Portfolios

Christopher Ting

http://www.mysmu.edu/faculty/christophert/

✉: christopherting@smu.edu.sg

☎: 6828 0364

: LKCSB 5036

Table of Contents

Pre-U Scalar and Vector

- A scalar a is a single number or one dimension.
  - $\mathfrak{N}$: Natural numbers
  - $\mathfrak{Z}$: Integers
  - $\mathfrak{Q}$: Rational numbers
  - $\mathfrak{R}$: Real numbers
- A vector $\boldsymbol{a}$ is a $k \times 1$ list of numbers arranged in a column.

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} \in \mathfrak{R}^k.$$

Pre-U Matrix

- A matrix $\boldsymbol{A}$ is a $k \times r$ rectangular array of numbers arranged in k rows and r columns.

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1r} \\ A_{21} & A_{22} & \cdots & A_{2r} \\ \vdots & \vdots & \vdots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kr} \end{bmatrix}$$

- Regular letters are used for scalars, lower case bold letters for vectors, and upper case bold letters for matrices.
- By convention, a_i is the element of a vector in the i-th row, and A_{ij} refers to the element of a matrix in the i-th row and j-th column.

Matrix Transpose

- The transpose of matrix is the matrix obtained by rotating each row into a column (clockwise) with the first element as pivot, i.e., for each row i,

$$\begin{bmatrix} A_{i1} & A_{i2} & \cdots & A_{ir} \end{bmatrix} \curvearrowright \begin{bmatrix} A_{1i} \\ A_{2i} \\ \vdots \\ A_{ri} \end{bmatrix}$$

- Notation-wise, we write the transpose of $\boldsymbol{A}$ as $\boldsymbol{A}^\top$ or $\boldsymbol{A}'$, and

$$\left[A_{ij}\right]^\top = \left[A_{ji}\right].$$

Question

If $\boldsymbol{A}$ is a $k \times r$ matrix, what is the length and width of $\boldsymbol{A}^\top$?

Matrix Addition

- Two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are addable only when they are of the same order (rows × column):

$$\boldsymbol{A} + \boldsymbol{B} = [A_{ij}] + [B_{ij}] = [A_{ij} + B_{ij}].$$

- Matrix addition follows the commutative and associative laws.

$$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}.$$
$$\left(\boldsymbol{A} + \boldsymbol{B}\right) + \boldsymbol{C} = \boldsymbol{A} + \left(\boldsymbol{B} + \boldsymbol{C}\right).$$

Inner or Dot Product

- If $\boldsymbol{a}$ and $\boldsymbol{b}$ are both $k \times 1$, then the sum of element-wise products is called the inner product or dot product.

$$\boldsymbol{a}^\top \boldsymbol{b} = \sum_{h=1}^{k} a_h b_h.$$

- The result of an inner product is a scalar.
- Dot product is commutative:

$$\boldsymbol{a}^\top \boldsymbol{b} = \boldsymbol{b}^\top \boldsymbol{a}.$$

- In Python, inner product is coded as `numpy.inner(a, b)`.

Linear Combination

- A vector can be written as a linear combination:

$$\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_k \end{bmatrix} = a_1 \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \vdots + a_k \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

- The vectors with 1 in the i-th row and the rest 0 are called the basis vectors, and denoted by $\boldsymbol{e}_i$.
- Compactly,
  - the column vector is expressed as $\boldsymbol{a} = \sum_{i=1}^{k} a_i \boldsymbol{e}_i$;
  - and the row vector $\boldsymbol{a}^\top$ as $\boldsymbol{a}^\top = \sum_{i=1}^{k} a_i \boldsymbol{e}_i^\top$.

Linear Dependence and Independence

1. A set of vectors $\boldsymbol{a}_i$, $i = 1, 2, \ldots, n$ are said to be linearly dependent if there exist scalars $c_1, c_2, \ldots, c_n$ such that

$$c_1 \boldsymbol{a}_1 + c_2 \boldsymbol{a}_2 + \cdots c_n \boldsymbol{a}_n = c^i \boldsymbol{a}_i = \boldsymbol{o}. \qquad (1)$$

2. A set of vectors $\boldsymbol{a}_i$, $i = 1, 2, \ldots, n$ are said to be linearly independent if they are not linearly dependent.

