Portfolios that Contain Risky Assets

Portfolio Models 3.

Appendix: Covariance Matrix

C. David Levermore and Paul Smith
University of Maryland, College Park

Math 420: *Mathematical Modeling*
February 21, 2016 version

Portfolio Models
1. Risk and Reward
2. Markowitz Portfolios
3. Basic Markowitz Portfolio Theory
4. Portfolios with Risk-Free Assets
5. Long Portfolios
6. Long Portfolios with a Safe Investment

Stochastic Models
1. Independent, Identically-Distributed Models
2. Law of Large Numbers (Kelly) Objectives
3. Growth Rate Estimators
4. Central Limit Theorem Objectives
5. Optimization of Mean-Variance Objectives
6. Utility Function Objectives

Portfolio Models. Covariance Matrix

1. Introduction
2. Symmetry and Definiteness
3. Eigenpairs and Diagonalization
4. Solving Linear Algebraic Systems
5. Simple Example Considered for Large N
6. Conclusion

Portfolio Models. Covariance Matrix

Introduction. Suppose that we are considering the return rate histories $\{r_i(d)\}_{d=1}^{D}$ for assets $i = 1, \cdots, N$ over a period of D trading days and assign day d a weight $w(d) > 0$ such that the weights $\{w(d)\}_{d=1}^{D}$ satisfy

$$\sum_{d=1}^{D} w(d) = 1\,.$$

Then the return rate means and covariances are given by

$$m_i = \sum_{d=1}^{D} w(d)\, r_i(d)\,,$$

$$v_{ij} = \sum_{d=1}^{D} w(d) \Big(r_i(d) - m_i\Big)\Big(r_j(d) - m_j\Big)\,.$$

The return rate history can be expressed as $\{\mathbf{r}(d)\}_{d=1}^{D}$ where

$$\mathbf{r}(d) = \begin{pmatrix} r_1(d) \\ \vdots \\ r_N(d) \end{pmatrix} .$$

Then the N-vector of return rate means $\mathbf{m}$ and the $N \times N$-matrix of return rate covariances $\mathbf{V}$ can be expressed as

$$\mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix} = \sum_{d=1}^{D} w(d)\, \mathbf{r}(d) \,,$$

$$\mathbf{V} = \begin{pmatrix} v_{11} & \cdots & v_{1N} \\ \vdots & \ddots & \vdots \\ v_{N1} & \cdots & v_{NN} \end{pmatrix} = \sum_{d=1}^{D} w(d) \Big(\mathbf{r}(d) - \mathbf{m}\Big) \Big(\mathbf{r}(d) - \mathbf{m}\Big)^{\top} .$$

Here we give some properties of $\mathbf{V}$ that will be used to extract statistical information from it.

Symmetry and Definiteness. The most important properties of $\mathbf{V}$ are that it is always symmetric and that it is almost always positive definite. These properties are taught in elementary linear algebra courses, but are so important that we review them here.

Definition 1. A real $N \times N$-matrix $\mathbf{A}$ is said to be symmetric if $\mathbf{A}^\mathsf{T} = \mathbf{A}$. It is said to be nonnegative definite if

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq 0 \quad \text{for every } \mathbf{x} \in \mathbb{R}^N .$$

It is said to be positive definite if

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} > 0 \quad \text{for every nonzero } \mathbf{x} \in \mathbb{R}^N .$$

Remarks. Clearly, every positive definite matrix is nonnegative definite. A nonnegative matrix is positive definite if and only if

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = 0 \quad \Longrightarrow \quad \mathbf{x} = \mathbf{0} . \tag{1}$$

Fact 1. The covariance matrix $\mathbf{V}$ is symmetric.

Proof. Exercise.

Fact 2. The covariance matrix $\mathbf{V}$ is nonegative definite.

Proof. Let $\mathbf{x} \in \mathbb{R}^N$ be arbitrary. Then

$$\begin{aligned}
\mathbf{x}^\mathsf{T}\mathbf{V}\mathbf{x} &= \mathbf{x}^\mathsf{T}\left(\sum_{d=1}^{D} w(d)\left(\mathbf{r}(d)-\mathbf{m}\right)\left(\mathbf{r}(d)-\mathbf{m}\right)^\mathsf{T}\right)\mathbf{x} \\
&= \sum_{d=1}^{D} w(d)\,\mathbf{x}^\mathsf{T}\left(\mathbf{r}(d)-\mathbf{m}\right)\left(\mathbf{r}(d)-\mathbf{m}\right)^\mathsf{T}\mathbf{x} \\
&= \sum_{d=1}^{D} w(d)\left(\left(\mathbf{r}(d)-\mathbf{m}\right)^\mathsf{T}\mathbf{x}\right)^2 \geq 0\,.
\end{aligned}$$

Fact 3. The covariance matrix $\mathbf{V}$ is positive definite if and only if the vectors $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ span $\mathbb{R}^N$.

