A Note on "Risk reduction in large portfolios: Why imposing the wrong constraints helps."

RAVI JAGANNATHAN, TONGSHU MA, and JIAQI ZHANG *

ABSTRACT

This note corrects an error in the proof of Proposition 2 of "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraint Helps" that appeared in the Journal of Finance, August 2003.

*Ravi Jagannathan and Jiaqi Zhang: Kellogg School, Northwestern University, Evanston, IL 60208; Tongshu Ma: Binghamton University, Binghamton, NY 13902-6000.

Proposition 1 (page 1655) of Jagannathan and Ma (2003) shows that the global minimum variance portfolio of a given a sample covariance matrix when there are "no short sale" constraints and upper bounds on portfolio weights is the unconstrained global minimum variance portfolio when the covariance matrix is adjusted by shrinking its elements that are large in magnitude in a particular manner.

Proposition 2 (page 1656), says that: (1) The adjusted covariance matrix in Proposition 1 is the maximum likelihood estimate of the covariance matrix subject to the constraint that its implied global minimum variance portfolio satisfies the portfolio weight constraints; and (2) The global minimum variance portfolio of the constrained maximum likelihood estimate of the covariance matrix is also the global minimum variance of the sample covariance matrix when there are "no short sale" constraints and upper bounds on portfolio weights.

Unfortunately, there is an error in the proof of Proposition 2, and in this note we correct the error. Equation (8) in Jagannathan and Ma (2003) should have been Equation (10) in this note. The two equations are the same except for the fourth term on the right side of Equation (10) in this note. The fourth term was omitted in Jagannathan and Ma (2003) by mistake, and that mistake carried over to the proofs of the propositions. Because of this error, the proofs in the paper are correct only when there are lower bounds on portfolio weights (i.e., "no short sale" constraints only) but not when there are upper bounds on portfolio weights as well. We now have the fourth term incorporated, and the proofs of the propositions have been corrected. The propositions as originally stated in Jagannathan and Ma (2003) remain valid. In the following we present a self-contained discussion of the effects of portfolio weight constraints that corrects the omission in Section I of the original paper.

Given an estimated covariance matrix S, the global portfolio variance minimization problem when portfolio weights are constrained to satisfy both a lower bound of zero and a upper bound of $\bar{\omega}$ is given by

$$\min_{\omega} \; \frac{1}{2}\,\omega' S\omega, \tag{1}$$

$$s.t. \; \sum_i \omega_i = 1. \tag{2}$$

$$\omega_i \geq 0, \;\; i = 1, 2, ..., N. \tag{3}$$

$$\omega_i \leq \bar{\omega}, \;\; i = 1, 2, ..., N. \tag{4}$$

Then, the Lagrangian expression is

$$\mathcal{L} = \frac{1}{2}\omega' S\omega - \lambda_0\Big(\sum_i \omega_i - 1\Big) - \sum_i \lambda_i \omega_i - \sum_i \delta_i(\bar{\omega} - \omega_i). \tag{5}$$

Here $\lambda = (\lambda_1, ..., \lambda_N)'$ are the Lagrange multipliers for the nonnegativity constraints (3), $\delta = (\delta_1, ..., \delta_N)'$ the multipliers for the constraints (4), and λ_0 is the multiplier for (2).

It is well-known that the Karush-Kuhn-Tucker conditions below are sufficient and necessary for the above optimization problem.

$$\sum_j S_{i,j}\omega_j - \lambda_i + \delta_i - \lambda_0 = 0, \tag{6}$$

$$\lambda_0, \lambda_i, \;\delta_i, \,\omega_i, \,(\bar{\omega} - \omega_i) \geq 0, \tag{7}$$

$$\lambda_i \omega_i = 0, \tag{8}$$

$$\delta_i(\bar{\omega} - \omega_i) = 0. \tag{9}$$

All 4 equations above hold for all $i = 1, 2, ..., N$.

Note that (1) is concave as a function of ω and the constraints are all convex, so the solution to the constrained portfolio variance minimization problem (1) is unique. Denote it as $\omega^{++}(S)$. Let $\mathbf{1}$ denote the column vector of ones. Then we have the following proposition.