3. Hence, $\boldsymbol{a}_i$, $i = 1, 2, \ldots, n$ are linearly independent iff equation (1) has only the trivial solution:

$$c_i = 0, \qquad \text{for all } i = 1, 2, \ldots, n.$$

Einstein's Notation

- In Einstein's notation, the inner product is written as

$$\boldsymbol{a}^\top \boldsymbol{b} = \boldsymbol{e}^{i^\top} a_i b^j \boldsymbol{e}_j = a_i b^j \boldsymbol{e}^{i^\top} \boldsymbol{e}_j = a_i b^j \delta^i_j = a_i b^i,$$

where δ^i_j is called the Kronecker delta, which is defined as

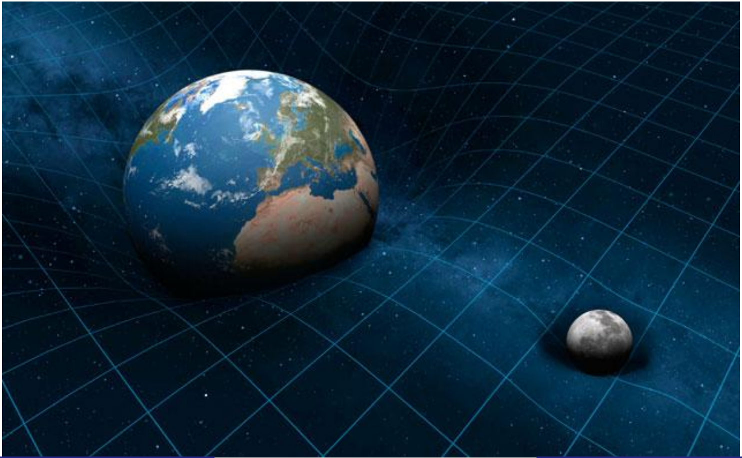
$$\boldsymbol{e}^{i^\top} \boldsymbol{e}_j = \delta^i_j := \begin{cases} 1 & \text{when } i = j; \\ 0 & \text{otherwise.} \end{cases}$$

- When $\boldsymbol{a}^\top \boldsymbol{b} = 0$, we say that $\boldsymbol{a}$ and $\boldsymbol{b}$ are orthogonal.

Inspirational Digression: General Relativity

- Space-Time Curvature = Energy-Momentum tensor

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}$$



Matrix Multiplication

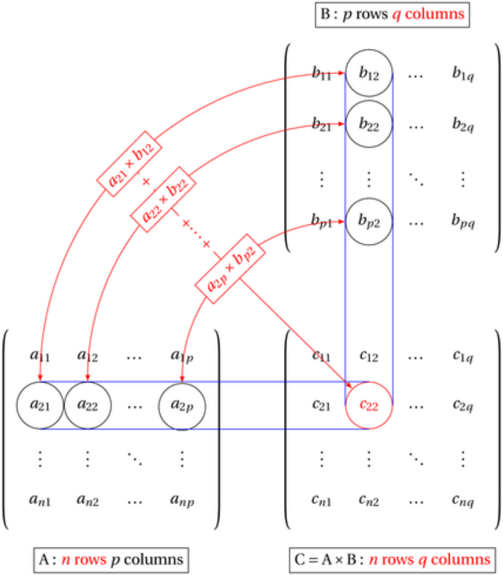
- Multiplication by a scalar: If c is a non-zero and finite scalar,

$$c\boldsymbol{A} = \boldsymbol{A}c = \left[A_{ij}c\right].$$

- If $\boldsymbol{A}$ is $k \times r$ and $\boldsymbol{B}$ is $r \times s$ so that the number of columns of $\boldsymbol{A}$ equals the number of rows of $\boldsymbol{B}$, we say that $\boldsymbol{A}$ and $\boldsymbol{B}$ are conformable, and the matrix product $\boldsymbol{AB}$ can be defined.

$$\boldsymbol{AB} = \begin{bmatrix} \boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \vdots \\ \boldsymbol{a}_k^\top \end{bmatrix} \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_s \end{bmatrix} = \begin{bmatrix} \boldsymbol{a}_1^\top \boldsymbol{b}_1 & \boldsymbol{a}_1^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1^\top \boldsymbol{b}_s \\ \boldsymbol{a}_2^\top \boldsymbol{b}_1 & \boldsymbol{a}_2^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2^\top \boldsymbol{b}_s \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{a}_k^\top \boldsymbol{b}_1 & \boldsymbol{a}_k^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_k^\top \boldsymbol{b}_s \end{bmatrix}$$

Matrix Multiplication



Square Matrix and Trace

- A matrix is said to be square when $r = k$.
- A square matrix is said to symmetric when $\boldsymbol{A}^\top = \boldsymbol{A}$.
- The trace of a $k \times k$ square matrix $\boldsymbol{A}$ is defined as

$$\operatorname{tr} \boldsymbol{A} := \sum_{i=1}^{k} A_{ii} = A_{ij} \delta^{ij},$$

i.e., the sum of its diagonal elements.

