

## 4. Matrix inverses

- left and right inverse
- linear independence
- nonsingular matrices
- matrices with linearly independent columns
- matrices with linearly independent rows

# Left and right inverse

$AB \neq BA$  in general, so we have to distinguish two types of inverses

**Left inverse:**  $X$  is a *left inverse* of  $A$  if

$$XA = I$$

$A$  is *left-invertible* if it has at least one left inverse

**Right inverse:**  $X$  is a *right inverse* of  $A$  if

$$AX = I$$

$A$  is *right-invertible* if it has at least one right inverse

## Examples

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

- $A$  is left-invertible; the following matrices are left inverses:

$$\frac{1}{9} \begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1/2 & 3 \\ 0 & 1/2 & -2 \end{bmatrix}$$

- $B$  is right-invertible; the following matrices are right inverses:

$$\frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

# Some immediate properties

## Dimensions

a left or right inverse of an  $m \times n$  matrix must have size  $n \times m$

## Left and right inverse of (conjugate) transpose

- $X$  is a left inverse of  $A$  if and only if  $X^T$  is a right inverse of  $A^T$

$$A^T X^T = (XA)^T = I$$

- $X$  is a left inverse of  $A$  if and only if  $X^H$  is a right inverse of  $A^H$

$$A^H X^H = (XA)^H = I$$

# Inverse

if  $A$  has a left **and** a right inverse, then they are equal and unique:

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

- in this case, we call  $X = Y$  the **inverse** of  $A$  (notation:  $A^{-1}$ )
- $A$  is *invertible* if its inverse exists

## Example

$$A = \begin{bmatrix} -1 & 1 & -3 \\ 1 & -1 & 1 \\ 2 & 2 & 2 \end{bmatrix}, \quad A^{-1} = \frac{1}{4} \begin{bmatrix} 2 & 4 & 1 \\ 0 & -2 & 1 \\ -2 & -2 & 0 \end{bmatrix}$$

# Linear equations

set of  $m$  linear equations in  $n$  variables

$$\begin{aligned}A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\&\vdots \\A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m\end{aligned}$$

- in matrix form:  $Ax = b$
- may have no solution, a unique solution, infinitely many solutions

# Linear equations and matrix inverse

**Left-invertible matrix:** if  $X$  is a left inverse of  $A$ , then

$$Ax = b \quad \implies \quad x = XAx = Xb$$

there is *at most one* solution (if there is a solution, it must be equal to  $Xb$ )

**Right-invertible matrix:** if  $X$  is a right inverse of  $A$ , then

$$x = Xb \quad \implies \quad Ax = AXb = b$$

there is *at least one* solution (namely,  $x = Xb$ )

**Invertible matrix:** if  $A$  is invertible, then

$$Ax = b \quad \iff \quad x = A^{-1}b$$

there is a *unique* solution

# Outline

- left and right inverse
- **linear independence**
- nonsingular matrices
- matrices with linearly independent columns
- matrices with linearly independent rows



# Linear combination

a linear combination of vectors  $a_1, \dots, a_n$  is a sum of scalar-vector products

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n$$

- the scalars  $x_i$  are the *coefficients* of the linear combination
- can be written as a matrix-vector product

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

- the *trivial* linear combination has coefficients  $x_1 = \cdots = x_n = 0$

(same definition holds for real and complex vectors/scalars)

# Linear dependence

a collection of vectors  $a_1, a_2, \dots, a_n$  is *linearly dependent* if

$$x_1 a_1 + x_2 a_2 + \dots + x_n a_n = 0$$

for some scalars  $x_1, \dots, x_n$ , not all zero

- the vector 0 can be written as a nontrivial linear combination of  $a_1, \dots, a_n$
- equivalently, at least one vector  $a_i$  is a linear combination of the other vectors:

$$a_i = -\frac{x_1}{x_i} a_1 - \dots - \frac{x_{i-1}}{x_i} a_{i-1} - \frac{x_{i+1}}{x_i} a_{i+1} - \dots - \frac{x_n}{x_i} a_n$$