PROPOSITION 1: *Let*

$$\tilde{S} = S + (\delta\mathbf{1}' + \mathbf{1}\delta') - (\lambda\mathbf{1}' + \mathbf{1}\lambda') - 2\bar{\omega}(\delta'\mathbf{1})(\mathbf{1}\mathbf{1}'). \tag{10}$$

Then $\tilde{S}$ is symmetric and positive semi-definite, and $\omega^{++}(S)$ is one of its global minimum variance portfolios.

Proof: The matrix $\tilde{S}$ is obviously symmetric. Now we prove that it is positive semi-definite. Suppose that

$$(\omega_1, ..., \omega_N, \lambda_1, ..., \lambda_N, \delta_1, ..., \delta_N, \lambda_0) \equiv ((\omega^{++})', \lambda', \delta', \lambda_0)$$

is a solution to the portfolio variance minimization problem given in equation (1) with constraints given in equations (2)-(4). In order to prove $\tilde{S}$ is positive semi-definite, for any vector x, we need to check that $x'\tilde{S}x \geq 0$.

For notational simplicity, write (10) as

$$\tilde{S} = S - L.$$

where

$$L = -(\delta\mathbf{1}' + \mathbf{1}\delta') + (\lambda\mathbf{1}' + \mathbf{1}\lambda') + 2\bar{\omega}(\delta'\mathbf{1})(\mathbf{1}\mathbf{1}'). \tag{11}$$

We can easily check that for any z,

$$z'Lz = 0 \text{ if } z'\mathbf{1} = 0. \tag{12}$$

Now consider two situations.

1. $x'\mathbf{1} = 0$.

We have $x'Lx = 0$, hence $x'\tilde{S}x = x'Sx \geq 0$. This inequality shows that $\tilde{S}$ is positive semi-definite. We know that for any x, $x'Sx \geq (\omega^{++})'S\omega^{++}$. Then for any x, $x'\tilde{S}x \geq (\omega^{++})'\tilde{S}\omega^{++}$, and $\omega^{++}(S)$ is one of the global minimum variance portfolios of $\tilde{S}$. Q.E.D.

2. $x'\mathbf{1} \neq 0$.

Without loss of generality, we can assume

$$x'\mathbf{1} = 1. \tag{13}$$

We first state some results. From the Karush-Kuhn-Tucker conditions (6), (8) and (9), we have correspondingly

$$\lambda - \delta = S\omega^{++} - \lambda_0\mathbf{1}, \tag{14}$$

$$(\omega^{++})'\lambda = 0, \tag{15}$$

$$(\omega^{++})'\delta = \bar{\omega}\delta'\mathbf{1}. \tag{16}$$

Then, noticing that from (2) we have $(\omega^{++})'\mathbf{1} = 1$,

$$
\begin{aligned}
(\omega^{++})'L &= -(\omega^{++})'\delta\mathbf{1}' - (\omega^{++})'\mathbf{1}\delta' + (\omega^{++})'\lambda\mathbf{1}' + (\omega^{++})'\mathbf{1}\lambda' \\
&\quad + 2\bar{\omega}(\delta'\mathbf{1})(\omega^{++})'(\mathbf{1}\mathbf{1}') \\
&= -\bar{\omega}(\delta'\mathbf{1})\mathbf{1}' - \delta' + 0 + \lambda' + 2\bar{\omega}(\delta'\mathbf{1})\mathbf{1}' \\
&= \lambda' - \delta' + \bar{\omega}(\delta'\mathbf{1})\mathbf{1}'. \qquad (17)
\end{aligned}
$$

From (15), (16) and (17),

$$
(\omega^{++})'L\omega^{++} = \lambda\omega^{++} - \delta\omega^{++} + \bar{\omega}(\delta'\mathbf{1}) = 0. \qquad (18)
$$

So

$$
\begin{aligned}
(\omega^{++})'S\omega^{++} &= (\omega^{++})'\tilde{S}\omega^{++} + (\omega^{++})'L\omega^{++} \\
&= (\omega^{++})'\tilde{S}\omega^{++}. \qquad (19)
\end{aligned}
$$

Now we are ready to check that $x'\tilde{S}x \geq 0$.

