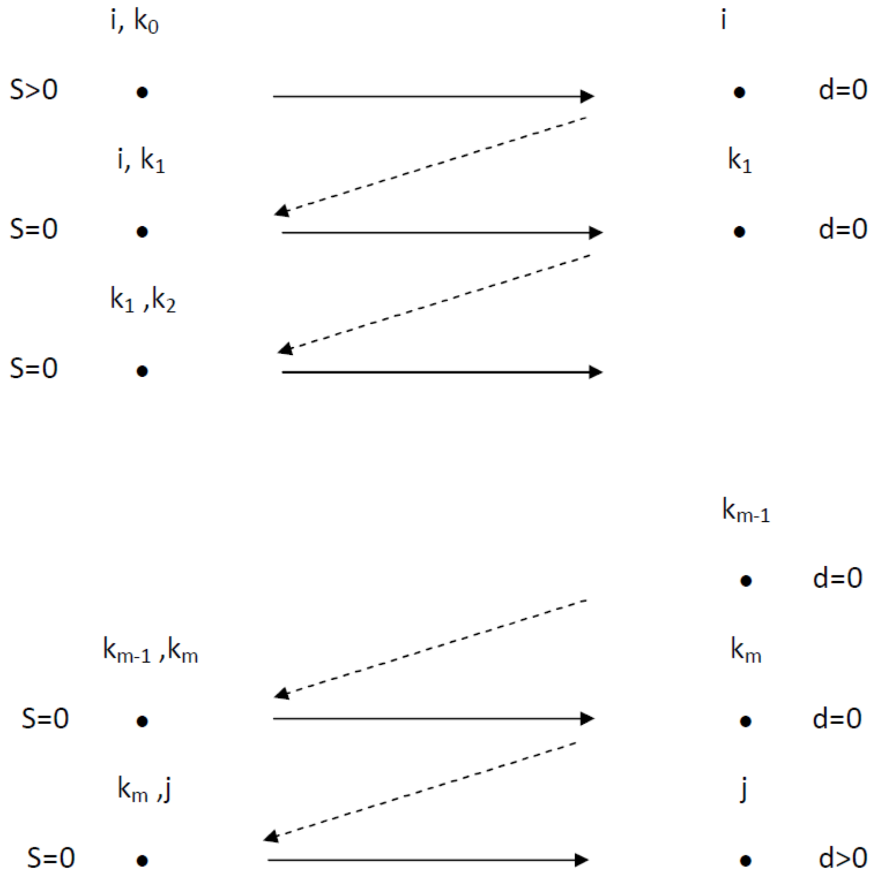

Figure 2. An opportunity of refinancing: solid arrows illustrate new transfers, and dashed lines illustrate transfer returns.

First, we remove any asset i which is not fully financed, that is, every asset with $d_i > 0$ in the moment in which all possibilities are exhausted. Then, we remove any other asset k, that was (entirely or partially) financed by a transfer from ik. We continue removing assets k which are financed by a transfer from ik for some previously removed assets i. We stop when no other asset can be removed. The remaining portfolio is the one selected by our algorithm, and will be denoted by S^*.

Returning to Example 2, asset 2 is fully financed, but asset 3 is not. Thus, we remove asset 3. However, we must now remove asset 2 as well, because asset 2 was partially financed by a transfer from 23. Thus, the optimal portfolio S^* consists only of asset 1.

Although the optimality of portfolio S^* selected by the financing algorithm follows from the optimality of the project selected by the Ford and Fulkerson algorithm (see Section 3), it is instructive to provide a simple direct proof of this result.

Proposition 1 *The financing algorithm selects a portfolio S^* which maximizes the payoff function given by (1).*

Proof. First, notice that

$$a\left(\sum_{i\in S^*}\mu_i-\sum_{i\in S^*}c_i\right)-b\sum_{i,j\in S^*}\sigma_{ij}\geq 0,$$

because the deficits of all assets from S^* are financed by the surpluses of ij for $i,j\in S^*$. Thus, (at least) some optimal portfolio contains as a subset the portfolio selected by the financing algorithm.