Proof. Because $w(d) > 0$, the calculation in the previous proof shows that $\mathbf{x}^\mathsf{T}\mathbf{V}\mathbf{x} = 0$ if and only if

$$\Big(\mathbf{r}(d) - \mathbf{m}\Big)^\mathsf{T}\mathbf{x} = 0 \quad \text{for every } d = 1, \cdots, D\,. \tag{2}$$

First, suppose that $\mathbf{V}$ is not positive definite. Then by (1) there exists an $\mathbf{x} \in \mathbb{R}^N$ such that $\mathbf{x}^\mathsf{T}\mathbf{V}\mathbf{x} = 0$ and $\mathbf{x} \neq \mathbf{0}$. This implies by (2) that the vectors $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ lie in the hyperplane orthogonal (normal) to $\mathbf{x}$. Therefore the vectors $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ do not span $\mathbb{R}^N$.

Conversely, suppose that the vectors $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ do not span $\mathbb{R}^N$. Then there must be a nonzero vector $\mathbf{x}$ that is orthogonal to their span. This implies that $\mathbf{x}$ satisfies (2), whereby $\mathbf{x}^\mathsf{T}\mathbf{V}\mathbf{x} = 0$. Therefore $\mathbf{V}$ is not positive definite by (1). □

Remark. The set of vectors $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ can span $\mathbb{R}^N$ only if $D \geq N$. Therefore we require that $D \geq N$.

Remark. In practice D will be much larger than N. In the homework and projects for this course usually $N \leq 10$ while $D \geq 42$ (often $D = 252$). When D is so much greater than N the covariance matrix $\mathbf{V}$ will almost always be positive definite.

Remark. If $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ spans $\mathbb{R}^N$ then $\{\mathbf{r}(d)\}_{d=1}^{D}$ also spans $\mathbb{R}^N$. However, the converse need not hold. A counterexample for $N = 2$ and any $D \geq 2$ can be constructed as follows. Let $\{\mathbf{m}, \mathbf{n}\}$ span $\mathbb{R}^2$. Let $\mathbf{r}(d) = \mathbf{m} + h(d)\mathbf{n}$ where $h(d) \neq 0$ and

$$\sum_{d=1}^{D} w(d)h(d) = 0\,.$$

Then $\{\mathbf{r}(d)\}_{d=1}^{D}$ spans $\mathbb{R}^2$ while $\{\mathbf{r}(d) - \mathbf{m}\}_{d=1}^{D}$ does not span $\mathbb{R}^2$.

Eigenpairs and Diagonalization. Let us recall from linear algebra that an *eigenpair* $(\lambda, \mathbf{q})$ of a real $N \times N$ matrix $\mathbf{A}$ is a scalar λ (possibly complex) and a nonzero vector $\mathbf{q}$ (possibly with complex entries) such that

$$\mathbf{A}\mathbf{q} = \lambda \mathbf{q}\,. \tag{3}$$

An eigenpair is called a real eigenpair when λ and every entry of $\mathbf{q}$ is real.

An important fact from linear algebra is that if $\mathbf{A}$ is symmetric then it has N real eigenpairs

$$(\lambda_1, \mathbf{q}_1)\,, \quad (\lambda_2, \mathbf{q}_2)\,, \quad \cdots \quad (\lambda_N, \mathbf{q}_N)\,, \tag{4}$$

such that the eigenvectors $\{\mathbf{q}_i\}_{i=1}^N$ are an orthonormal set. This means that they satisfy the orthonormality conditions

$$\mathbf{q}_i^T \mathbf{q}_j = \delta_{ij}\,. \tag{5}$$

Because the $\{\mathbf{q}_i\}_{i=1}^N$ satisfy the orthonormality conditions (5), they form an orthonormal basis of $\mathbb{R}^N$. Every $\mathbf{x} \in \mathbb{R}^N$ can be expanded as

$$\mathbf{x} = \mathbf{q}_1\,\mathbf{q}_1^\mathsf{T}\mathbf{x} + \mathbf{q}_2\,\mathbf{q}_2^\mathsf{T}\mathbf{x} + \cdots + \mathbf{q}_N\,\mathbf{q}_N^\mathsf{T}\mathbf{x}\,. \tag{6}$$