$$
\begin{aligned}
x'\tilde{S}x &= x'Sx - x'Lx \\
&= x'Sx - (\omega^{++} + (x - \omega^{++}))'L(\omega^{++} + (x - \omega^{++})) \\
&= x'Sx - (\omega^{++})'L\omega^{++} - 2(\omega^{++})'L(x - \omega^{++}) \\
&\quad - (x - \omega^{++})'L(x - \omega^{++}) \\
&= x'Sx - 2(\omega^{++})'L(x - \omega^{++}). \qquad (20)
\end{aligned}
$$

This is because $(x - \omega^{++})'\mathbf{1} = 1 - 1 = 0$, and then $(x - \omega^{++})'L(x - \omega^{++}) = 0$ according to (12). $(\omega^{++})'L\omega^{++} = 0$ is from (18). Then,

$$
\begin{aligned}
x'\tilde{S}x &= x'Sx - 2(\lambda' - \delta' + \bar{\omega}\delta'\mathbf{1}\mathbf{1}')(x - \omega^{++}) \\
&= x'Sx - 2(\lambda' - \delta')(x - \omega^{++}) \\
&= x'Sx - 2(S\omega^{++} - \lambda_0\mathbf{1})'(x - \omega^{++}) \\
&= x'Sx - 2(\omega^{++})'S(x - \omega^{++}) \\
&= x'Sx - 2(\omega^{++})'Sx + (\omega^{++})'S\omega^{++} + (\omega^{++})'S\omega^{++} \\
&= (x - \omega^{++})'S(x - \omega^{++}) + (\omega^{++})'S\omega^{++} \\
&\geq (\omega^{++})'S\omega^{++} \\
&= (\omega^{++})'\tilde{S}\omega^{++} \qquad (21) \\
&\geq 0. \qquad (22)
\end{aligned}
$$

The first equality follows from (17). The second and fourth equality holds because $(x-\omega^{++})'\mathbf{1} = 0$, and S is symmetric. The third equality follows from (14). The last equality follows from (19). Both inequalities hold because S is positive semidefinite.

From (22), $\tilde{S}$ is positive semi-definite.

To prove that ω^{++} is a global minimum variance portfolio of $\tilde{S}$, it is sufficient to show that the variance of all portfolios, x, such that $x'\mathbf{1} = 1$, is not smaller than $(\omega^{++})'\tilde{S}\omega^{++}$. This is exactly what equation (21) implies. Q.E.D.

The proposition is proved by combining two situations discussed above.

This result shows that constructing a constrained global minimum variance portfolio from S is equivalent to constructing a (unconstrained) minimum variance portfolio from $\tilde{S} = S + (\delta\mathbf{1}' + \mathbf{1}\delta') - (\lambda\mathbf{1}' + \mathbf{1}\lambda') - 2\bar{\omega}(\delta'\mathbf{1})(\mathbf{1}\mathbf{1}')$.

In general, given a constrained optimal portfolio $\omega^{++}(S)$, there are many covariance matrix estimates that have $\omega^{++}(S)$ as their (unconstrained) minimum variance portfolio. Is there anything special about $\tilde{S}$? We do have an answer to this question when returns are jointly normal and S is the MLE of the population covariance matrix.

Let the $N \times 1$ return vector $h_t = (r_{1t}, r_{2t}, ..., r_{Nt})'$ be i.i.d. normal $N(\mu, \Omega)$. Then the MLE of Ω is the sample covariance matrix $S = \sum_{t=1}^{T}(h_t - \bar{h})(h_t - \bar{h})'/T$. The likelihood function depends on both μ and Ω, even though we want to estimate Ω only. To get rid of the dependence on μ, recall that for any estimate of the covariance matrix, the MLE of the mean is always the sample mean (Morrison 2005). With this estimate of the mean, the log-likelihood (as a function of the covariance matrix alone) becomes

$$l(\Omega) = \text{CONST} - \frac{T}{2}\ln|\Omega| - \frac{T}{2}\text{tr}(S\Omega^{-1}). \tag{23}$$

This can also be considered as the likelihood function of Ω^{-1} and is defined for nonsingular Ω.

