To: High Dimensional Workshop, University of Connecticut

From: Nick Gunther

Subject: Geometry Of Minimum Variance Portfolio Estimation Error

Date: August 2, 2024

---

1 Introduction

1.1 Summary

This note investigates the changes in the solutions to a quadratic optimization program subject to a linear equality constraint resulting from a series of equally-spaced small rotational changes in the quadratic objective function. It shows how the magnitude and direction of the solution changes are not themselves equally-spaced, but depend heavily on the orientation of the objective function's level sets relative to the constraint.

The motivating use case, from finance, is the estimation of a fully-invested minimum variance portfolio of risky assets.

1.2 Preliminaries

Consider a positive definite symmetric $n \times n$ covariance matrix $\mathbb{M}$. For concreteness, we may take as an example $\mathbb{M} \equiv$ the return covariance matrix for portfolios in n securities, with r_i the dollar return on the investment x_i in security s_i. We then associated a point $x \in \mathbb{R}^n$ with a portfolio of total size (or market value) Σx_i. For similarity with the traditional business school framework of percentage returns and portfolio weights, we may take for analysis portfolios of size $1, so the $ investments translate easily into percentage weights. The hyperplane $\mathbb{H}_1 \equiv \{x|\Sigma x_i = 1\}$ represents such portfolios.

The traditional framework's main benefit is that by dispensing with the absolute size of investments, both return and risk for all investments, individual securities and portfolios alike, can be graphed in two dimensions. However, the price of this simplicity is projection and the associated loss of some interesting geometry, as we show below. By contrast, the price of that geometry is representing portfolios as points in $\mathbb{R}^n$, with n potentially large.

Instead of representing risk on a one-dimensional x-axis as in the traditional framework, we represent risk by variance, shown by n-1 dimensional ellipsoidal contours in $\mathbb{R}^n$, equivalently, isovariance *ellipsoids*,[1] given by $\{x|x'\mathbb{M}x = c\}$, c a constant. Noting that scaling a portfolio by k increases its variance by k^2, we see that the variances of portfolios with size $1 are represented by the intersections of all such ellipsoidal contours above some threshold. That threshold is identified with the smallest contour that intersects $\mathbb{H}_1$, which is the unique ellipsoid tangent to $\mathbb{H}_1$.[2] Note that $\mathbb{H}_1 \equiv \{x|e'x = 1\}$, where e is the vector with all its entries $= 1$.[3]

[1] Although the traditional framework is often associated with Harry Markowitz, the inventor of modern mean-variance financial analysis, his early work uses the same isovariance ellipsoids as we discuss in this note. See, e.g., Markowitz, *Portfolio Selection*, J. Finance (Mar. 1952) 84-86.

[2] This analysis assumes the matrix $\mathbb{M}$ is the true covariance matrix, and that the estimates of $\mathbb{M}$ discussed later are "good", in particular unbiased, estimates. Obtaining such "good" estimates is discussed elsewhere, such as in the papers of Goldberg, Shkolnik, Kercheval et al. listed in the Bibliography at the end.

[3] Although we work exclusively with this "full investment" constraint, other constraints could be handled similarly.

2 Analysis

2.1 Orthogonal Transformation

In our framework, the hyperplane $\mathbb{H}_1$ is simple (like all hyperplanes, it may be identified with its normal vector) but the ellipsoids are complex, representing as they do level contours for a random, and thus potentially arbitrary, positive definite symmetric (covariance) matrix. We therefore simplify $\mathbb{M}$ by diagonalizing it via an orthogonal transformation.[4] In the transformed coordinate system the matrix $\mathbb{M}$ becomes a diagonal matrix of eigenvalues $\equiv \Lambda$ while $\mathbb{H}_1$ transforms to a hyperplane $\mathbb{H}_m$ which we identify with its normal vector, no longer e but instead an arbitrary (random) vector $m|m'm = n$.

Figure 1 shows a conceptual visualization of a representative transformation in two dimensions:

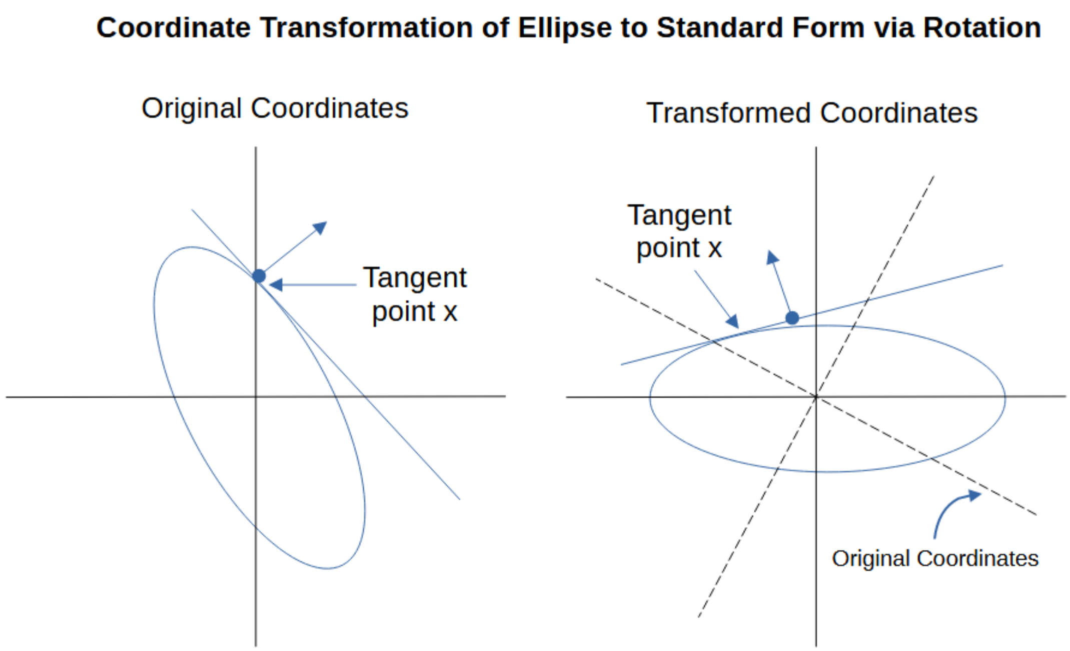

Figure 1: Conceptual Graph of Rotational Transformation to Standard Form

This transformation reduces our problem from studying a standard hyperplane tangent to a random ellipse, to studying a random hyperplane tangent to an ellipse in standard form. In the transformed coordinates, the matrix $\mathbb{M}$ is always the same diagonal matrix Λ, while the constraint transforms. Referring to our motivating use case, the optimal portfolio in transformed coordinates is the minimum variance portfolio satisfying the transformed, but still linear, constraint. The associated portfolio weights (and thus the optimal portfolio) will naturally differ in the original coordinate system, where the constraint is always that of full investment but the covariance matrix orientation can be arbitrary. The orthogonal transformation between the coordinate systems converts one portfolio's weights to those of the other. Naturally, the variances of the original and transformed portfolios and isovariance ellipsoids are the same; if O is the diagonalizing orthogonal transformation, $x'(O'\Lambda O)x = (Ox)'\Lambda Ox$.

[4]In general, there may be many such transformations; we choose one (unique up to sign) that transforms the ellipsoid axes to the standard axes in $\mathbb{R}^n$ given by $\{e_i, i = 1...n\}$ in decreasing order of axis length.

Except where otherwise stated, we will work in the transformed coordinate system.

2.2 Tangent Point

In the transformed coordinate system, denote by $\mathbb{L}_m$ the tangent ellipsoid to $\mathbb{H}_m$. Let x_m be the point at which $\mathbb{L}_m$ and $\mathbb{H}_m$ are tangent and let n be the normal to (the tangent to) $\mathbb{L}_m$ at x_m. We have the following three equations linking x_m, n and Λ:

$$
\begin{aligned}
2\Lambda x_m &= n && (1)\\
n &= km && (2)\\
m' x_m &= 1 && (3)
\end{aligned}
$$

Equation (1) follows from standard matrix differentiation and Equation (2) says that at x_m the normal to $\mathbb{L}_m$ is a multiple of the vector m, the transform of the vector $e \equiv \mathbf{1}$, which is normal to $\mathbb{H}_m \equiv$ the transformed constraint line, while the Equation (3) says that x_m is on $\mathbb{H}_m$, and follows because an orthogonal transformation is an isometry.