The numbers $\{\mathbf{q}_i^\mathsf{T}\mathbf{x}\}_{i=1}^N$ are called the coordinates of $\mathbf{x}$ for the orthonormal basis $\{\mathbf{q}_i\}_{i=1}^N$. The square of the Euclidean norm of $\mathbf{x}$ is given by

$$\|\mathbf{x}\|^2 = \mathbf{x}^\mathsf{T}\mathbf{x} = \left(\mathbf{q}_1^\mathsf{T}\mathbf{x}\right)^2 + \left(\mathbf{q}_2^\mathsf{T}\mathbf{x}\right)^2 + \cdots + \left(\mathbf{q}_N^\mathsf{T}\mathbf{x}\right)^2. \tag{7}$$

Because the $\{\mathbf{q}_i\}_{i=1}^N$ are eigenvectors of $\mathbf{A}$, we see from (6) that

$$\begin{aligned}\mathbf{A}\mathbf{x} &= \mathbf{A}\mathbf{q}_1\,\mathbf{q}_1^\mathsf{T}\mathbf{x} + \mathbf{A}\mathbf{q}_2\,\mathbf{q}_2^\mathsf{T}\mathbf{x} + \cdots + \mathbf{A}\mathbf{q}_N\,\mathbf{q}_N^\mathsf{T}\mathbf{x} \\ &= \lambda_1\mathbf{q}_1\,\mathbf{q}_1^\mathsf{T}\mathbf{x} + \lambda_2\mathbf{q}_2\,\mathbf{q}_2^\mathsf{T}\mathbf{x} + \cdots + \lambda_N\mathbf{q}_N\,\mathbf{q}_N^\mathsf{T}\mathbf{x}\,.\end{aligned} \tag{8}$$

Hence, the $\{\lambda_i\mathbf{q}_i^\mathsf{T}\mathbf{x}\}_{i=1}^N$ are the coordinates of $\mathbf{A}\mathbf{x}$ for the orthonormal basis $\{\mathbf{q}_i\}_{i=1}^N$. Therefore by (7) we have

$$\|\mathbf{A}\mathbf{x}\|^2 = \lambda_1^2\left(\mathbf{q}_1^\mathsf{T}\mathbf{x}\right)^2 + \lambda_2^2\left(\mathbf{q}_2^\mathsf{T}\mathbf{x}\right)^2 + \cdots + \lambda_N^2\left(\mathbf{q}_N^\mathsf{T}\mathbf{x}\right)^2. \tag{9}$$

Moreover, $\mathbf{A}$ can be expressed in the factored form $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}}$ where $\mathbf{D}$ and $\mathbf{Q}$ are the real $N \times N$ matrices constructed from the eigenpairs (4) as

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_N \end{pmatrix}, \qquad \mathbf{Q} = \begin{pmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_N \end{pmatrix}.$$

Because the matrix $\mathbf{D}$ is a *diagonal matrix*, this factorization of $\mathbf{A}$ is called diagonalization. The orthonormality conditions (5) satisfied by the vectors $\{\mathbf{q}_i\}_{i=1}^{N}$ imply that $\mathbf{Q}$ is an orthogonal matrix. This means that $\mathbf{Q}$ satisfies

$$\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{I} = \mathbf{Q}\mathbf{Q}^{\mathsf{T}}.$$

The relation $\mathbf{Q}^{\mathsf{T}}\mathbf{Q} = \mathbf{I}$ is a recasting of the orthonormality conditions (5). The relation $\mathbf{I} = \mathbf{Q}\mathbf{Q}^{\mathsf{T}}$ is equivalent to $\mathbf{x} = \mathbf{Q}\mathbf{Q}^{\mathsf{T}}\mathbf{x}$, which is a recasting of expansion (6). These relations show that $\mathbf{Q}$ and $\mathbf{Q}^{\mathsf{T}}$ are inverses of each other — i.e. that $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}$ and that $\mathbf{Q}^{-\mathsf{T}} = \mathbf{Q}$.

Other important facts from linear algebra are that if $\mathbf{A}$ is a real symmetric matrix then:

- it is nonnegative definite if and only if all its eigenvalues are nonnegative;

- it is positive definite if and only if all its eigenvalues are positive.