if  $x_i \neq 0$

## Example

the vectors

$$a_1 = \begin{bmatrix} 0.2 \\ -7 \\ 8.6 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -0.1 \\ 2 \\ -1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ -1 \\ 2.2 \end{bmatrix}$$

are linearly dependent

- 0 can be expressed as a nontrivial linear combination of  $a_1, a_2, a_3$ :

$$0 = a_1 + 2a_2 - 3a_3$$

- $a_1$  can be expressed as a linear combination of  $a_2, a_3$ :

$$a_1 = -2a_2 + 3a_3$$

(and similarly  $a_2$  and  $a_3$ )

# Linear independence

vectors  $a_1, \dots, a_n$  are *linearly independent* if they are not linearly dependent

- the zero vector cannot be written as a nontrivial linear combination:

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0 \quad \implies \quad x_1 = x_2 = \cdots = x_n = 0$$

- none of the vectors  $a_i$  is a linear combination of the other vectors

## Matrix with linearly independent columns

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

has linearly independent columns if

$$Ax = 0 \quad \implies \quad x = 0$$

## Example

the vectors

$$a_1 = \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad a_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

are linearly independent:

$$x_1 a_1 + x_2 a_2 + x_3 a_3 = \begin{bmatrix} x_1 - x_2 \\ -2x_1 + x_3 \\ x_2 + x_3 \end{bmatrix} = 0$$

holds only if  $x_1 = x_2 = x_3 = 0$

## Dimension inequality

if  $n$  vectors  $a_1, a_2, \dots, a_n$  of length  $m$  are linearly independent, then

$$n \leq m$$

(proof is in textbook)

- if an  $m \times n$  matrix has linearly independent columns then  $m \geq n$
- if an  $m \times n$  matrix has linearly independent rows then  $m \leq n$

# Outline

- left and right inverse
- linear independence
- **nonsingular matrices**
- matrices with linearly independent columns
- matrices with linearly independent rows

# Nonsingular matrix

for a **square** matrix  $A$  the following four properties are equivalent

1.  $A$  is left-invertible
2. the columns of  $A$  are linearly independent
3.  $A$  is right-invertible
4. the rows of  $A$  are linearly independent

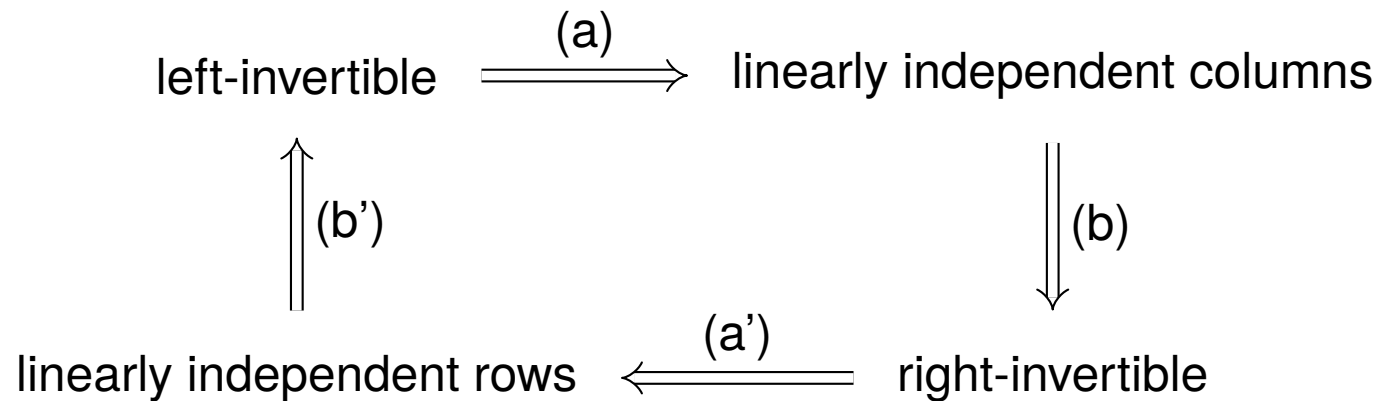
a square matrix with these properties is called **nonsingular**