Properties of Trace

- For square matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and a real and finite constant c, we have

$$\begin{aligned} \operatorname{tr}\left(c\boldsymbol{A}\right) &= c\operatorname{tr}\boldsymbol{A}; \\ \operatorname{tr}\boldsymbol{A}^\top &= \operatorname{tr}\boldsymbol{A}; \\ \operatorname{tr}\left(\boldsymbol{A}+\boldsymbol{B}\right) &= \operatorname{tr}\boldsymbol{A}+\operatorname{tr}\boldsymbol{B}; \\ \operatorname{tr}\boldsymbol{I}_k &= k. \end{aligned}$$

- For example, the trace of the matrix

$$\operatorname{Tr}(\boldsymbol{A}) = \begin{bmatrix} 3 & 4 \\ 7 & 9 \end{bmatrix} = 3+9 = 12.$$

Commutativity of Trace

Theorem

For $k \times r$ $\boldsymbol{A}$ and $r \times k$ $\boldsymbol{B}$, we have $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$.

- Proof:

$$\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}\begin{bmatrix} \boldsymbol{a}_1^\top \boldsymbol{b}_1 & \boldsymbol{a}_1^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_1^\top \boldsymbol{b}_k \\ \boldsymbol{a}_2^\top \boldsymbol{b}_1 & \boldsymbol{a}_2^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_2^\top \boldsymbol{b}_k \\ \vdots & \vdots & \cdots & \vdots \\ \boldsymbol{a}_k^\top \boldsymbol{b}_1 & \boldsymbol{a}_k^\top \boldsymbol{b}_2 & \cdots & \boldsymbol{a}_k^\top \boldsymbol{b}_k \end{bmatrix}$$
$$= \boldsymbol{a}^{i^\top} \boldsymbol{b}_i = \boldsymbol{b}^{i^\top} \boldsymbol{a}_i$$
$$= \operatorname{tr}(\boldsymbol{BA}).$$

- We have applied the fact that the dot product is commutative.

Rank of a Matrix

1. The rank of a matrix is defined as the number of its linearly independent rows, which is equal to the number of its linearly independent columns, i.e.,

   row rank = column rank.

2. The rank of a matrix A is given by the maximum number of linearly independent rows (or columns). For example,

$$\text{rank}\begin{bmatrix} 3 & 4 \\ 7 & 9 \end{bmatrix} = 2,$$

$$\text{rank}\begin{bmatrix} 3 & 6 \\ 2 & 4 \end{bmatrix} = 1.$$

Properties of Rank

- A matrix with a rank equal to its dimension is a matrix of full rank.
- A matrix that is not full rank is known as a short rank matrix, and is singular (non-invertible).
- Four important properties:
  - $\text{rank}(\boldsymbol{A}) = \text{rank}\left(\boldsymbol{A}'\right)$.
  - $\text{rank}(c\boldsymbol{A}) = \text{rank}\left(\boldsymbol{A}\right)$, where c is a constant.
  - $\text{rank}(\boldsymbol{AB}) \leq \min\left(\text{rank}(\boldsymbol{A}), \text{rank}(\boldsymbol{B})\right)$
  - $\text{rank}\left(\boldsymbol{A}'\boldsymbol{A}\right) = \text{rank}\left(\boldsymbol{A}\boldsymbol{A}'\right) = \text{rank}\left(\boldsymbol{A}\right)$.

Rank and Inverse Matrix

- The rank of the $k \times r$ matrix ($r \leq k$)

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_r \end{bmatrix}$$

  is the number of linearly independent columns $\boldsymbol{a}_j$, and is written as $\operatorname{rank}\boldsymbol{A}$. We say that $\boldsymbol{A}$ has full rank if $\operatorname{rank}\boldsymbol{A} = r$.

- A square $k \times k$ matrix $\boldsymbol{A}$ is said to be non-singular if it is has full rank, e.g. $\operatorname{rank}\boldsymbol{A} = k$.