From Equation (1) and (2):

$$x_m = \frac{k}{2}\Lambda^{-1} m \tag{4}$$

Plugging this into Equation (3):

$$
\begin{aligned}
&\frac{k}{2} m' \Lambda^{-1} m = 1\\
\implies\ & k = \frac{2}{m'\Lambda^{-1} m} && (5)
\end{aligned}
$$

From Equation (4) and (5) we derive the well-known expression for the portfolio weights x_m:

$$x_m = \frac{\Lambda^{-1} m}{m'\Lambda^{-1} m} \tag{6}$$

The portfolio x_m therefore represents the fully-invested *characteristic portfolio* $\equiv$ *cp*(m) associated to m.[5]

Note also that by pre-multiplying Equation (4) by $x_m'\Lambda$ we also have another expression relating k to the variance $\equiv V(m)$ of *cp*(m):

$$
\begin{aligned}
V(m) = x_m' \Lambda x_m &= \frac{k}{2} x_m' \Lambda \Lambda^{-1} m\\
&= \frac{k}{2} m' x_m = \frac{k}{2} \times 1 = \frac{k}{2}\\
&\implies k = 2V(m) && (7)
\end{aligned}
$$

Comparing with Equation (5), we also have the following expression for V(m):

[5] *See, eg.*, Grinold and Kahn, *Active Portfolio Management*, 2d ed., McGraw-Hill (1999) 28.

$$V(m) = \frac{1}{m'\Lambda^{-1}m} \tag{8}$$

leading from Equation (6) to the following expression for x_m in terms of m and V(m):

$$x_m = V(m)\Lambda^{-1}m \tag{9}$$

2.3 Example

We consider a computational example in $\mathbb{R}^2$ demonstrating in our framework how quadratic optimization introduces varying and asymmetric error, leading to bias.

Recall that the vector m denotes the vector **1** in the transformed coordinates that diagonalize the covariance matrix, and that $|m| = |\mathbf{1}| = \sqrt{2}$ because the transformation $\equiv \mathcal{T}$ is orthogonal.

Let $\theta \equiv \theta(\mathcal{T}) \equiv$ the angle of rotation for $\mathcal{T}$ that puts the (ground truth) isovariance ellipses into standard form: $x'\Lambda x = c$ with Λ diagonal. This rotation will transform the standard constraint hyperplane for full investment $\equiv \mathbb{H}_\mathbf{1}$ to a transformed hyperplane $\mathbb{H}_m$, where m is the normal vector **1** rotated by θ. We write: $m \equiv m(\theta) = \sqrt{2}(\alpha, \bar{\alpha})', \alpha = cos(\theta), \bar{\alpha} = \pm\sqrt{1-\alpha^2}$. Note that $m(\frac{\pi}{4}) \iff \alpha = \frac{1}{\sqrt{2}}$ corresponds to the standard full investment constraint for two securities $m(\frac{\pi}{4}) = \mathbf{1}$.

We also set $\Lambda \equiv \Lambda_Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & Z \end{bmatrix}, Z > 1 \implies \Lambda_Z^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{Z} \end{bmatrix}$

From Equation (8), after some manipulation:

$$V(m(\alpha)) = \frac{1}{2}\frac{1}{\alpha^2 + \bar{\alpha}^2/Z} \tag{10}$$

2.4 Three Rotations: $\theta = 0, \pm\frac{\pi}{4}$,

Figure 2 below shows on the transformed axes[6] the isovariance ellipses specified by Λ_{10} for three different orientations of the constraint. We choose Λ_{10} because it corresponds to eigenvalues that differ by an order of magnitude. The three orientations for the hyperplane normal relative to the ellipse axes are as follows, describing each orientation by the indicated α:

$$\alpha_1 = 1 \implies \bar{\alpha}_1 = 0$$
$$\alpha_2 = \frac{1}{\sqrt{2}} \implies \bar{\alpha}_2 = \frac{1}{\sqrt{2}}$$
$$\alpha_3 = 0 \implies \bar{\alpha}_3 = 1$$

Thus:

1. α_1 represents the tangent normal $m_1 = (\sqrt{2}, 0)$ and a rotation of the constraint line of $-\frac{\pi}{4}$,
2. α_2 represents the tangent normal $m_2 = (1, 1)$ and no rotation, while
3. α_3 represents the tangent normal $m_3 = (0, \sqrt{2})$ and a rotation of the constraint line of $\frac{\pi}{4}$.

[6] The axes that standardize the ellipses and diagonalize the associated matrix.

All the m fall on the circle with radius $\sqrt{2}$

The three associated constraint hyperplanes $\mathbb{H}_{m_1}, \mathbb{H}_{m_2}, \mathbb{H}_{m_3}$ are tangent to the circle (displayed in purple in Figure 2) with radius $\frac{1}{\sqrt{2}}$.[7] As shown in Figure 2, the minimum variance characteristic portfolios correspond to:

1. $\theta = \pm\frac{\pi}{4}$, equal investments in the two securities corresponding to the original (pre-transformation) xy axes. (Note that the original coordinate axes are at angles of $\frac{\pi}{4}, \frac{3\pi}{4}$ to the Figure 2 axes for these two values of θ. However, the minimum variances differ by an order of magnitude between the two values for θ.)

