MANAGEMENT SCIENCE
DOI 10.1287/mnsc.1080.0986ec
pp. ec1–ec6

e-companion
ONLY AVAILABLE IN ELECTRONIC FORM

Electronic Companion—"A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms" by Victor DeMiguel, Lorenzo Garlappi, Francisco J. Nogales, and Raman Uppal, *Management Science*, DOI 10.1287/mnsc.1080.0986.

Material for on-line appendix

EC.1. Appendix: Details of the Partial Minimum-Variance Portfolios

In this appendix we explain the method used to compute the partial minimum-variance portfolios. Because the conjugate-gradient method can be used to solve only unconstrained optimization problems, we first show in Section EC.1.1 how to eliminate the constraint that the portfolio weights sum to one. Section EC.1.2 gives the method statement.

EC.1.1. Expressing Minimum-Variance Problem Without Adding-Up Constraint

To eliminate the constraint $\mathrm{w}^\top e = 1$, we will use the following proposition showing that any portfolio whose weights sum to one can be rewritten as the $1/N$ portfolio plus a portfolio whose weights sum to zero, herein a *zero-cost* portfolio.

Proposition EC.1. *Let* $\mathrm{w} \in \mathbb{R}^N$ *and* $\mathrm{w}^\top e = 1$*, then the following holds:*

1. *The portfolio* w *may be expressed as*

$$\mathrm{w} = \frac{e}{N} + \mathrm{w}_0, \tag{EC.1}$$

in which w_0 *is a zero-cost portfolio; that is,* $\mathrm{w}_0^\top e = 0$.

2. *There exists a matrix* $Z \in \mathbb{R}^{N\times(N-1)}$ *whose columns form an orthonormal basis for the subspace of zero-cost portfolios (that is, an orthonormal basis for the subspace of all portfolios satisfying* $\mathrm{w}^\top e = 0$*), and this matrix satisfies* $e^\top Z = 0$ *and* $Z^\top Z = I$.

3. *There exists a vector* $\mathrm{w}_Z \in \mathbb{R}^{N-1}$ *such that*

$$\mathrm{w} = \frac{e}{N} + Z\mathrm{w}_Z. \tag{EC.2}$$

Part 1: Let w_0 be the portfolio defined as $\mathrm{w}_0 = \mathrm{w} - e/N$. Then note that if the weights of the portfolio w sum up to one, $\mathrm{w}^\top e = 1$, then the weights of the portfolio w_0 must sum up to zero because $e^\top \mathrm{w}_0 = e^\top \mathrm{w} - e^\top e/N = 0$.

Part 2: The existence of an orthonormal basis for the subspace of zero-cost portfolios follows from general vector space theory (Nocedal and Wright (1999)). Let $Z \in \mathbb{R}^{N\times(N-1)}$ be the matrix whose columns are the vectors composing the orthonormal basis for the zero-cost portfolio subspace, then, because all zero-cost portfolios are orthogonal to e, we must have $e^\top Z = 0$. Finally, because the columns of Z are orthogonal to each other and their 2-norm is equal to one, we have that $Z^\top Z = I$.

Part 3: The result follows from Part 1 because w_0 is a zero-cost portfolio and Z is a basis of the zero-cost portfolio subspace.

■

We define *portfolio space* to be the $\mathbb{R}^N$ space, *zero-cost portfolio subspace* the subspace in $\mathbb{R}^N$ that is formed by all linear combinations of the columns of Z, and *reduced space* the $N-1$ dimensional space $\mathbb{R}^{N-1}$, which w_Z inhabits. The following proposition shows that expression (EC.2) may be used to eliminate the constraint from the minimum-variance problem.

Proposition EC.2. *The sample minimum-variance portfolio can be written as* $\mathrm{w}_{MINU} = \frac{e}{N} + Z\mathrm{w}_Z$*, where* w_{MINU} *is the solution to the* unconstrained minimum-variance problem*:*

$$\min_{\mathrm{w}_Z} \quad \left(\frac{e}{N} + Z\mathrm{w}_Z\right)^\top \hat{\Sigma} \left(\frac{e}{N} + Z\mathrm{w}_Z\right). \tag{EC.3}$$

S ubstituting (EC.1) into (1)–(2) yields the result.