Proof. The ($\Longrightarrow$) directions of these characterizations follow from the fact that if $(\lambda, \mathbf{q})$ is an eigenpair of $\mathbf{A}$ that is normalized so that $\mathbf{q}^\mathsf{T}\mathbf{q} = 1$ then

$$\lambda = \lambda \mathbf{q}^\mathsf{T}\mathbf{q} = \mathbf{q}^\mathsf{T}(\lambda \mathbf{q}) = \mathbf{q}^\mathsf{T}(\mathbf{A}\mathbf{q}) = \mathbf{q}^\mathsf{T}\mathbf{A}\mathbf{q}\,.$$

The ($\Longleftarrow$) directions of these characterizations use the full power of the orthonormality conditions (5) as embodied by expansion (8),

$$\mathbf{A}\mathbf{x} = \lambda_1 \mathbf{q}_1 \, \mathbf{q}_1^{\mathsf{T}} \mathbf{x} + \lambda_2 \mathbf{q}_2 \, \mathbf{q}_2^{\mathsf{T}} \mathbf{x} + \cdots + \lambda_N \mathbf{q}_N \, \mathbf{q}_N^{\mathsf{T}} \mathbf{x} \, .$$

By taking the scalar product of this expansion with $\mathbf{x}$ we obtain

$$\mathbf{x}^{\mathsf{T}} \mathbf{A} \mathbf{x} = \lambda_1 \left(\mathbf{q}_1^{\mathsf{T}} \mathbf{x} \right)^2 + \lambda_2 \left(\mathbf{q}_2^{\mathsf{T}} \mathbf{x} \right)^2 + \cdots + \lambda_N \left(\mathbf{q}_N^{\mathsf{T}} \mathbf{x} \right)^2 .$$

It is thereby clear that:

- if $\lambda_i \geq 0$ for every $i = 1, \cdots, N$ then $\mathbf{A}$ is nonnegative definite;

- if $\lambda_i > 0$ for every $i = 1, \cdots, N$ then $\mathbf{A}$ is positive definite.

This proves the ($\Longleftarrow$) directions of the characterizations. $\square$

Solving Linear Algebraic Systems. In order to compute the frontier Markowitz portfolios when $\mathbf{V}$ is positive definite, we must solve the linear algebraic systems

$$\mathbf{V}\mathbf{y} = \mathbf{1}\,, \qquad \mathbf{V}\mathbf{z} = \mathbf{m}\,.$$

In principle this task can be done with the Matlab command backslash. However, in practice things may not be so simple because the matrix $\mathbf{V}$ can be ill-conditioned. This concept is usually presented in numerical analysis courses. Here we review it.

Consider an $N \times N$ real linear algebraic system in the form

$$\mathbf{A}\mathbf{x} = \mathbf{b}\,, \quad \text{where } \mathbf{A} \text{ is symmetric and positive definite}\,. \tag{10}$$

We say that $\mathbf{A}$ is ill-conditioned for a given numerical method for solving system (10) when small changes in $\mathbf{b}$ can lead to significant changes in the solution $\mathbf{x}$.

The conditioning of the symmetric, positive definite matrix $\mathbf{A}$ is measured by its so-called condition number, which is defined as

$$\kappa(\mathbf{A}) = \frac{\lambda_{\max}(\mathbf{A})}{\lambda_{\min}(\mathbf{A})}, \tag{11}$$

where

$$\begin{aligned} \lambda_{\max}(\mathbf{A}) &= \max\left\{\lambda \,:\, \lambda \text{ is an eigenvalue of } \mathbf{A}\right\}, \\ \lambda_{\min}(\mathbf{A}) &= \min\left\{\lambda \,:\, \lambda \text{ is an eigenvalue of } \mathbf{A}\right\}. \end{aligned} \tag{12}$$

Notice that $\kappa(\mathbf{A}) \geq 1$. Notice too that $\kappa(\mathbf{A}) = 1$ if and only if $\mathbf{A}$ is a positive multiple of $\mathbf{I}$. As $\kappa(\mathbf{A})$ increases, $\mathbf{A}$ becomes less well-conditioned.

The condition number arises when considering how small errors in $\mathbf{b}$ can lead to errors in the solution $\mathbf{x}$ of system (10). Let $\mathbf{x}$ and $\mathbf{x} + \tilde{\mathbf{x}}$ solve

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \qquad \mathbf{A}(\mathbf{x} + \tilde{\mathbf{x}}) = \mathbf{b} + \tilde{\mathbf{b}}.$$

Here $\tilde{\mathbf{x}}$ is the error in the solution of system due to the error $\tilde{\mathbf{b}}$ in the forcing.