Now consider the constrained MLE of Ω, subject to the constraint that the global minimum variance portfolio constructed from Ω satisfies the weight constraints (3)-(4). Let $\Omega_{i,j}$ denote the (i,j)-th element of Ω and $\Omega^{i,j}$ denote the (i,j)-th element of Ω^{-1}, then the constraints are

$$\sum_j \Omega^{i,j} \geq 0, \;\; i = 1, 2, ..., N. \tag{24}$$

$$\sum_j \Omega^{i,j} \leq \bar{\omega}\sum_k\sum_j \Omega^{k,j}, \;\; i = 1, 2, ..., N. \tag{25}$$

So the constrained ML problem is to maximize (23), subject to constraints (24)-(25). The Lagrangian expression for this is

$$\mathcal{L} = l(\Omega) + \sum_i \tilde{\lambda}_i \sum_j \Omega^{i,j} + \sum_i \tilde{\delta}_i \left(\bar{\omega}\sum_k\sum_j \Omega^{k,j} - \sum_j \Omega^{i,j}\right). \tag{26}$$

We have the following proposition.

Proposition 2: *Assume that returns are jointly i.i.d. normal $N(\mu, \Omega)$. Let S be the unconstrained MLE of Ω.*

1. *Given S, let $\{\lambda_i, \delta_i, \omega_i\}_{i=1,...,N}$ be a solution to the constrained portfolio variance minimization problem (1)-(4), and construct $\tilde{S}$ according to (10). Assume $\tilde{S}$ is non-singular. Then, $\tilde{S}$ and $\{\tilde{\lambda}_i, \tilde{\delta}_i\} = \{\lambda_i, \delta_i\}_{i=1,...,N}$ jointly satisfy the first-order conditions for the constrained ML problem.*

2. *Let $\tilde{S}$ and $\{\tilde{\lambda}_i, \tilde{\delta}_i\}$ jointly satisfy the first-order conditions for the constrained ML problem. For $i = 1, ..., N$, define $\omega_i = \sum_j \tilde{S}^{i,j} / \sum_{k,l}\tilde{S}^{k,l}$, the normalized row sums of $\tilde{S}^{-1}$. Then, the sample covariance matrix $S = \tilde{S} - (\tilde{\delta}\mathbf{1}' + \mathbf{1}\tilde{\delta}') + (\tilde{\lambda}\mathbf{1}' + \mathbf{1}\tilde{\lambda}') + 2\bar{\omega}(\tilde{\delta}'\mathbf{1})(\mathbf{1}\mathbf{1}')$; and $\{\lambda_0, \lambda_i, \delta_i, \omega_i\}_{i=1,...,N}$ is a solution to the constrained portfolio variance minimization problem (1)-(4), where $\lambda_0 = \frac{1}{\mathbf{1}'\tilde{S}^{-1}\mathbf{1}} + \bar{\omega}\tilde{\delta}'\mathbf{1}$, and $\{\lambda_i, \delta_i\} = \{\tilde{\lambda}_i, \tilde{\delta}_i\}, \;\; i = 1, ..., N$.*

Proof of Proposition 2:

Before the proof we will show the expressions of the Karush-Kuhn-Tucker conditions.

Let the solution to the constrained ML problem be $\hat{\Omega}$. Morrison (2005) (equation (6) on page 23) shows that the derivatives of the likelihood function are

$$\frac{\partial l}{\partial \Omega^{i,i}} = \frac{1}{2}\Omega_{i,i} - \frac{1}{2}S_{i,i}, \text{ all } i. \tag{27}$$

$$\frac{\partial l}{\partial \Omega^{i,j}} = \Omega_{i,j} - S_{i,j}, \text{ all } i < j. \tag{28}$$

Note that the log likelihood is concave as a function of Ω^{-1} (Zwiernik, Uhler, and Richards (2017)) and the constraints (24)-(25) are all convex, so the estimate satisfying the Karush-Kuhn-Tucker conditions is the constrained MLE.