- If $\boldsymbol{A}$ is non-singular then there exists a unique $k \times k$ matrix $\boldsymbol{A}^{-1}$ called the inverse of $\boldsymbol{A}$ that satisfies

$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I}_k.$$

Properties of Matrix Inverse

- For non-singular $\boldsymbol{A}$ and $\boldsymbol{C}$,

$$\left(\boldsymbol{A}^{-1}\right)^{\top} = \left(\boldsymbol{A}^{\top}\right)^{-1};$$
$$\left(\boldsymbol{A}\boldsymbol{C}\right)^{-1} = \boldsymbol{C}^{-1}\boldsymbol{A}^{-1};$$
$$\left(\boldsymbol{A}+\boldsymbol{C}\right)^{-1} = \boldsymbol{A}^{-1}\left(\boldsymbol{A}^{-1}+\boldsymbol{C}^{-1}\right)^{-1}\boldsymbol{C}^{-1};$$
$$\boldsymbol{A}^{-1} - \left(\boldsymbol{A}+\boldsymbol{C}\right)^{-1} = \boldsymbol{A}^{-1}\left(\boldsymbol{A}^{-1}+\boldsymbol{C}^{-1}\right)^{-1}\boldsymbol{A}^{-1}.$$

- If $\boldsymbol{A}$ is an orthogonal matrix, then $\boldsymbol{A}^{-1} = \boldsymbol{A}^{\top}$.

Determinant of Square Matrix

- Let $\boldsymbol{A}$ be a general $k \times k$ matrix. Let $(j_1, j_2, \ldots, j_k)$ denote a permutation of $(1, 2, \ldots, k)$. There are $k!$ permutations.
- There is a unique count of the number of *inversions* of the indices of such permutations relative to the natural order $(1, 2, \ldots, k)$, and let $\epsilon_{(j_1, j_2, \ldots, j_k)} = +1$ if this count is even and $\epsilon_{(j_1, j_2, \ldots, j_k)} = -1$ if the count is odd. Then the determinant of $\boldsymbol{A}$ is defined as

$$\begin{aligned} \det \boldsymbol{A} &:= \sum_{\pi} \epsilon_{\pi} A_{1j_1} A_{2j_2} \cdots A_{kj_k} \\ &= \xi^{j_1 j_2 \cdots j_k} A_{1j_1} A_{2j_2} \cdots A_{kj_k}, \end{aligned}$$

where

$$\xi^{j_1 j_2 \cdots j_k} := \begin{cases} 1 & \text{if } (j_1, j_2, \ldots, j_k) \text{ is even;} \\ -1 & \text{if } (j_1, j_2, \ldots, j_k) \text{ is odd.} \end{cases}$$

Properties of Determinant

- For example, if $k=2$, then the two permutations of $(1,2)$ are $(1,2)$ and $(2,1)$, for which $\epsilon_{(1,2)}=1$ and $\epsilon_{(2,1)}=-1$. Thus,

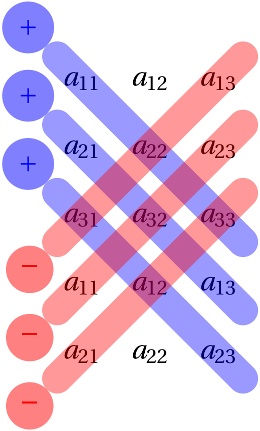
$$\det \boldsymbol{A} = \epsilon_{(1,2)} A_{11} A_{22} + \epsilon_{(2,1)} A_{21} A_{12} = A_{11} A_{22} - A_{21} A_{12}.$$

- If $\boldsymbol{A}$ is non-singular, $\det \boldsymbol{A} \neq 0$.
- If $\boldsymbol{A}$ is orthogonal, then $\det \boldsymbol{A} = \pm 1$.
- If $\boldsymbol{A}$ is triangular (upper or lower), then $\det \boldsymbol{A} = \prod_{i=1}^{k} A_{ii}$.
- Some other properties of a $k \times k$ square matrix $\boldsymbol{A}$ include

$$\begin{aligned}
\det \boldsymbol{A} &= \det \boldsymbol{A}^{\top}; \\
\det\left(c\boldsymbol{A}\right) &= c^{k} \det \boldsymbol{A}; \\
\det\left(\boldsymbol{A}\boldsymbol{B}\right) &= \det \boldsymbol{A} \det \boldsymbol{B}; \\
\det\left(\boldsymbol{A}^{-1}\right) &= \left(\det \boldsymbol{A}\right)^{-1}; \\
\det \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{bmatrix} &= \det \boldsymbol{D} \det\left(\boldsymbol{A} - \boldsymbol{B}\boldsymbol{D}^{-1}\boldsymbol{C}\right), \quad \text{if } \det \boldsymbol{D} \neq 0.
\end{aligned}$$

Determinant of a 3 × 3 Matrix

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \left(a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}\right) - \left(a_{13}a_{22}a_{31} + a_{23}a_{32}a_{11} + a_{33}a_{12}a_{21}\right)$$



Calculating the Inverse of a 2×2 Matrix

- The inverse of a 2×2 non-singular matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{is} \quad \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

- The expression in the denominator, $ad-bc$, is the determinant of the matrix.
- If the matrix is

$$\begin{bmatrix} 2 & 1 \\ 4 & 6 \end{bmatrix},$$

then the inverse will be

$$\frac{1}{8}\begin{bmatrix} 6 & -1 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{8} \\ -\frac{1}{2} & \frac{1}{4} \end{bmatrix}.$$

- As a check, multiply the two matrices together and it should give the identity matrix $\boldsymbol{I}$.