2. $\theta = 0$, a 91% investment in the security corresponding to the x-axis (there is no rotation) and a 9% investment in the security corresponding to the y-axis. This is intuitive; the first security is significantly less risky and the two securities are uncorrelated. In this case, the original axes and the Figure 2 axes coincide. The minimum variance of the fully-invested portfolio is 0.91

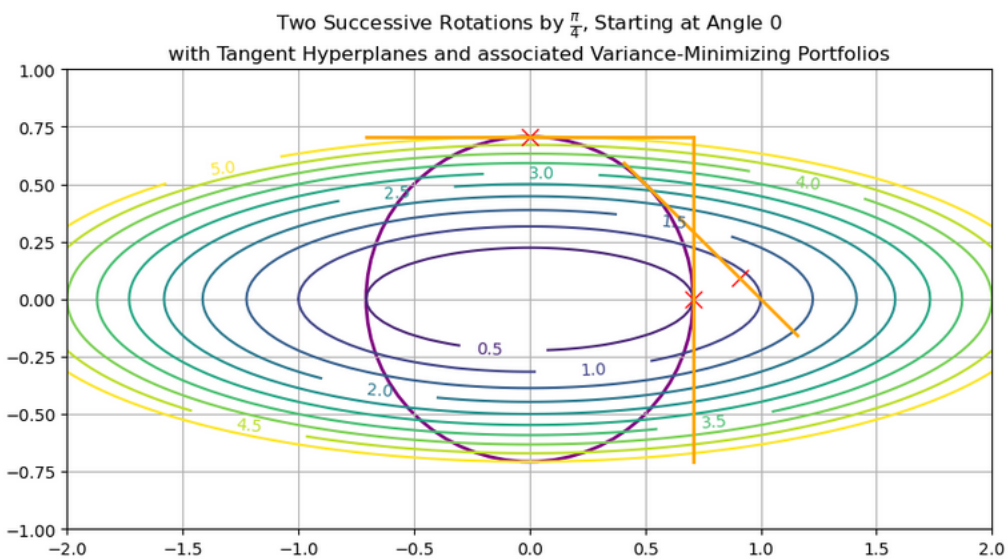

Figure 2: Three Rotations of Equal Size (Three Ellipse Axes Orientations)

The curvature of the ellipse at m_3 is a low 0.1[8]. At a point where ellipsoid curvature is low, the slope of the ellipsoid tangent changes little even in a large neighborhood of that point. Since the slope is determined by m, even a small error in m can move the tangent point (which corresponds to the portfolio weights) outside of that neighborhood, resulting in a large error in the portfolio weights and associated variance.[9]

As m changes from m_2 to m_3 in Figure 2, the minimum variance changes from 0.91 to 5, an increase by a factor of 5.5

[7] Cf. with the standard $m = (1, 1)$, the closest point to the origin = (0.5,0.5) at a distance of $\frac{1}{\sqrt{2}}$.

[8] There are many ways to compute curvature, including parameterizing the ellipse as $(\sqrt{10}rcos(\theta), rsin(\theta))$ and using the formula $curvature = \frac{|x'y''-y'x''|}{(x'^2+y'^2)^{\frac{3}{2}}}$

[9] Slightly more rigorously, low curvature implies that a set of errors in m of small measure can correspond to a set of errors of large measure in the tangent point.

By contrast, the curvature of the ellipse at m_1 is 1, 10 times greater than at m_3; correspondingly, as m changes from m_1 to m_2, the minimum variance changes from 0.5 to 0.91, an increase by a factor of only 1.8

Curvature measures the rate of change of the tangent slope to the isovariance ellipsoid; if the tangent slope at a particular m changes rapidly, small errors in m will not result in large changes in the tangent point, the portfolio weights and the variance, since even nearby points will have a significantly different tangent slope. The curvature variation over the ellipsoid depends on the relative sizes of the ellipsoid axes, that is to say, the ellipsoid's *eccentricity* or *aspect ratio*, which in turn depends on the dispersion of the eigenvalues.

We therefore see that the effect of eigenvector estimation error on estimated minimum variance portfolio weights error depends on the ground truth; more specifically, the magnitude of the probable error in portfolio weights depends on the curvature of the isovariance ellipsoid at the ground truth tangency point, which in turn depends on eigenvalue dispersion. Since the constraint imposes a tangency condition, this conclusion is intuitive - the curvature is the rate of change in tangency slope, and thus as discussed above is inversely related to the rate of change in the location of the tangent point. Moreover, because for ellipsoids the curvature itself changes from point to point, the portfolio weights and variance estimates both exhibit bias - eigenvector direction errors in one direction will result in consistently larger tangency point errors than eigenvector direction errors in the opposite direction.

