Matrix Inverse

CE40282-1: Linear Algebra

Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

Inverse

- For $A \in M_{n\times n}$, if there exists a matrix $B \in M_{n\times n}$ such that $AB = BA = I_n$, then:
  - A is invertible (or nonsingular)
  - B is the inverse of A
- A square matrix that does not have an inverse is called noninvertible (or singular)
- Note
  - The definition of the inverse of a matrix is similar to that of the inverse of a scalar, i.e., $c \cdot (1/c) = 1$
  - Since there is no inverse (or said multiplicative inverse for the real number 0, you can "imagine" that noninvertible matrices act a similar role to the real number 0 is some sense

Inverse

- The inverse of A is denoted by A^{-1}
- Theorem: The inverse of a matrix is unique
  - Proof?

- Find the inverse of a matrix by the Gauss-Jordan Elimination:

$$[A \quad | \quad I] \xrightarrow{\text{Gauss-Jordan Elimination}} [I \quad | \quad A^{-1}]$$

Gauss-Jordan Elimination for finding the Inverse of a Matrix

- Let A be an $n \times n$ matrix.
  - Adjoin the identity $n \times n$ matrix I_n to A to form the matrix $[A : I_n]$.
  - Compute the reduced echelon form of $[A : I_n]$.
- <u>If the reduced echelon form is of the type $[I_n : B]$, then B is the inverse of A.</u>
- If the reduced echelon form is not of the type $[I_n : B]$, in that the first $n \times n$ submatrix is not I_n, then A has no inverse.

An $n \times n$ matrix A is invertible if and only if its reduced echelon form is I_n.

A is row equivalent to I_n

Inverse (Example)

$$A = \begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}$$

$$AX = I$$

$$\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} x_{11}+4x_{21} & x_{12}+4x_{22} \\ -x_{11}-3x_{21} & -x_{12}-3x_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

by equating corresponding entries

$$\Rightarrow \begin{array}{rcrcl} x_{11} & + & 4x_{21} & = & 1 \\ -x_{11} & - & 3x_{21} & = & 0 \end{array} \quad (1)$$

$$\begin{array}{rcrcl} x_{12} & + & 4x_{22} & = & 0 \\ -x_{12} & - & 3x_{22} & = & 1 \end{array} \quad (2)$$

This two systems of linear equations have the same coefficient matrix, which is exactly the matrix A.

$$(1) \Rightarrow \left[\begin{array}{cc:c} 1 & 4 & 1 \\ -1 & -3 & 0 \end{array}\right] \xrightarrow{A_{1,2}^{(1)},\ A_{2,1}^{(-4)}} \left[\begin{array}{cc:c} 1 & 0 & -3 \\ 0 & 1 & 1 \end{array}\right] \Rightarrow x_{11} = -3,\ x_{21} = 1$$

$$(2) \Rightarrow \left[\begin{array}{cc:c} 1 & 4 & 0 \\ -1 & -3 & 1 \end{array}\right] \xrightarrow{A_{1,2}^{(1)},\ A_{2,1}^{(-4)}} \left[\begin{array}{cc:c} 1 & 0 & -4 \\ 0 & 1 & 1 \end{array}\right] \Rightarrow x_{12} = -4,\ x_{22} = 1$$

Perform the Gauss-Jordan elimination on the matrix A with the same row operations

Thus

$$X = A^{-1} = \begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}$$

$$\underset{A}{\begin{bmatrix} 1 & 4 \\ -1 & -3 \end{bmatrix}} \vdots \underset{I}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \xrightarrow[A_{1,2}^{(1)},A_{2,1}^{(-4)}]{\text{Gauss-Jordan Elimination}} \underset{I}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \vdots \underset{A^{-1}}{\begin{bmatrix} -3 & -4 \\ 1 & 1 \end{bmatrix}}$$

solution for $\begin{bmatrix} x_{11} \\ x_{21} \end{bmatrix}$ solution for $\begin{bmatrix} x_{12} \\ x_{22} \end{bmatrix}$

Inverse

- Properties (If A is an invertible matrix, k is a positive integer, and c is a scalar):
  - A^{-1} is invertible and $(A^{-1})^{-1} = A$
  - A^k is invertible and $(A^k)^{-1} = A^{-k} = (A^{-1})^k$
  - cA is invertible if $c \neq 0$ and $(cA)^{-1} = \frac{1}{c}A^{-1}$
  - A^T is invertible and $(A^T)^{-1} = (A^{-1})^T$
- Theorem: If A and B are invertible matrices of order n, then AB is invertible and $(\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1}$
  - Proof?

$$(\boldsymbol{A_1 A_2 A_3 \cdots A_n})^{-1} = \boldsymbol{A_n^{-1} \cdots A_3^{-1} A_2^{-1} A_1^{-1}}$$

Inverse

- Theorem: Let $AX = B$ be a system of n linear equations in n variables.
  If A^{-1} exists, the solution is unique and is given by $X = A^{-1}B$.

Invertible Matrix

Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $ad - bc \neq 0$, then A is invertible and

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

If $ad - bc = 0$, then A is not invertible.

$$\det A = ad - bc$$

2×2 matrix A is invertible if and only if $\det A \neq 0$.

Elementary Matrices

Each elementary matrix E is invertible. The inverse of E is the elementary matrix of the same type that transforms E back into I.

- Example

Find the inverse of $E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}$.