Next, observe that when no financing opportunity and no opportunity of refinancing exist, then at the time the algorithm stops, we have that $s_{ik}=0$ for all pairs i and k, except possibly pairs $i,k\in S^*$. Indeed, suppose that $s_{ik}>0$ for a pair i, k, such that, say, $i\notin S^*$. Consider the following two cases:

(i) If $d_i>0$ (again, at the time the algorithm stops), then there would exist a financing opportunity, since i's deficit could be financed from ik' surplus;

(ii) Suppose that $d_i=0$, and consider the set of all j with the following property:

(*) there exists a sequence $i,k_0,k_1,...,,k_m,j$, where $k_0=k$, such that i's deficit is (entirely or partially) financed from ik_1's surplus, k_l's deficit is financed from k_lk_{l+1}'s surplus for $l=1,...,m-1$, and k_m's deficit is financed from k_mj's surplus.

Since there is no opportunity of refinancing, $d_j=0$ for every j with property (*). By definition, the deficit of every j with property (*) was financed only from the surpluses of jk such that k also has property (*). Thus, since $d_j=0$ for every j with (*), the set of all j with (*) should be included into S^*. This, however, contradicts the assumption that $i\notin S^*$.

Consider any portfolio S that contains S^*. Denote by S' the portfolio obtained from S by including all assets which were removed from the grand set N later than the first removed asset that belongs to S; moreover, assume that every asset i included to S' has cost c_i', possibly lower than its actual cost c_i, such that its deficit (at the time when the first asset from S was removed) is equal to 0. Observe that $U(S)\leq U(S')$. This follows from the fact that the deficits of all assets $i\in S'-S$ (when their costs are c_i') are fully financed from the surpluses of pairs ij such that $i,j\in S'$.

Finally, notice that $U(S')=U(S^*)$. This follows because $s_{ij}=0$ for all pairs ij, possibly except pairs $i,j\in S^*$. ■

5 An example of comparative statics

In this section, we provide an example of comparative statics. We consider an increase in the cost of investing in an asset, and compare two assets: one with positive individual return but no hedging value, and the other with positive hedging value but no individual return. We argue that if both assets belong to the optimal portfolio, and the optimal portfolio contains a large number of assets, then the same increase of the cost is more likely to call for removing the former asset from the optimal portfolio than to call for removing the latter asset.

The intuition can be roughly explained as follows: A larger number of assets in the portfolio increases the chance that a given asset is highly correlated with another asset (or other assets), and so the chance that the asset's hedging value will exceeds the cost of investing in the asset, even when this cost increases by a given value. In contrast, the individual return of an asset is unaffected by the number of other assets in a portfolio.

This observation can be expressed in the form of the following proposition:

Proposition 2 *Suppose that the expected values of assets are drawn from some interval, variances are drawn from some interval, and covariances are drawn from an interval $[-\rho, 0]$, where $\rho \geq 0$; all draws are independent and uniform.*[5] *Consider two assets: 1 and 2 such that both assets cost $c_1 = c_2 = c$, and belong to the optimal portfolio for some unobserved draw of expected values, variances and covariances. Suppose that asset 1 has positive individual return and zero hedging value, while asset 2 has positive hedging value and zero individual return.*

If the number of assets in the optimal portfolio is sufficiently large, and cost $c+\Delta c$ is sufficiently small, then the chance that asset 1 will be removed from the optimal portfolio in response to an increase Δc in c_1 is higher than the chance that asset 2 will be removed from the optimal portfolio in response to an equal increase Δc in c_2.

Proof. It suffices to show that the probability of removing asset 2 tends to 0 as the number of assets in the optimal portfolio tends to ∞. Indeed, the probability of removing asset 1 is a positive number, independent the size of the optimal portfolio.