## Nonsingular = invertible

- if properties 1 and 3 hold, then  $A$  is invertible (page 4.5)
- if  $A$  is invertible, properties 1 and 3 hold (by definition of invertibility)



# Proof



- we show that (a) holds in general
- we show that (b) holds for square matrices
- (a') and (b') follow from (a) and (b) applied to  $A^T$

**Part a:** suppose  $A$  is left-invertible

- if  $B$  is a left inverse of  $A$  (satisfies  $BA = I$ ), then

$$\begin{aligned} Ax = 0 &\implies BAx = 0 \\ &\implies x = 0 \end{aligned}$$

- this means that the columns of  $A$  are linearly independent: if

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

then

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = 0$$

holds only for the trivial linear combination  $x_1 = x_2 = \cdots = x_n = 0$

**Part b:** suppose  $A$  is square with linearly independent columns  $a_1, \dots, a_n$

- for every  $n$ -vector  $b$  the vectors  $a_1, \dots, a_n, b$  are linearly dependent  
(from dimension inequality on page 4.13)
- hence for every  $b$  there exists a nontrivial linear combination

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n + x_{n+1} b = 0$$

- we must have  $x_{n+1} \neq 0$  because  $a_1, \dots, a_n$  are linearly independent
- hence every  $b$  can be written as a linear combination of  $a_1, \dots, a_n$
- in particular, there exist  $n$ -vectors  $c_1, \dots, c_n$  such that

$$Ac_1 = e_1, \quad Ac_2 = e_2, \quad \dots, \quad Ac_n = e_n,$$

- the matrix  $C = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$  is a right inverse of  $A$ :

$$A \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix} = \begin{bmatrix} e_1 & e_2 & \cdots & e_n \end{bmatrix} = I$$

## Examples

$$A = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

- $A$  is nonsingular because its columns are linearly independent:

$$x_1 - x_2 + x_3 = 0, \quad -x_1 + x_2 + x_3 = 0, \quad x_1 + x_2 - x_3 = 0$$

is only possible if  $x_1 = x_2 = x_3 = 0$

- $B$  is singular because its columns are linearly dependent:

$$Bx = 0 \quad \text{for } x = (1, 1, 1, 1)$$

## Example: Vandermonde matrix

$$A = \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^{n-1} \\ 1 & t_2 & t_2^2 & \cdots & t_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & t_n^2 & \cdots & t_n^{n-1} \end{bmatrix} \quad \text{with } t_i \neq t_j \text{ for } i \neq j$$

we show that  $A$  is nonsingular by showing that  $Ax = 0$  only if  $x = 0$

- $Ax = 0$  means  $p(t_1) = p(t_2) = \cdots = p(t_n) = 0$  where

$$p(t) = x_1 + x_2t + x_3t^2 + \cdots + x_nt^{n-1}$$

$p(t)$  is a polynomial of degree  $n - 1$  or less

- if  $x \neq 0$ , then  $p(t)$  can not have more than  $n - 1$  distinct real roots
- therefore  $p(t_1) = \cdots = p(t_n) = 0$  is only possible if  $x = 0$

# Inverse of transpose and product

## Transpose and conjugate transpose

if  $A$  is nonsingular, then  $A^T$  and  $A^H$  are nonsingular and

$$(A^T)^{-1} = (A^{-1})^T, \quad (A^H)^{-1} = (A^{-1})^H$$

we write these as  $A^{-T}$  and  $A^{-H}$

## Product

if  $A$  and  $B$  are nonsingular and of equal size, then  $AB$  is nonsingular with

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

# Outline

- left and right inverse
- linear independence
- nonsingular matrices
- **matrices with linearly independent columns**
- matrices with linearly independent rows