Eigenvalues

- The characteristic equation of a $k \times k$ square matrix $\boldsymbol{A}$ is

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}_k) = 0.$$

- It is a polynomial of degree k in λ, so it has exactly k roots, which are not necessarily distinct and may be real or complex.
- They are called the latent roots or characteristic roots or eigenvalues of $\boldsymbol{A}$.
- When the eigenvalues are real, they are written in descending order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$. We also write $\lambda_{\min}(\boldsymbol{A}) = \lambda_k$ and $\lambda_{\max}(\boldsymbol{A}) = \lambda_1$.

Eigenvectors

- If λ_i is an eigenvalue of $\boldsymbol{A}$, then $\boldsymbol{A} - \lambda_i \boldsymbol{I}_k$ is singular, i.e., there exists a non-zero vector $\boldsymbol{h}_i$ such that

$$\left(\boldsymbol{A} - \lambda_i \boldsymbol{I}_k\right)\boldsymbol{h}_i = \boldsymbol{o}.$$

- The vector $\boldsymbol{h}_i$ is called a latent vector or characteristic vector or eigenvector of $\boldsymbol{A}$ corresponding to λ_i.
- It is a fundamental result of linear algebra that an equation $\boldsymbol{M}\boldsymbol{v} = \boldsymbol{o}$ has a non-zero solution $\boldsymbol{v}$ if and only if the determinant $\det(\boldsymbol{M})$is zero. Hence,

$$f(\lambda) := \det\left(\boldsymbol{A} - \lambda \boldsymbol{I}\right) = 0.$$

Calculating Eigenvalues: An Example

- Let A be the 2 × 2 matrix $A = \begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix}$
- Then the characteristic equation is $|A - \lambda I_2| = 0$. That is

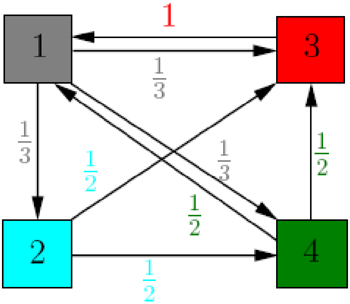
$$\begin{bmatrix} 5 & 1 \\ 2 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$$

$$\Longrightarrow \begin{bmatrix} 5-\lambda & 1 \\ 2 & 4-\lambda \end{bmatrix} = (5-\lambda)(4-\lambda) - 2 = \lambda^2 - 9\lambda + 18$$

- The solutions are $\lambda = 6$ and $\lambda = 3$.
- The characteristic roots are also known as ___________.

A Simple PageRank Algorithm

- Transition matrix of the directed graph of 4 web sites:



$$A = \begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix}$$

A Simple PageRank Algorithm (Cont'd)

- Denote by x_1, x_2, x_3, and x_4 the importance of the four sites. Analyzing the situation at each node, we get the system:

$$
\begin{aligned}
x_1 &= 1 \cdot x_3 + \frac{1}{2} \cdot x_4 \\
x_2 &= \frac{1}{3} \cdot x_1 \\
x_3 &= \frac{1}{3} \cdot x_1 + \frac{1}{2} \cdot x_2 + \frac{1}{2} \cdot x_4 \\
x_4 &= \frac{1}{3} \cdot x_1 + \frac{1}{2} \cdot x_2
\end{aligned}
$$

- It is equivalent to $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{x}$, where $\boldsymbol{x}^\top := [x_1 \quad x_2 \quad x_3 \quad x_4]$.

Finding the Eigenvector

- The PageRank algorithm involves finding the eigenvector corresponding to the eigenvalue of 1!

$$\begin{bmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = 1 \times \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

- Since $x_2 = x_1/3$, we substitute it in x_4 to find that $x_4 = x_1/2$. In turn, we find that $x_3 = \frac{x_1}{3} + \frac{1}{2} \times \frac{x_1}{3} + \frac{1}{2} \times \frac{x_1}{2} = \frac{3x_1}{4}$.