Let n be the number of assets in the optimal portfolio for the initial cost structure, that is, when the cost of asset 2 is c. Denote by $\rho_1, ..., \rho_{n-1}$ the realizations of covariances of asset 2 with other assets in the optimal portfolio. For any $r/b \in (0, \rho)$, if $c + \Delta c < r/b$, then the realizations of covariances for which asset 2 will be removed from the optimal portfolio belong to the simplex

$$\{(-\rho_1, ..., -\rho_{n-1}) \in [-\rho, 0]^{n-1} : \sum_{j=1}^{n-1} \rho_j < r/b\}.$$

The volume of this simplex is

$$\frac{1}{(n-1)!} \left(\frac{r}{b}\right)^{n-1},$$

so the probability that the realizations of covariances belong to this simplex is

$$\frac{1}{(n-1)!} \left(\frac{r}{b}\right)^{n-1} \frac{1}{\rho^{n-1}},$$

[5] We implicitly assume here that the intervals from which the parameters are drawn are such that every draw yields the covariance matrix of some random variables.

and tends to 0 as n tends to ∞. ■

We conjecture that the assumption that $c + \Delta c$ is sufficiently small is dispensable. The conjecture is likely to follow from an analogous argument. The only difficulty is that the proof would require computing the measure of some n-dimensional polytopes contained in the unit cube of more complicated form than simplices.

6 Positive correlations

In this section, we show that for the general covariance matrix, the problem of selecting an optimal portfolio is in NP-hard. More specifically, we will reduce our problem to the following partition problem, which is well-known to be NP-hard. In the partition problem, we are given a set $\{b_1, b_2, \ldots, b_m\}$ of positive integers and we would like to know if there is a subset S of this set such that

$$\sum_{i \in S} b_i = \sum_{i \notin S} b_i.$$

Given an instance of the partition problem, we construct an instance of portfolio selection whose solution will also resolve the partition problem. So, given positive integers $b_1, b_2, \ldots, b_m$, consider the grand set of $n = 2m$ assets, and an agent with mean-variance utility function $\mu - \sigma^2$, where μ is the mean and σ^2 is the variance of the portfolio the agent holds. (That is, in terms of our model, we set $a = b = 1$.) Let the means, costs and covariance matrix be such that:

1. (a) $\mu_i - \sigma_i^2 = 1$ and $c_i = 0$ for all $i = 1, \ldots, m$; (b) $\mu_i - \sigma_i^2 = 0$ and $c_i = b_{i-m}^2$ for all $i = m+1, \ldots, 2m$.

2. (a) $\sigma_{ij} = 0$ for $i \neq j$ and $i, j = 1, ..., m$; (b) $\sigma_{ij} = b_{i-m} b_{j-m}$ for $i \neq j$ and $i, j = m+1, ..., 2m$; (c) $\sigma_{ij} = -b_i b_{j-m}/2$ if $i = 1, ..., m$, $j = m+1, ..., 2m$, and $j - i \neq m$; (d) $\sigma_{ij} = -b_{i-m} b_j/2$ if $i = m+1, ..., 2m$, $j = 1, ..., m$, and $i - j \neq m$.

3. (a) $\sigma_{ij} = -(\sigma_j^2 + b_i b_i)/2$ if $i = 1, ..., m$, $j = m+1, ..., 2m$, and $j - i = m$; (b) $\sigma_{ij} = -(\sigma_i^2 + b_j b_j)/2$ if $i = m+1, ..., 2m$, $j = 1, ..., m$, and $j - i = m$.