# Gram matrix

the *Gram matrix* associated with a matrix

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$$

is the matrix of column inner products

- for real matrices:

$$A^T A = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{bmatrix}$$

- for complex matrices

$$A^H A = \begin{bmatrix} a_1^H a_1 & a_1^H a_2 & \cdots & a_1^H a_n \\ a_2^H a_1 & a_2^H a_2 & \cdots & a_2^H a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^H a_1 & a_n^H a_2 & \cdots & a_n^H a_n \end{bmatrix}$$



# Nonsingular Gram matrix

the Gram matrix is nonsingular if only if  $A$  has linearly independent columns

- suppose  $A \in \mathbf{R}^{m \times n}$  has linearly independent columns:

$$\begin{aligned} A^T A x = 0 &\implies x^T A^T A x = (Ax)^T (Ax) = \|Ax\|^2 = 0 \\ &\implies Ax = 0 \\ &\implies x = 0 \end{aligned}$$

therefore  $A^T A$  is nonsingular

- suppose the columns of  $A \in \mathbf{R}^{m \times n}$  are linearly dependent

$$\exists x \neq 0, Ax = 0 \implies \exists x \neq 0, A^T A x = 0$$

therefore  $A^T A$  is singular

(for  $A \in \mathbf{C}^{m \times n}$ , replace  $A^T$  with  $A^H$  and  $x^T$  with  $x^H$ )

# Pseudo-inverse

- suppose  $A \in \mathbf{R}^{m \times n}$  has linearly independent columns
- this implies that  $A$  is tall or square ( $m \geq n$ ); see page 4.13

the *pseudo-inverse* of  $A$  is defined as

$$A^\dagger = (A^T A)^{-1} A^T$$

- this matrix exists, because the Gram matrix  $A^T A$  is nonsingular
- $A^\dagger$  is a left inverse of  $A$ :

$$A^\dagger A = (A^T A)^{-1} (A^T A) = I$$

(for complex  $A$  with linearly independent columns,  $A^\dagger = (A^H A)^{-1} A^H$ )

# Summary

the following three properties are equivalent for a real matrix  $A$

1.  $A$  is left-invertible
2. the columns of  $A$  are linearly independent
3.  $A^T A$  is nonsingular

- $1 \Rightarrow 2$  was already proved on page 4.16
- $2 \Rightarrow 1$ : we have seen that the pseudo-inverse is a left inverse
- $2 \Leftrightarrow 3$ : proved on page 4.22
- a matrix with these properties must be tall or square
- for complex matrices, replace  $A^T A$  in property 3 by  $A^H A$

# Outline

- left and right inverse
- linear independence
- nonsingular matrices
- matrices with linearly independent columns
- **matrices with linearly independent rows**

## Pseudo-inverse

- suppose  $A \in \mathbf{R}^{m \times n}$  has linearly independent rows
- this implies that  $A$  is wide or square ( $m \leq n$ ); see page 4.13

the *pseudo-inverse* of  $A$  is defined as

$$A^\dagger = A^T(AA^T)^{-1}$$

- $A^T$  has linearly independent columns
- hence its Gram matrix  $AA^T$  is nonsingular, so  $A^\dagger$  exists
- $A^\dagger$  is a right inverse of  $A$ :

$$AA^\dagger = (AA^T)(AA^T)^{-1} = I$$

(for complex  $A$  with linearly independent rows,  $A^\dagger = A^H(AA^H)^{-1}$ )

# Summary

the following three properties are equivalent

1.  $A$  is right-invertible
  2. the rows of  $A$  are linearly independent
  3.  $AA^T$  is nonsingular
- 
- $1 \Rightarrow 2$  and  $2 \Leftrightarrow 3$ : by transposing result on page 4.24
  - $2 \Rightarrow 1$ : we have seen that the pseudo-inverse is a right inverse
  - a matrix with these properties must be wide or square
  - for complex matrices, replace  $AA^T$  in property 3 by  $AA^H$