■

EC.1.2. Applying the Conjugate-Gradient Method

To compute the partial minimum-variance portfolios, we apply the conjugate-gradient method to solve the first-order optimality conditions for the *unconstrained* minimum-variance problem:

$$Z^\top \hat{\Sigma} Z \mathrm{w}_Z = -Z^\top \hat{\Sigma} \frac{e}{N}.$$

In our implementation, we use the $1/N$ portfolio as the starting point for the conjugate-gradient method. The method commences by finding a zero-cost portfolio w_{CG1} that induces the maximum marginal rate of decrease in sample variance when combined with the $1/N$ portfolio. In other words, w_{CG1} is a zero-cost portfolio such that when an infinitesimally small multiple of this zero-cost portfolio is added to the $1/N$ portfolio, it induces the maximum rate of decrease in the overall portfolio sample variance. Mathematically, this implies that the first conjugate portfolio must be proportional to the negative of the projection of the gradient of the portfolio sample variance onto the zero-cost portfolio subspace:

$$\mathrm{w}_{CG1} \propto -ZZ^\top \nabla_{\mathrm{w}} \left(\mathrm{w}^\top \hat{\Sigma} \mathrm{w} \right) \big|_{e/N}, \tag{EC.4}$$

where $ZZ^\top$ is the projection matrix that projects any portfolio onto the zero-cost subspace and $\nabla_{\mathrm{w}} \left(\mathrm{w}^\top \hat{\Sigma} \mathrm{w} \right) \big|_{e/N}$ is the gradient of the portfolio sample variance evaluated at the $1/N$ portfolio. Evaluating this gradient we have that the first conjugate portfolio can be written as:[18]

$$\mathrm{w}_{CG1} = -ZZ^\top \hat{\Sigma} \frac{e}{N}. \tag{EC.5}$$

Then the conjugate-gradient method finds the combination of the $1/N$ portfolio and the first conjugate-gradient portfolio that minimizes the sample portfolio variance. We term the resulting portfolio the first partial minimum-variance portfolio (w_{PAR1}), where the term "partial" refers to the fact that the portfolio minimizes the sample variance within the subset of portfolios formed by combinations of the $1/N$ portfolio and the first conjugate portfolio.

One may feel tempted to follow the same strategy in the successive iterations, and hence find the zero-cost portfolio that would induce the maximum marginal rate of decrease in sample variance when combined with w_{PAR1}, but it can be shown (Nocedal and Wright (1999)) that it is more efficient to find the zero-cost portfolio w_{CG2} that induces the maximum marginal rate decrease in sample variance, subject to the condition that it is *conjugate* with respect to w_{CG1} (i.e., $\mathrm{w}_{CG2} \hat{\Sigma} \mathrm{w}_{CG1} = 0$). Then the method finds the optimal combination of w_{PAR1} and w_{CG2}, which is termed the second partial minimum-variance portfolio. If this process is iterated $N-1$ times, we recover the minimum-variance portfolio. If instead we perform this operation only $K < N-1$ times, we obtain a portfolio that lies between the $1/N$ portfolio and the minimum-variance portfolio.

The following proposition describes the relation between the partial minimum-variance portfolios and the 2-norm-constrained portfolios.

[18] For notational convenience we have divided the gradient of the sample variance by a factor of 2. Note that this is correct because in (EC.4) we only state that w_{CG1} is proportional to the expression in the right-hand side.

Proposition EC.3. *Provided the matrix $\hat{\Sigma}$ is nonsingular, the 2-norm-constrained portfolios form a continuously differentiable curve for δ ranging from $1/N$ to $\|w_{MINU}\|_2^2$, and the difference between the first partial minimum-variance portfolio and the $1/N$ portfolio is tangential to this curve at $\delta = 1/N$.*

U sing the unconstrained minimum-variance problem (EC.3), we can rewrite the 2-norm-constrained minimum-variance problem as

$$\min_{\mathrm{w}_Z} \quad \left(\frac{e}{N} + Z\mathrm{w}_Z\right)^\top \hat{\Sigma} \left(\frac{e}{N} + Z\mathrm{w}_Z\right) \tag{EC.6}$$

$$\text{s.t.} \quad \|e/N\|_2^2 + \|\mathrm{w}_Z\|_2^2 \leq \delta. \tag{EC.7}$$

Moreover, because $\|e/N\|_2^2 = 1/N$ we can then rewrite this problem as

$$\min_{\mathrm{w}_Z} \quad \left(\frac{e}{N} + Z\mathrm{w}_Z\right)^\top \hat{\Sigma} \left(\frac{e}{N} + Z\mathrm{w}_Z\right) \tag{EC.8}$$

