6: Inverse matrices

- left and right inverse
- nonsingular matrices
- orthogonal matrices

Left and right inverse

a matrix inverse has come up several times so far – we now make its meaning precise

- in general, matrices do not commute: $AX \neq XA$
- in contrast, for scalars, the multiplicative inverse is straight forward to define

$$a \cdot b = 1 \quad \iff \quad b = \frac{1}{a} \quad \text{and} \quad a = \frac{1}{b}$$

- for matrices, two types of inverse are possible:

$$\underbrace{XA = I}_{X \text{ is a left inverse}} \qquad \text{and} \qquad \underbrace{AX = I}_{X \text{ is a right inverse}}$$

- A may have neither, one, or both types of inverse
- if $A \in \mathbb{F}^{m \times n}$, then a left or right inverse must have dimension $n \times m$

Linear independence

recall (p.3-18), the vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ are **linearly independent** iff

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_n = 0 \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$$

- in matrix notation, let $A = [v_1 \ v_2 \ \ldots \ v_n]$, if $\{v_i\}_{i=1}^n$ are linearly independent

$$Ax = 0 \quad \Longrightarrow \quad x = 0$$

Linear independence

recall (p.3-18), the vectors $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ are **linearly independent** iff

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_k v_n = 0 \quad \Longrightarrow \quad \lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$$

- in matrix notation, let $A = [v_1 \ v_2 \ \ldots \ v_n]$, if $\{v_i\}_{i=1}^n$ are linearly independent

$$Ax = 0 \quad \Longrightarrow \quad x = 0$$

Independence-dimension inequality (see p.3-20)

If the set of vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ are linearly independent, then $n \leq m$.

- an $m \times n$ matrix with **linearly independent columns** then $m \geq n$ (must be tall)
- an $m \times n$ matrix with **linearly independent rows** then $m \leq n$ (must be wide)

$CA = I$

Left invertibility

given an $m \times n$ matrix A, suppose $CA = I$, then A must be tall (or square), and have independent columns

$CA = I$

Left invertibility

given an $m \times n$ matrix A, suppose $CA = I$, then A must be tall (or square), and have independent columns

to see this:

1. to have independent columns, A must be tall or square (ind-dim inequality)
2. suppose $Ax = 0$, left multiply by C gives

$$0 = C(Ax) = (CA)x = Ix = x$$

   thus $x_i = 0$ are the only coefficients of the columns that gives 0_m

the converse is also true: a left inverse exists, iff the columns are linearly independent

$AB = I$

Right invertibility

given an $m \times n$ matrix A, if $AB = I$, then A must be wide (or square), and have have independent rows

$AB = I$

Right invertibility

given an $m \times n$ matrix A, if $AB = I$, then A must be wide (or square), and have have independent rows

to see this we do not need to develop a new set of results, it follows from

- A must be square or wide to have a right inverse
- if $AB = I$ then $B^T A^T = I$ (because $B^T A^T = (AB)^T = I$)

Inverse

if A has a left and a right inverse, i.e., $XA = I$ and $AY = I$, then $X = Y$:

$$XA = I, \quad AY = I \quad \Longrightarrow \quad X = X(AY) = (XA)Y = Y$$

- the matrix $X = Y$ is called the **inverse** of A, denoted A^{-1}
- the inverse is **unique**
- A is said to be **invertible** or **non-singular** if A^{-1} exists
- $(A^{-1})^{-1} = A$
- invertible matrices **must be square** (tall: no right inverse, wide: no left inverse)
- it is always true that $A^{-1}A = AA^{-1} = I$

Examples

- **diagonal matrix**

$$\begin{bmatrix} 2 & & \\ & -4 & \\ & & \frac{1}{3} \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & & \\ & -\frac{1}{4} & \\ & & 3 \end{bmatrix}$$

- **unstructured matrix**

$$\begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 2 \\ -3 & -4 & -4 \end{bmatrix}^{-1} = \frac{1}{30} \begin{bmatrix} 0 & -20 & -10 \\ -6 & 5 & -2 \\ 6 & 10 & 2 \end{bmatrix}$$

- **permutation matrix**

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

- **left inverse**

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}, \quad B = \frac{1}{9}\begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \quad C = \frac{1}{2}\begin{bmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{bmatrix}$$

$BA = CA = I$, both B and C are left inverses of A

- **inverse of a transpose**

$$(A^T)^{-1} = (A^{-1}) \quad \text{often written as} \quad A^{-T}$$

- **inverse of a product**

$$(AB)^{-1} = B^{-1}A^{-1}$$

2×2 matrix inverse

for this special case, there is a manageable closed-form solution:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