Finding the Eigenvector (Cont'd)

- So the eigenvector is

$$\boldsymbol{v} := \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 1/3 \\ 3/4 \\ 1/2 \end{bmatrix}.$$

- To find x_1, we impose the condition that the sum of eigenvector's entries is equal to 1. Therefore

$$x_1 \left(1 + \frac{1}{3} + \frac{3}{4} + \frac{1}{2} \right) = 1.$$

Finding the Eigenvector (Cont'd)

- So $x_1 = 0.3871$. Finally, the unique eigenvector is

$$\begin{bmatrix} 0.3871 \\ 0.1290 \\ 0.2903 \\ 0.1935 \end{bmatrix}.$$

- Therefore, the PageRank is, in declining order of importance, Site 1, Site 3, Site 4, and Site 2.

Final Result

- So $x_1 = 0.3871$. Finally, the unique eigenvector is

$$\begin{bmatrix} 0.3871 \\ 0.1290 \\ 0.2903 \\ 0.1935 \end{bmatrix}.$$

- Therefore, the PageRank is, in declining order of importance, Site 1, Site 3, Site 4, and Site 2.
- Reference: PageRank Algorithm - The Mathematics of Google Search
- "...some of you may have heard of quants but at Google, they're just called employees, because they're all quants..."

James Simon

Definition of Vector Differentiation

- Let $\boldsymbol{x}$ be a column k-vector. Consider the function

$$g(\boldsymbol{x}) = g(x_1, x_2, \ldots, x_k) : \Re^k \longrightarrow \Re.$$

- The vector derivative is

$$\frac{\partial}{\partial \boldsymbol{x}} g(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} g(\boldsymbol{x}) \\ \frac{\partial}{\partial x_2} g(\boldsymbol{x}) \\ \vdots \\ \frac{\partial}{\partial x_k} g(\boldsymbol{x}) \end{bmatrix}.$$

and

$$\frac{\partial}{\partial \boldsymbol{x}^\top} g(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial}{\partial x_1} g(\boldsymbol{x}) & \frac{\partial}{\partial x_2} g(\boldsymbol{x}) & \cdots & \frac{\partial}{\partial x_k} g(\boldsymbol{x}) \end{bmatrix}$$

Basic Properties

- For constant vector $\boldsymbol{a}$ and matrix $\boldsymbol{A}$,

$$\frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{a}^\top \boldsymbol{x}\right) = \frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^\top \boldsymbol{a}\right) = \boldsymbol{a}, \qquad \frac{\partial}{\partial \boldsymbol{x}^\top}\left(\boldsymbol{a}^\top \boldsymbol{x}\right) = \boldsymbol{a}^\top$$

$$\frac{\partial}{\partial \boldsymbol{x}^\top}\left(\boldsymbol{A}\boldsymbol{x}\right) = \boldsymbol{A}$$

$$\frac{\partial}{\partial \boldsymbol{x}}\left(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}\right) = \left(\boldsymbol{A} + \boldsymbol{A}^\top\right)\boldsymbol{x}$$

$$\frac{\partial^2}{\partial \boldsymbol{x}\partial \boldsymbol{x}^\top}\left(\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}\right) = \boldsymbol{A} + \boldsymbol{A}^\top$$

Assets

- There are n assets whose expected returns are denoted by

$$\mu_i := \mathbb{E}(r_i), \quad i = 1, 2, \ldots, n.$$

- The covariance between asset i and asset j is denoted as σ_{ij}. Arrange the covariances into a $n \times n$ matrix:

$$\boldsymbol{\Sigma} := [\sigma_{ij}] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

  - The diagonal element σ_{ii} is the variance of asset i.
  - Note that the covariance matrix $\boldsymbol{\Sigma}$ is symmetric.

Investment

- For each dollar invested, a fraction w_i is invested in asset i. It must be that
$$\sum_{i=1}^{n} w_i = 1.$$
- The weights are arranged as a n-vector $\boldsymbol{w}$.
- The portfolio's expected return and variance are, respectively,

$$\mu_p := \mathbb{E}(r_p) = \sum_{i=1}^{n} w_i \mathbb{E}(r_i) = w_i \mu^i = \boldsymbol{w}^\top \boldsymbol{\mu};$$
$$\sigma_p^2 := \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij} = w_i \Sigma^i_j w_j = \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}.$$

Numerical Illustration

- Suppose there are two assets. $\mu_1 = 5\%$ and $\mu_2 = 8\%$ per annum.
- The covariance is a 2 by 2 matrix.