$$\text{s.t.} \quad \|\mathrm{w}_Z\|_2^2 \leq \tilde{\delta}, \tag{EC.9}$$

in which $\tilde{\delta} = \delta - 1/N$. Also, provided the matrix $\hat{\Sigma}$ is nonsingular, then we also have that $Z^\top \hat{\Sigma} Z$ is nonsingular; hence, the solution to problem (EC.8)–(EC.9) is unique and is the unique solution to the first-order optimality conditions for problem (EC.8)–(EC.9). Then, following an argument similar to that in the proof of Proposition 2, we have that, provided the matrix $\hat{\Sigma}$ is nonsingular, for each δ there exists a $\nu \geq 0$ such that the solution to the unique 2-norm-constrained problem coincides with the unique solution to the following problem:

$$\min_{\mathrm{w}_Z} \quad \left(\frac{e}{N} + Z\mathrm{w}_Z\right)^\top \hat{\Sigma} \left(\frac{e}{N} + Z\mathrm{w}_Z\right) + \nu \mathrm{w}_Z^\top \mathrm{w}_Z, \tag{EC.10}$$

in which $\mathrm{w}_Z^\top \mathrm{w}_Z = \|\mathrm{w}_Z\|_2^2$. Moreover, the uniqueness of the minimizer to this problem ensures that the curve of 2-norm-constrained portfolios for different values of δ or ν is continuously differentiable (see Fiacco and McCormick (1968, Theorem 14)). The first-order conditions give the unique solution to problem (EC.10):

$$\mathrm{w}_Z = -(Z^\top \hat{\Sigma} Z + \nu I)^{-1} Z^\top \hat{\Sigma} \frac{e}{N}. \tag{EC.11}$$

We are interested in the tangent to the curve of solutions to the 2-norm-constrained problem. This is given by the differential of expression (EC.11) with respect to ν. We are interested in this differential at the $1/N$ portfolio, which is obtained for very large values of the penalty parameter ν.[19] For very large values of ν, we have that

$$(Z^\top \hat{\Sigma} Z + \nu I) \approx \nu I, \tag{EC.12}$$

and hence,

$$\frac{d}{d\nu}(Z^\top \hat{\Sigma} Z + \nu I)^{-1} \approx \frac{d}{d\nu}(\nu I)^{-1} = \frac{d}{d\nu}(\frac{1}{\nu} I) = -\frac{1}{\nu^2} I. \tag{EC.13}$$

Therefore,

$$\frac{d\mathrm{w}_Z}{d\nu} = \frac{1}{\nu^2} Z^\top \hat{\Sigma} \frac{e}{N}, \tag{EC.14}$$

and thus,

$$\frac{d\mathrm{w}}{d\nu} = \frac{1}{\nu^2} Z Z^\top \hat{\Sigma} \frac{e}{N}. \tag{EC.15}$$

By equation (EC.5), we know that the right-hand side in (EC.15) is proportional to the first conjugate portfolio. Moreover, by definition, the first conjugate portfolio is proportional to the difference between the first partial minimum-variance portfolio and the $1/N$ portfolio.

■

We now give a brief mathematical description of the algorithm followed to compute the partial minimum-variance portfolios; for a rigorous treatment of the conjugate-gradient method see Nocedal and Wright (1999, Chapter 5). Let w_{PARk} be the k^{th} partial minimum-variance portfolio and let $\epsilon^k = -Z^\top \hat{\Sigma} \mathrm{w}_{PARk}$ denote the residual at the k^{th} partial minimum-variance portfolio. Select the number of conjugate portfolios $K \leq N$, and let $\mathrm{w}_{PAR0} = e/N$ and $k = 0$. The $(k+1)$th partial minimum-variance portfolio is defined as $\mathrm{w}_{PARk+1} = \mathrm{w}_{PARk} + \alpha_k \mathrm{w}_{CG(k+1)}$, in which $\mathrm{w}_{CG(k+1)} = Z\epsilon^k + \beta^k \mathrm{w}_{CGk}$ is the $(k+1)^{\text{th}}$ conjugate portfolio; for $k = 0$ we set $\beta^k = 0$ and for $k > 0$ we set $\beta^k = \frac{(\epsilon^k)^\top \epsilon^k}{(\epsilon^{k-1})^\top \epsilon^{k-1}}$, and the step-size is $\alpha_k = \frac{(\epsilon_k)^\top \epsilon^k}{(\mathrm{w}_{CG(k+1)})^\top \hat{\Sigma} \mathrm{w}_{CG(k+1)}}$.

[19] This follows from basic penalty parameter theory (see Nocedal and Wright (1999, Chapter 17)).