As Figure 3 indicates, a finer rotation grid shows more clearly the path of minimum variance portfolios subject to the changing isovariance ellipse axes orientation relative to the constraint. Figure 4 shows on the original (not transformed) axes the sequence of portfolio weights from ten equally-sized rotations from 0 to $\frac{\pi}{2}$. There are only four portfolios with short positions, and the largest short position is slightly over 20%. The first and last portfolio weights are equal; we have jittered the last slightly to avoid occlusion. Note that the change in portfolio weights from rotation 10 to rotation 11 is particularly large.

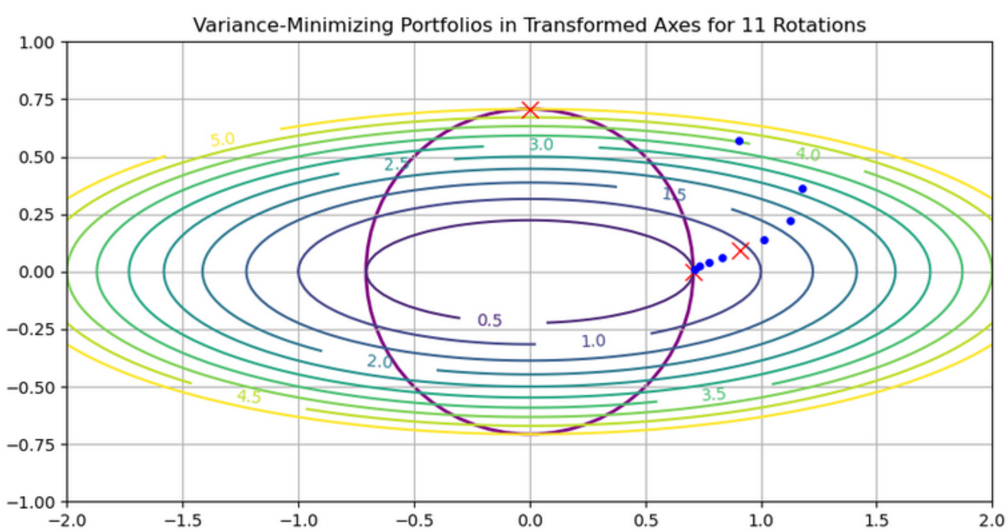

Figure 3: Portfolio Path over Ten Equally-Sized Rotations (Eleven Ellipse Axes Orientations)

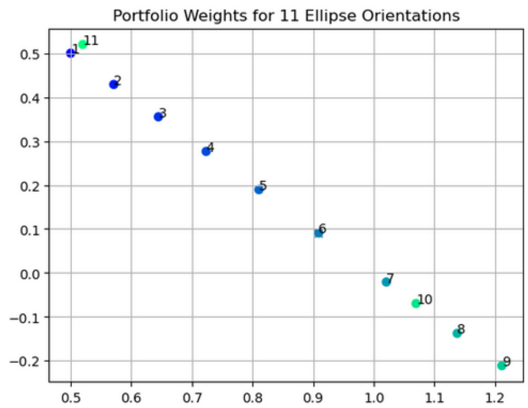

Figure 4: Variance-Minimizing Portfolio Weights Graphed on Original (True) Axes

References

[1] Lisa R. Goldberg, Hubeyb Gurdogan, and Alec Kercheval. "Portfolio optimization via strategy-specific eigenvector shrinkage". Aug. 2024.

[2] Lisa R. Goldberg and Alec N. Kercheval. "James-Stein for the leading eigenvector". In: *Proceedings of the National Academy of Sciences* 120.2 (2023).

[3] Lisa R. Goldberg, Alex Papanicolau, and Alex Shkolnik. "The Dispersion Bias". In: *SIAM Journal of Financial Mathematics* 13.2 (2022), pp. 521–550.

[4] Lisa R. Goldberg et al. "Better betas". In: *The Journal of Portfolio Management* 47.1 (2020), pp. 119–136.

[5] H. Gurdogan. "Eigenvector shrinkage for estimating covariance matrices". In: *PhD thesis, Florida State University* (2021).

[6] Hubeyb Gurdogan and Alec Kercheval. *Multi anchor point shrinkage for the sample covariance matrix*. Tech. rep. working paper, 2021.

[7] Alex Shkolnik. "James-Stein estimation of the first principal component". In: *Stat* 11.1 (2022).

[8] Alex Shkolnik et al. "Portfolio Selection Revisited". 2024.