- recall that $ad - bc$ is the **determinant**
- clearly if $ad - bc = 0$ then the matrix is **singular**

Systems of linear equations

linear independence has clear implications for linear systems

- a set of m linear equations with n unknowns

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \end{array}$$

- in matrix form: $Ax = b$ with $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n$, and $b \in \mathbb{R}^m$

- can replace $\mathbb{R}$ with $\mathbb{C}$ and nothing changes

- from p.6-3, if A is wide, the columns are linearly dependent, and

$$Ax = 0$$

  has non-trivial solutions

Portfolio allocation

an investor must choose how to allocate money among several assets

each asset has returns that vary under different market conditions.

pick weights x (how much to invest in each asset) to meet a desired portfolio return b

consider n assets and m possible market scenarios

Scenario	**Stock**	**Bond**	**Real Estate**
1 (bull)	0.10	0.02	0.04
2 (bear)	-0.05	0.03	0.06
3 (stable)	0.08	0.01	0.03

model as a linear system $Ax = b$:

$$A = \begin{bmatrix} 0.10 & 0.02 & 0.04 \\ -0.05 & 0.03 & 0.06 \\ 0.08 & 0.01 & 0.03 \end{bmatrix}, \quad x = \begin{bmatrix} x_{\text{stock}} \\ x_{\text{bond}} \\ x_{\text{real estate}} \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Interpretation

- A_{ij}: return of asset j in scenario i (as a $\%$)
- x_j: fraction of portfolio invested in asset j
- b_i: portfolio return in scenario i

Interpretation

- A_{ij}: return of asset j in scenario i (as a $\%$)
- x_j: fraction of portfolio invested in asset j
- b_i: portfolio return in scenario i

Matrix-vector multiplication

given data A and portfolio weight x, $b = Ax$ produces returns given each market condition

Interpretation

- A_{ij}: return of asset j in scenario i (as a %)
- x_j: fraction of portfolio invested in asset j
- b_i: portfolio return in scenario i

Matrix-vector multiplication

given data A and portfolio weight x, $b = Ax$ produces returns given each market condition

Solving $Ax = b$ for x

determine portfolio weight to generate returns b under m market scenarios

$$\begin{bmatrix} 0.10 & 0.02 & 0.04 \\ -0.05 & 0.03 & 0.06 \\ 0.08 & 0.01 & 0.03 \end{bmatrix} \begin{bmatrix} x_{\text{stock}} \\ x_{\text{bond}} \\ x_{\text{real estate}} \end{bmatrix} = \begin{bmatrix} 0.06 \\ 0.02 \\ 0.04 \end{bmatrix}$$

solution given by $x = [0.4 \;\; 0.3 \;\; 0.3]^T$

Solution via a right inverse

consider the set of linear equations $Ax = b$, with $AB = I$, then

- A must be square or wide

$$\begin{bmatrix} & & A & & \end{bmatrix} \begin{bmatrix} \\ x \\ \\ \end{bmatrix} = \begin{bmatrix} b \end{bmatrix}$$

- for any n-vector b, $x = Bb$ satisfies $Ax = b$, to see this:

$$Ax = A(Bb) = (AB)b = Ib = b$$

Conclusion

if A is wide and has linearly independent rows, then $Ax = b$ has a solution for any b

Solution via a left inverse

consider the set of linear equations $Ax = b$, with $CA = I$, then

- A must be square or tall

$$\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

- multiplying $Ax = b$ on the left by C gives

$$Cb \ = \ C(Ax) \ = \ (CA)x \ = \ Ix \ = \ x$$

so $x = Cb$ is a solution to $Ax = b$

Conclusion

when A is tall, $x = Cb$ is the unique solution to $Ax = b$, if $A(Cb) \neq b$, then there is no solution

Solutions to square systems

Theorem:

Consider a system with n equations and n variables, $Ax = b$, if A^{-1} exists, then

$$x = A^{-1}b,$$

and the solution is unique.