We want to bound the relative error of the solution, $\|\tilde{\mathbf{x}}\|/\|\mathbf{x}\|$, in terms of the relative error of the forcing, $\|\tilde{\mathbf{b}}\|/\|\mathbf{b}\|$. By linearity we have

$$\mathbf{A}\mathbf{x} = \mathbf{b}\,, \qquad \mathbf{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}\,.$$

We see from (9) and from definitions (12) that

$$\|\mathbf{b}\| = \|\mathbf{A}\mathbf{x}\| \le \lambda_{\mathsf{max}}(\mathbf{A})\|\mathbf{x}\|\,, \qquad \|\tilde{\mathbf{b}}\| = \|\mathbf{A}\tilde{\mathbf{x}}\| \ge \lambda_{\mathsf{min}}(\mathbf{A})\|\tilde{\mathbf{x}}\|\,.$$

These inequalities imply that

$$\frac{\|\tilde{\mathbf{x}}\|}{\|\mathbf{x}\|} \le \frac{\lambda_{\mathsf{max}}(\mathbf{A})}{\lambda_{\mathsf{min}}(\mathbf{A})}\,\frac{\|\tilde{\mathbf{b}}\|}{\|\mathbf{b}\|} = \kappa(\mathbf{A})\,\frac{\|\tilde{\mathbf{b}}\|}{\|\mathbf{b}\|}\,.$$

Therefore $\kappa(\mathbf{A})$ gives an upper bound on the factor by which relative errors can increase when solving system (10) numerically.

Many numerical methods for solving such a system are associated with some factorization of the matrix $\mathbf{A}$.

For example, consider the diagonalization $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}}$, where $\mathbf{Q}$ is an orthogonal matrix and $\mathbf{D}$ is a diagonal matrix with positive diagonal entries. Because $\mathbf{Q}$ is orthogonal we have $\mathbf{Q}^{-1} = \mathbf{Q}^{\mathsf{T}}$ and $\mathbf{Q}^{-\mathsf{T}} = \mathbf{Q}$, whereby

$$\mathbf{A}^{-1} = \left(\mathbf{Q}\mathbf{D}\mathbf{Q}^{\mathsf{T}}\right)^{-1} = \mathbf{Q}^{-\mathsf{T}}\mathbf{D}^{-1}\mathbf{Q}^{-1} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{Q}^{\mathsf{T}}\,.$$

This factorization suggests solving system (10) by the recipe

$$\mathbf{c} = \mathbf{Q}^{\mathsf{T}}\mathbf{b}\,, \qquad \mathbf{D}\mathbf{y} = \mathbf{c}\,, \qquad \mathbf{x} = \mathbf{Q}\mathbf{y}\,. \tag{13}$$

The difficulty arises because computers perform approximate arithmetic, which means that the solutions to these systems are approximate. Small errors in $\mathbf{c}$ can create large errors in $\mathbf{y}$, which will lead to large errors in $\mathbf{x}$. To see why, suppose that

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_N \end{pmatrix}, \quad \text{where } \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N > 0\,.$$

Suppose that the result of the calculation $\mathbf{Q}^{\mathsf{T}}\mathbf{b}$ is $\mathbf{c}+\tilde{\mathbf{c}}$ where $\tilde{\mathbf{c}}$ is much smaller than $\mathbf{c}$. Then the result of solving $\mathbf{D}\mathbf{y}=\mathbf{c}$ with $\mathbf{c}+\tilde{\mathbf{c}}$ replacing $\mathbf{c}$ is $\mathbf{y}+\tilde{\mathbf{y}}$ where

$$\mathbf{y}=\begin{pmatrix} y_1 & y_2 & \cdots & y_N \end{pmatrix}^{\mathsf{T}} = \mathbf{D}^{-1}\mathbf{c} = \begin{pmatrix} \frac{c_1}{\lambda_1} & \frac{c_2}{\lambda_2} & \cdots & \frac{c_N}{\lambda_N} \end{pmatrix}^{\mathsf{T}},$$

$$\tilde{\mathbf{y}}=\begin{pmatrix} \tilde{y}_1 & \tilde{y}_2 & \cdots & \tilde{y}_N \end{pmatrix}^{\mathsf{T}} \approx \mathbf{D}^{-1}\tilde{\mathbf{c}} = \begin{pmatrix} \frac{\tilde{c}_1}{\lambda_1} & \frac{\tilde{c}_2}{\lambda_2} & \cdots & \frac{\tilde{c}_N}{\lambda_N} \end{pmatrix}^{\mathsf{T}}.$$