That is, the covariance matrix $(\sigma_{ij})_{i,j=1}^n$ has the following form

$$\begin{bmatrix} \sigma_1^2 & ... & 0 & -(\sigma_{m+1}^2 + b_1 b_1)/2 & ... & -b_1 b_m \\ ... & & ... & ... & & ... \\ 0 & ... & \sigma_m^2 & -b_m b_1 & ... & -(\sigma_{2m}^2 + b_m b_m)/2 \\ -(\sigma_{m+1}^2 + b_1 b_1)/2 & ... & -b_1 b_m & \sigma_{m+1}^2 & ... & b_1 b_m \\ ... & & ... & ... & & ... \\ -b_m b_1 & ... & -(\sigma_{2m}^2 + b_m b_m)/2 & b_m b_1 & ... & \sigma_{2m}^2 \end{bmatrix}.$$

Note that we have not specified the variances of our assets. They must only be sufficiently large. This guarantees that $(\sigma_{ij})_{i,j=1}^n$ is positive semi-definite. And since $(\sigma_{ij})_{i,j=1}^n$ is also symmetric, it is the covariance matrix of a set of $n = 2m$ random variables.

Suppose that we solved this particular instance of portfolio selection. Let

$$B = \sum_{i=1}^{m} b_i.$$

Since, assets $1, ..., m$ belong to the optimal portfolio, no matter what other assets are included, our problem reduces to finding a set $S \subseteq \{m+1, \ldots, 2m\}$ that maximizes

$$U(\{1, ..., m\} \cup S) = m + 2 \sum_{i \in S} b_i (B - \sum_{i \in S} b_i).$$

Observe that if the answer to the partition problem is 'YES', then the utility of optimal portfolio for this particular instance of portfolio selection will be $m + B^2/2$. If the answer to the partition problem is 'NO', then the utility of optimal portfolio will be strictly less than $m + B^2/2$. That is, knowing the optimal portfolio, we can immediately say if the answer to the partition problem is 'YES' or 'NO'.

7 Conclusions

We conclude with spelling out the contribution of the present paper. First, the paper offers a rigorous analysis of what we think is a basic issue in economics and finance: how to choose a portfolio out of a finite set of available assets. This question was studied by Reiter (1963), but no formal solution concept was applied, and no formal result was proved. Reiter's pioneering analysis is restricted to showing that some simple algorithms perform very well in some numerical examples. Our paper also offers an algorithm that solves the portfolio problem in an important class of cases, when all pairs of assets are negatively correlated. This algorithm is a special case of the Ford and Fulkerson algorithm, and the contribution is rather in the recognition that the portfolio selection problem is a special case of the project selection problem. Note, in addition, that the original Ford and Fulkerson algorithm was somewhat simplified in our paper, which was possible due to the specific structure of Reiter's problem. Finally, the paper performs a comparative statics exercise concerning the optimal portfolio, and shows that for the case of general correlations the problem is NP-hard.

8 References

Chade H. and L. Smith. 2005. "Simultaneous Search," **Econometrica** 74, 1293-1307.

Dinic, E. A. 1970. "Algorithm for solution of a problem of maximum flow in a network with power estimation," **Soviet Math. Doklady (Doklady)**, 11, 1277-1280.

Edmonds, J. and R. M. Karp. 1972. "Theoretical improvements in algorithmic efficiency for network flow problems," **Journal of the ACM**, 19 (2), 248-264.

Ford, L. R. and D. R. Fulkerson. 1956. "Maximal flow through a network," **Canadian Journal of Mathematics**, 8, 399-404.

Gale D. and L. S. Shapley. 1962. "College Admissions and the Stability of Marriage", **American Mathematical Monthly** 69, 9-14.

Hansen, P. and B. Simeone. 1986. "Unimodular functions," **Discrete Applied Mathematics**, 14, 269-281.

Hochbaum, D. S. 2004. "Selection, Provisioning, Shared Fixed Costs, Maximum Closure, and Implications on Algorithmic Methods Today," **Management Science**, 50 (6), 709-723.

Milgrom P. and I. Segal. 2014. Deferred-acceptance auctions and radio spectrum reallocation, mimeo.

Rhys, J. M. W. 1970. "A Selection Problem of Shared Fixed Costs and Network Flows," **Management Science**, 17 (3), 200-207.

Reiter, S. 1963. "Choosing an Investment Program Among Interdependent Projects," **Review of Economic Studies**, 30, 32-36.