$$\mathbf{\Sigma} := \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} 0.0625 & -0.01 \\ -0.01 & 0.16 \end{bmatrix}$$

- Given that the variance of asset 1 is 0.0625, its volatility is $\sqrt{0.0625} = 25\%$ per annum.
  The volatility of asset 2 is ____________________.
- The portfolio's expected return and variance are, respectively,

$$\mu_p = 0.05 w_1 + 0.08 w_2$$
$$\sigma_p^2 = 0.0625 w_1^2 - 0.01 w_1 w_2 - 0.01 w_2 w_1 + 0.16 w_2^2$$

Optimization

- Minimize half the portfolio variance under two constrains:

$$\begin{aligned}\sum_{i=1}^{n} w_i \mathbb{E}(r_i) &= \mathbb{E}(r_p). \\ \sum_{i=1}^{n} w_i &= 1.\end{aligned}$$

- Constrained optimization with Lagrange multipliers λ and ψ:

$$\min_{w_1, w_2, \ldots, w_n} L = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} - \lambda \left(\sum_{i=1}^{n} w_i \mu_i - \mu_p \right) - \psi \left(\sum_{i=1}^{n} w_i - 1 \right)$$

In Matrix Form

- The Lagrangian L is

$$L = \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} - \lambda\left(\boldsymbol{w}^\top \boldsymbol{\mu} - \mu_p\right) - \psi(\boldsymbol{w}^\top \mathbf{1} - 1).$$

- The first-order conditions with respect to $\boldsymbol{w}$ are

$$\begin{aligned} \boldsymbol{\Sigma} \boldsymbol{w} - \lambda \boldsymbol{\mu} - \psi \mathbf{1} &= \mathbf{0} \\ \boldsymbol{w}^\top \boldsymbol{\mu} &= \mu_p \\ \boldsymbol{w}^\top \mathbf{1} &= 1 \end{aligned}$$

Solution of First FOC

- The first FOC gives the solution for the weight vector

$$\boldsymbol{w}^* = \boldsymbol{\Sigma}^{-1}\left(\lambda\boldsymbol{\mu} + \psi\mathbf{1}\right).$$

- But what are the values of the Lagrange multipliers λ and ψ?
- To solve for λ and ψ, substitute the optimal weight vector above into the last two FOC's,

$$\begin{aligned} \boldsymbol{\mu}^\top \boldsymbol{w}^* &= \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\left(\lambda\boldsymbol{\mu} + \psi\mathbf{1}\right) = \mu_p \\ \mathbf{1}^\top \boldsymbol{w}^* &= \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\left(\lambda\boldsymbol{\mu} + \psi\mathbf{1}\right) = 1 \end{aligned}$$

Solution of Second and Third FOCs

- Let

$$A := \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

$$B := \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}$$

$$C := \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}$$

  The last two FOCs can be written as

$$\begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \lambda \\ \psi \end{bmatrix} = \begin{bmatrix} \mu_p \\ 1 \end{bmatrix}$$

- Solving these two linear equations, we obtain

$$\lambda = \frac{C\mu_p - B}{AC - B^2}, \qquad \psi = \frac{A - B\mu_p}{AC - B^2}$$

Optimal Weight Vector and Portfolio Variance

- The optimal weight is then solved as

$$\boldsymbol{w}^* = \frac{\boldsymbol{\Sigma}^{-1}\Big(\big(C\mu_p - B\big)\boldsymbol{\mu} + \big(A - B\mu_p\big)\mathbf{1}\Big)}{AC - B^2}.$$

- The portfolio variance is a quadratic function of the mean portfolio return μ_p:

$$\begin{aligned}\mathbb{V}\big(r_p\big) = \boldsymbol{w}^{*\top}\boldsymbol{\Sigma}\boldsymbol{w}^* &= \frac{C\mu_p^2 - 2B\mu_p + A}{AC - B^2} \\ &= \frac{C}{AC - B^2}\mu_p^2 - \frac{2B}{AC - B^2}\mu_p + \frac{A}{AC - B^2}.\end{aligned}$$

Global Minimum Variance Portfolio

- The global minimum variance portfolio is obtained by minimizing $\mathbb{V}(r_p)$ with respect to μ_p.

$$\frac{d\mathbb{V}(r_p)}{d\mu_p} = \frac{2C}{AC - B^2}\mu_p - \frac{2B}{AC - B^2}.$$