The derivatives of the following two terms of the Lagrangian expression are

$$\begin{aligned}&\frac{\partial}{\partial \Omega^{i,i}}\left[\sum_k \tilde{\lambda}_k \sum_j \Omega^{k,j} + \sum_k \tilde{\delta}_k(\bar{\omega}\sum_l \sum_j \Omega^{l,j} - \sum_j \Omega^{k,j})\right]\\ =&\tilde{\lambda}_i - \tilde{\delta}_i + \bar{\omega}\sum_k \tilde{\delta}_k,\end{aligned} \tag{29}$$

$$\begin{aligned}&\frac{\partial}{\partial \Omega^{i,j}}\left[\sum_k \tilde{\lambda}_k \sum_m \Omega^{k,m} + \sum_k \tilde{\delta}_k(\bar{\omega}\sum_l \sum_m \Omega^{l,m} - \sum_m \Omega^{k,m})\right]\\ =&(\tilde{\lambda}_i + \tilde{\lambda}_j) - (\tilde{\delta}_i + \tilde{\delta}_j) + 2\bar{\omega}\sum_k \tilde{\delta}_k.\end{aligned} \tag{30}$$

Combining (27)-(30), we have that the Karush-Kuhn-Tucker conditions for the constrained ML problem (23)-(25) are

$$\frac{1}{2}\hat{\Omega}_{i,i} - \frac{1}{2}S_{i,i} = -\tilde{\lambda}_i + \tilde{\delta}_i - \bar{\omega}\sum_k \tilde{\delta}_k, \text{ all } i. \tag{31}$$

$$\hat{\Omega}_{i,j} - S_{i,j} = -(\tilde{\lambda}_i + \tilde{\lambda}_j) + (\tilde{\delta}_i + \tilde{\delta}_j) - 2\bar{\omega}\sum_k \tilde{\delta}_k, \text{ all } i < j. \tag{32}$$

$$\tilde{\lambda}_i,\ \tilde{\delta}_i,\ \sum_j \hat{\Omega}^{i,j},\ \left(\bar{\omega}\sum_k \sum_j \hat{\Omega}^{k,j} - \sum_j \hat{\Omega}^{i,j}\right) \geq 0, \text{ all } i. \tag{33}$$

$$\tilde{\lambda}_i \sum_j \hat{\Omega}^{i,j} = 0, \text{ all } i. \tag{34}$$

$$\tilde{\delta}_i(\sum_j \hat{\Omega}^{i,j} - \bar{\omega}\sum_k \sum_j \hat{\Omega}^{k,j}) = 0, \text{ all } i. \tag{35}$$

Conditions (31) and (32) imply that the constrained MLE can be written as

$$\hat{\Omega} = S - (\tilde{\lambda}\mathbf{1}' + \mathbf{1}\tilde{\lambda}') + (\tilde{\delta}\mathbf{1}' + \mathbf{1}\tilde{\delta}') - 2\bar{\omega}(\tilde{\delta}'\mathbf{1})(\mathbf{1}\mathbf{1}'). \tag{36}$$

Note that $\hat{\Omega}$ has the same form as $\tilde{S}$ in (10). We will prove the proposition below using this observation.

The proof of Part 1:

Let $\{\lambda_i, \delta_i, \omega_i\}_{i=1,...,N}$ be a solution to the constrained portfolio variance minimization problem (1)-(4), given S, and construct $\tilde{S}$ according to (10). We want to prove that (31)-(35) still hold if replacing $\hat{\Omega}$ in them with $\tilde{S}$.

(33)-(35) directly follow (7)-(9), given $\omega_i = \sum_j \tilde{S}^{i,j} / \sum_{k,l} \tilde{S}^{k,l}$ already proved in Proposition 1.

If replacing $\hat{\Omega}$ in(36) with $\tilde{S}$ it would be the same as (10), so it holds by definition of $\tilde{S}$. This implies (31)-(32). Q.E.D.

The proof of Part 2:

Let $\tilde{S}$ and $\{\tilde{\lambda}_i, \tilde{\delta}_i\}_{i=1,...,N}$ jointly satisfy the first-order conditions for the constrained ML problem. Then, (31)-(35) hold if replacing $\hat{\Omega}$ in them with $\tilde{S}$, and (31)-(32) imply (36). To show that $\{\lambda_i, \delta_i, \omega_i\}_{i=1,...,N}$ is a solution to the constrained portfolio variance minimization problem (1)-(4), we need to prove (6)-(9).