If $\tilde{c}_1$ and $\tilde{c}_N$ are comparable in size and $\lambda_1 \gg \lambda_N$ then $\tilde{y}_N$ will be much larger than $\tilde{y}_1$ by the ratio λ_1/λ_N. When this ratio is large enough then $\tilde{y}_N \approx \tilde{c}_N/\lambda_N$ will be greater than $y_1 = c_1/\lambda_1$ and maybe greater than other entries of the exact solution $\mathbf{y}$.

Because the result of $\mathbf{x} = \mathbf{Q}\mathbf{y}$ with $\mathbf{y} + \tilde{\mathbf{y}}$ replacing $\mathbf{y}$ is $\mathbf{x} + \tilde{\mathbf{x}}$ where

$$\mathbf{x} + \tilde{\mathbf{x}} \approx \mathbf{Q}(\mathbf{y} + \tilde{\mathbf{y}}) = \mathbf{x} + \mathbf{Q}\tilde{\mathbf{y}}\,,$$

the vector $\tilde{\mathbf{y}}$ will taint the exact solution because some entries of $\tilde{\mathbf{y}}$ are larger than some entries of $\mathbf{y}$ and multiplication by $\mathbf{Q}$ generally will mix the entries of $\mathbf{y} + \tilde{\mathbf{y}}$.

Simple Example Considered for Large N. Consider the family of real $N \times N$ matrices $\mathbf{V}_c$ given by

$$\mathbf{V}_c = \begin{pmatrix} 1 & c & \cdots & c \\ c & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & c \\ c & \cdots & c & 1 \end{pmatrix} = (1-c)\mathbf{I} + c\mathbf{1}\mathbf{1}^{\mathsf{T}},$$

for some $-\frac{1}{N-1} < c < 1$. Notice $\mathbf{1}\mathbf{1}^{\mathsf{T}}$ is the $N \times N$ matrix with every entry equal to 1. It is clear that each $\mathbf{V}_c$ is symmetric.

Notice that $\mathbf{1}$ is an eigenvector of $\mathbf{V}_c$ with eigenvalue $1 + (N-1)c$.

Notice that every vector $\mathbf{q}$ that satisfies $\mathbf{1}^{\mathsf{T}}\mathbf{q} = 0$ is an eigenvector of $\mathbf{V}_c$ with eigenvalue $1-c$. Therefore the eigenvalue $1-c$ has multiplicity $N-1$.

Because $-\frac{1}{N-1} < c < 1$, both of these eigenvalues are positive. Hence, $\mathbf{V}_c$ is positive definite.

If $c \in (0,1)$ then $1+(N-1)c > 1-c > 0$, so that

$$\lambda_{\mathsf{max}}(\mathbf{V}_c) = 1+(N-1)c\,, \qquad \lambda_{\mathsf{min}}(\mathbf{V}_c) = 1-c\,.$$

Hence, the condition number of $\mathbf{V}_c$ is

$$\kappa(\mathbf{V}_c) = \frac{\lambda_{\mathsf{max}}(\mathbf{V}_c)}{\lambda_{\mathsf{min}}(\mathbf{V}_c)} = \frac{1+(N-1)c}{1-c}\,.$$

Notice that $\kappa(\mathbf{V}_c)$ grows linearly in N as N increases for fixed c.

Remark. The matrix $\mathbf{V}_c$ is the covariance matrix for N risky assets with identical variances equal to 1 and with identical covariances equal to c. Because the variances are all equal to 1 the matrix $\mathbf{V}_c$ is also the correlation matrix for these assets.

Conclusion. An important take-away message from this lecture is that even when $\mathbf{V}$ is symmetric and positive definite, it can be ill-conditioned — that is, it can have a large condition number. This becomes more likely for larger N. When this happens care must be taken to avoid inaccurate solutions of

$$\mathbf{V}\mathbf{y} = \mathbf{1}\,, \qquad \mathbf{V}\mathbf{z} = \mathbf{m}\,.$$

We will do this by decomposing $\mathbf{V}$ into the sum of a low rank matrix plus a well-conditioned matrix. We will show how this is done in a future lecture. We will also give a statistical interpretation of ill-conditioning.