Proof

1. the solution exists because A^{-1} is a right inverse of A (p.6-13)
2. the solution is unique because A^{-1} is a left inverse of A (p.6-14)

this leads to the obvious question: *when is a square matrix invertible?*

Nonsingular matrices

Theorem

Let A be a square matrix. The following properties are equivalent:

1. *A is left invertible*
2. *A has linearly independent columns*
3. *A is right invertible*
4. *A has linearly independent rows*

a square matrix with these properties is **invertible** or **nonsingular**

Proof

Proof strategy

$$(1) \quad \Longrightarrow \quad (2) \quad \Longrightarrow \quad (3) \quad \Longrightarrow \quad (4) \quad \Longrightarrow \quad (1)$$

(1) $\Longrightarrow$ (2)

- already shown this this for all tall and square matrices on p. 6-4

Proof

Proof strategy

$$(1) \implies (2) \implies (3) \implies (4) \implies (1)$$

(1) $\implies$ (2)

- already shown this this for all tall and square matrices on p. 6-4

(2) $\implies$ (3)

1. let $A = [a_1 \; a_2 \; \ldots \; a_n]$, by assumption columns are linearly independent
2. thus, $a_1, \ldots, a_n$ from a basis for $\mathbb{R}^n$ so all vectors in $\mathbb{R}^n$ can be expressed as a linear combination of the columns of A
3. specifically, $e_i = Ab_i$ for some vector b_i, define
$$B = [b_1 \; b_2 \; \ldots \; b_n]$$
4. B satisfies
$$AB = [Ab_1 \; Ab_2 \; \ldots \; Ab_n] = [e_1 \; e_2 \; \ldots \; e_n] = I$$

(3) $\Longrightarrow$ (4)

- apply transpose to arguments to $(1) \implies (2)$

(4) $\Longrightarrow$ (1)

- apply transpose to arguments to $(2) \implies (3)$

Example

Gram matrix

for an $m \times n$ matrix A with columns $a_1, \ldots, a_n$, the associated **Gram matrix** is

$$A^*A = \begin{bmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle & \ldots & \langle a_1, a_n \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle & \ldots & \langle a_2, a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_n, a_1 \rangle & \langle a_n, a_2 \rangle & \ldots & \langle a_n, a_n \rangle \end{bmatrix}$$

- the Gram matrix is nonsingular iff the columns of A are linearly independent

Invertibility and eigenvalues

Corollary

Let A be an $n \times n$ matrix with eigenvalues and vectors $Av_i = \lambda_i v_i$, for $i = 1, \ldots, n$. If $\lambda_i = 0$ for any i, then A is singular.

Invertibility and eigenvalues

Corollary

Let A be an $n \times n$ matrix with eigenvalues and vectors $Av_i = \lambda_i v_i$, for $i = 1, \ldots, n$. If $\lambda_i = 0$ for any i, then A is singular.

Proof.

1. from the theorem on p.6-16, A must have linearly independent columns
2. from p.6-3 if A does not have linearly independent columns then

$$Ax = 0 \quad \iff \quad x = 0 \tag{1}$$

3. assume $\lambda_j = 0$, then (1) holds with $x = v_j$.

□

Invertibility and eigenvalues

Corollary

Let A be an $n \times n$ matrix with eigenvalues and vectors $Av_i = \lambda_i v_i$, for $i = 1, \ldots, n$. If $\lambda_i = 0$ for any i, then A is singular.

Proof.

1. from the theorem on p.6-16, A must have linearly independent columns
2. from p.6-3 if A does not have linearly independent columns then

$$Ax = 0 \iff x = 0 \tag{1}$$

3. assume $\lambda_j = 0$, then (1) holds with $x = v_j$.

□

Alternative proof

Proof.

1. if λ is an eigenvalue then $\det(A - \lambda_i) = 0$
2. for $\lambda = 0$ this gives $\det(A) = 0$
3. if $\det(A) = 0$ then A is singular

□

Matrices with orthonormal columns

recall, a set of n-vectors $v_1, \ldots, v_k$ is **orthonormal** if

- $\|v_i\|_2 = 1$ for $i = 1, \ldots, k$
- all the vectors are mutually orthogonal

$$v_i^T v_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Example

the set of vectors

$$v_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad v_2 = \frac{1}{2} \begin{bmatrix} 1 \\ \sqrt{2} \\ 1 \end{bmatrix}, \quad v_3 = \frac{1}{2} \begin{bmatrix} 1 \\ -\sqrt{2} \\ 1 \end{bmatrix}$$

are orthonormal

sets of orthonormal vectors are **linearly independent**

- consider the orthonormal set $a_1, \ldots, a_k$ and assume

$$\sum_{i=1}^{k} a_i \beta_i = 0$$

- take the inner product of both sides of the inequality with some a_i

$$\begin{aligned} 0 &= a_i^T(\beta_1 a_1 + \ldots + \beta_k a_k) \\ &= \beta_1(a_1^T a_1) + \ldots \beta_k(a_k^T a_k) \\ &= \beta_i \end{aligned}$$