- The results of the first-order condition are

$$\mu_\star = \frac{B}{C}, \qquad \sigma^2_\star = \frac{1}{C}, \qquad \boldsymbol{w}_\star = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{C}.$$

Numerical Example

- For the two assets, compute the inverse matrix

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} 0.0625 & -0.01 \\ -0.01 & 0.16 \end{bmatrix}^{-1}$$

$$= \frac{1}{0.0625 \times 0.16 - (-0.01) \times (-0.01)} \begin{bmatrix} 0.16 & 0.01 \\ 0.01 & 0.0625 \end{bmatrix}$$

$$= \begin{bmatrix} 16.16 & 1.01 \\ 1.01 & 6.31 \end{bmatrix}$$

Values of A, B, and C

- So the three scalars A, B, and C are

$$
\begin{aligned}
A &= \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \\
&= \begin{bmatrix} 0.05 & 0.08 \end{bmatrix} \begin{bmatrix} 16.16 & 1.01 \\ 1.01 & 6.31 \end{bmatrix} \begin{bmatrix} 0.05 \\ 0.08 \end{bmatrix} \\
&= 0.088864; \\
B &= \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \\
&= \begin{bmatrix} 0.05 & 0.08 \end{bmatrix} \begin{bmatrix} 16.16 & 1.01 \\ 1.01 & 6.31 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= 1.4441 \\
C &= \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \\
&= 24.49.
\end{aligned}
$$

Global Minimum Variance Portfolio

- Let's consider the global minimum variance portfolio.

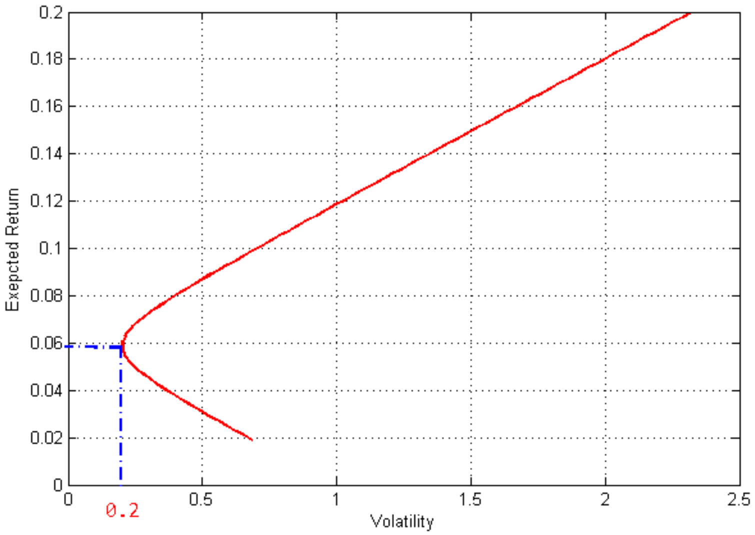
$$\mu_\star = \frac{B}{C} = 0.0590, \quad \sigma^2_\star = \frac{1}{C} = 0.0408$$

So the minimum volatility is $\sigma_\star = 20.21\%$.

- The weight vector for $w_\star$ for the global minimum variance portfolio is

$$\boldsymbol{w}_\star = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{C} = \frac{1}{24.49}\begin{bmatrix} 16.16 & 1.01 \\ 1.01 & 6.31 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 70.11\% \\ 29.89\% \end{bmatrix}$$

Efficient Frontier



Takeaways

- Quantitative (mathematics and programming) skills provide a competitive advantage.
- Ideas of eigenvalue and eigenvector appear in Google search engine!
- Investment optimization: Obtain highest possible return with minimal risk.
- Quants' way of deriving the efficient frontier

Week 12 Assignment

1. What are the eigenvalues and eigenvectors of a diagonal matrix?

2. Let $A = \begin{bmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{bmatrix}$.

   (A) Show that the characteristic equation is

   $$(\lambda - 1)(\lambda + 3)^2 = 0.$$

   (B) Find the eigenvector corresponding to $\lambda_1 = 1$.

   (C) Find the eigenvector corresponding to $\lambda = -3$.

Week 12 Additional Exercise

1. Consider the transformation of Cartesian coordinates (x, y) to the polar coordinates (r, φ):

$$x = r\cos\varphi, \qquad y = r\sin\varphi$$

   (A) Show that

$$\boldsymbol{J}(r,\varphi) = \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{bmatrix} = \begin{bmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{bmatrix}$$

   (B) Show that $\det \boldsymbol{J} = r$.

2. Apply the Jacobian $\det \boldsymbol{J}$ in Problem 1 to show that

$$I := \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{2\pi}.$$