(7)-(9) directly follow (33)-(35), given $\omega_i = \sum_j \tilde{S}^{i,j} / \sum_{k,l} \tilde{S}^{k,l}$ by definition.

The remained equation (6) can be written as

$$S\omega = \lambda - \delta + \lambda_0 \mathbf{1}. \tag{37}$$

From the definition of S here, $S = \tilde{S} - (\tilde{\delta}\mathbf{1'} + \mathbf{1}\tilde{\delta}') + (\tilde{\lambda}\mathbf{1'} + \mathbf{1}\tilde{\lambda}') + 2\bar{\omega}(\tilde{\delta}'\mathbf{1})(\mathbf{11'})$, we have

$$S\omega = \tilde{S}\omega + \tilde{\lambda}\mathbf{1}'\omega + \mathbf{1}\tilde{\lambda}'\omega - \tilde{\delta}\mathbf{1}'\omega - \mathbf{1}\tilde{\delta}'\omega + 2\bar{\omega}(\tilde{\delta}'\mathbf{1})\mathbf{11}'\omega. \tag{38}$$

And from definition of ω,

$$\tilde{S}\omega = \tilde{S}\frac{\tilde{S}^{-1}\mathbf{1}}{\mathbf{1}'\tilde{S}^{-1}\mathbf{1}} = \frac{1}{\mathbf{1}'\tilde{S}^{-1}\mathbf{1}}\mathbf{1}, \tag{39}$$

$$\mathbf{1}'\omega = \mathbf{1}'\frac{\tilde{S}^{-1}\mathbf{1}}{\mathbf{1}'\tilde{S}^{-1}\mathbf{1}} = 1. \tag{40}$$

Following (8)-(9)

$$\tilde{\lambda}'\omega = 0, \tag{41}$$

$$\tilde{\delta}'\omega = \bar{\omega}(\tilde{\delta}'\mathbf{1}). \tag{42}$$

Substituting (39)-(42) into (38), and using the definition of λ_0, we have

$$\begin{aligned} S\omega &= \frac{1}{\mathbf{1}'\tilde{S}^{-1}\mathbf{1}}\mathbf{1} + \tilde{\lambda} + 0 - \tilde{\delta} - \bar{\omega}(\tilde{\delta}'\mathbf{1})\mathbf{1} + 2\bar{\omega}(\tilde{\delta}'\mathbf{1})\mathbf{1} \\ &= \lambda - \delta + \lambda_0\mathbf{1}. \end{aligned} \tag{43}$$

Then (6) is also proved. Q.E.D.

According to this proposition the $\tilde{S}$ constructed from the solution to the constrained global variance minimization problem is the ML estimator of the covariance matrix, subject to the condition that the global minimum variance portfolio weights satisfy the nonnegativity and upper bound constraints. So we could impose the constraints in the estimation stage instead of the optimization stage and the result would be the same.

When only the nonnegativity constraint is imposed, the vector of Lagrange multipliers for the upper bound will be zero. So $\tilde{S} = S - (\lambda \mathbf{1}' + \mathbf{1}\lambda')$, and we can simplify the statements in the two Propositions in a straightforward way.

Jagannathan and Ma (2003) gave a shrinkage interpretation of the effects of portfolio weight constraints. Their interpretation of the terms $(\mathbf{1}\delta' + \delta\mathbf{1}')$ and $(\mathbf{1}\lambda' + \lambda\mathbf{1}')$ remain valid. For brevity we do not repeat these interpretations here. The originally omitted term, $-2\bar{\omega}(\delta'\mathbf{1})(\mathbf{1}\mathbf{1}')$, reduces the variances and covariances of all stocks by equal amount. Doing so does not change the global minimum variance portfolio. Hence after taking this term into account, the shrinkage interpretation would not change.

REFERENCES

Jagannathan, Ravi, and Tongshu Ma, 2003, Risk reduction in large portfolios: Why imposing the wrong constraints helps, *The Journal of Finance* 58, 1651–1683.

Morrison, Donald, 2005, Multivariate statistical methods.

Zwiernik, Piotr, Caroline Uhler, and Donald Richards, 2017, Maximum likelihood estimation for linear gaussian covariance models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79, 1269–1292.