- i.e., the only linear combination of the elements $\{a_i\}$ that gives zero is $\beta_i = 0$

Matrices with orthonormal columns

strangely, there is no name for matrices with orthonormal columns

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}, \quad a_i\text{'s mutually orthonormal}$$

- if $A \in \mathbb{R}^{m \times n}$ then $m \geq n$ (tall)
- the Gram matrix $A^T A = I$, this has several implications:

length preservation

$$\|Ax\|_2 \;=\; \sqrt{(Ax)^T(Ax)} \;=\; \sqrt{x^T(A^TA)x} \;=\; \|x\|_2$$

inner product preservation

$$(Ax)^T(Ay) = x^T(A^TA)y = x^Ty$$

Orthonormal summary

if A has the form

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}, \quad a_i\text{'s mutually orthonormal}$$

then

- A is tall or square
- A has linearly independent columns
- A has a left inverse, i.e., $BA = I$, moreover

$$A^T A = I$$

Orthogonal matrices

a real square matrix, A, is called **orthogonal** if it has orthonormal columns

- orthogonal matrices are nonsingular (follows from p. 6-16)
- from the Gram matrix property, it has a left inverse given by its transpose

$$A \in \mathbb{R}^{n\times n} \quad \text{and} \quad \{a_i\} \ \text{orthogonal} \quad \Longrightarrow \quad A^T A = I \quad \Longleftrightarrow \quad AA^T = I$$

- it is thus clear that $A^T = A^{-1}$

$$\begin{bmatrix} \text{---} & a_1^T & \text{---} \\ & \vdots & \\ \text{---} & a_n^T & \text{---} \end{bmatrix} \left[\begin{array}{c|c|c} a_1 & \cdots & a_n \end{array}\right] = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}$$

Examples

Permutation matrices

- map the vector $x = (6, 3, 2, 0)$ to $y = (3, 0, 2, 6)$

Examples

Permutation matrices

- map the vector $x = (6, 3, 2, 0)$ to $y = (3, 0, 2, 6)$
- using the notation from p.4-21: $\pi = (2, 4, 3, 1)$, then $x \mapsto Px = y$, with

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Examples

Permutation matrices

- map the vector $x = (6, 3, 2, 0)$ to $y = (3, 0, 2, 6)$
- using the notation from p.4-21: $\pi = (2, 4, 3, 1)$, then $x \mapsto Px = y$, with

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

- clearly $\|p_i\|_2 = 1$ for all i, $p_i \perp p_j$ for $i \neq j$, and $p_i^T p_i = 1$

Examples

Permutation matrices

- map the vector $x = (6, 3, 2, 0)$ to $y = (3, 0, 2, 6)$
- using the notation from p.4-21: $\pi = (2, 4, 3, 1)$, then $x \mapsto Px = y$, with

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

- clearly $\|p_i\|_2 = 1$ for all i, $p_i \perp p_j$ for $i \neq j$, and $p_i^T p_i = 1$
- to recover $x = (6, 3, 2, 0)$ from $y = (3, 0, 2, 6)$

$$x \;=\; P^{-1}y \;=\; P^T y \quad \text{with} \quad P^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = P^T$$

Examples

rotation matrices

consider the matrix R that rotates a vector counterclockwise in the $\mathbb{R}^2$ plane

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- $R^T R = I$ follows from $\cos^2\theta + \sin^2\theta = 1$

Examples

rotation matrices

consider the matrix R that rotates a vector counterclockwise in the $\mathbb{R}^2$ plane

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

- $R^T R = I$ follows from $\cos^2\theta + \sin^2\theta = 1$

Products of orthogonal matrices

let $A_1, \ldots, A_k$ all be orthogonal matrices with the same dimensions, then the product

$$A = A_1 A_2 \ldots A_k$$

is also orthogonal

$$\begin{aligned} A^T A &= (A_k^T \ldots A_2^T \underbrace{A_1^T)(A_1}_{I} A_2 \ldots A_k) \\ &= (A_k^T \ldots A_2^T)(A_2 \ldots A_k) \\ &\vdots \\ &= I \end{aligned}$$

Solving an orthogonal system of equations

if A is an $n \times n$ matrix with orthogonal columns, then the solution of

$$Ax = b$$

is

$$x = A^{-1}b = A^T b$$

- the algorithm for solving the system of equations consists of transposing A and performing a matrix-vector multiply on the result
- requires $2n^2$ floating-point operations (more on this later)

Further topics

- pseudo-inverse
- projection onto the range of a matrix
- orthogonal decomposition of a vector