Left inverses

- a number x that satisfies $xa = 1$ is called the inverse of a
- inverse (*i.e.*, $1/a$) exists if and only if $a \neq 0$, and is unique
- a matrix X that satisfies $XA = I$ is called a *left inverse* of A
- if a left inverse exists we say that A is *left-invertible*

$$A : m \times n \implies I : n \times n \implies X : n \times m$$

- example: the matrix

$$A = \begin{bmatrix} -3 & -4 \\ 4 & 6 \\ 1 & 1 \end{bmatrix}$$

has two different left inverses:

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

Left inverses of vector

- A non-zero column vector always has a left inverse.
- Left inverse is not unique.
- Example
  - $a = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$
    - Matrix with orthonormal columns
- Row vector does not have left inverse

$$a = \begin{bmatrix} 1 & 0 & 3 \end{bmatrix}$$

Left inverse and column independence

- if A has a left inverse C then the columns of A are linearly independent
- we'll see later the converse is also true, so

  a matrix is left-invertible if and only if its columns are linearly independent
- matrix generalization of

  a number is invertible if and only if it is nonzero

- so left-invertible matrices are tall or square
  - Wide matrix always are not left invertible.
  - Tall or square matrices can be left invertible.
  - Go to pervious slide

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & -1 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 4 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix}, \begin{bmatrix} 1 & -2 & -1 \\ 1 & 3 & 4 \\ -2 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 & -1 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Solving linear equations with a left inverse

- suppose $Ax = b$, and A has a left inverse C
- then $Cb = C(Ax) = (CA)x = Ix = x$
- so multiplying the right-hand side by a left inverse yields the solution

Right inverses

- Duality between right and left inverse

- a matrix X that satisfies $AX = I$ is a *right inverse* of A
- if a right inverse exists we say that A is *right-invertible*
- A is right-invertible if and only if A^T is left-invertible:

$$AX = I \iff (AX)^T = I \iff X^T A^T = I$$

- so we conclude

$$\boxed{A \text{ is right-invertible if and only if its rows are linearly independent}}$$

- right-invertible matrices are wide or square

Solving linear equations with a right inverse

- suppose A has a right inverse B
- consider the (square or underdetermined) equations $Ax = b$
- $x = Bb$ is a solution:

- so $Ax = b$ has a solution for *any* b

Example

- same A, B, C in example above
- C^T and B^T are both right inverses of A^T
- under-determined equations $A^T x = (1, 2)$ has (different) solutions

$$B^T(1,2) = (1/3, 2/3, -2/3), \qquad C^T(1,2) = (0, 1/2, -1)$$

(there are many other solutions as well)

Inverse

- The only matrix having both left and right inverse is ……
- Theorem: For a square matrix left and right inverse are the same. Rows and columns are linear independent
  - Proof? Next slides!

Inverse

- if A has a left and a right inverse, they are unique and equal (and we say that A is *invertible*)
- so A must be square
- to see this: if $AX = I$, $YA = I$
- we denote them by A^{-1}:

$$A^{-1}A = AA^{-1} = I$$

- inverse of inverse: $(A^{-1})^{-1} = A$

Invertible matrices

$$
\begin{tikzcd}
\text{left-invertible} \arrow[r, Rightarrow, "\text{(a)}"] & \text{linearly independent columns} \arrow[d, Rightarrow, "\text{(b)}", "\text{Holds for square matrices}"'] \\
\text{linearly independent rows} \arrow[u, Rightarrow, "\text{(b')}"'] & \text{right-invertible} \arrow[l, Rightarrow, "\text{(a')}"']
\end{tikzcd}
$$

Holds in general

(a') and (b') follow from (a) and (b) applied to A^T

Solving square systems of linear equations

- suppose A is invertible
- for any b, $Ax = b$ has the unique solution

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

- matrix generalization of simple scalar equation $ax = b$ having solution $x = (1/a)b$ (for $a \neq 0$)
- simple-looking formula $x = A^{-1}b$ is basis for many applications

Linear equations and matrix inverse

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

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

we must know in advance that there *exists* at least one solution.

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

- Why "at most"??

$$\mathbf{DC} = \mathbf{I}$$

$$\begin{cases} -y_1 + y_2 = -4 \\ 0y_1 - y_2 = 3 \\ 2y_1 + y_2 = 0 \end{cases} \quad \mathbf{C} = \begin{bmatrix} -1 & 1 \\ 0 & -1 \\ 2 & 1 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 5 & 2 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} -1 & 1 & -4 \\ 0 & -1 & 3 \\ 2 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{array}\right]$$

Linear equations and matrix inverse

If the system of equations $\mathbf{Ax} = \mathbf{b}$ *is consistent, and if a matrix* $\mathbf{B}$ *exists such that* $\mathbf{BA} = \mathbf{I}$, *then the system of equations has a unique solution, namely* $\mathbf{x} = \mathbf{Bb}$.

Linear equations and matrix inverse

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

$$x = Xb \qquad \Longrightarrow \qquad Ax = AXb = b$$

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

To pursue these ideas further, suppose that again we want to solve a system of linear equations, $\mathbf{Ax} = \mathbf{b}$. Assume now that we have another matrix, $\mathbf{B}$, such that $\mathbf{AB} = \mathbf{I}$. Then we can write $\mathbf{A}(\mathbf{Bb}) = (\mathbf{AB})\mathbf{b} = \mathbf{Ib} = \mathbf{b}$; whence $\mathbf{Bb}$ solves the equation $\mathbf{Ax} = \mathbf{b}$. This conclusion did *not* require an *a priori* assumption that a solution exist; we have *produced* a solution. The argument does not reveal whether $\mathbf{Bb}$ is the only solution. There may be others.

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

$$Ax = b \qquad \Longleftrightarrow \qquad x = A^{-1}b$$

there is a *unique* solution

Linear equations and matrix inverse

- system of linear equations $\mathbf{Ax} = \mathbf{b}$
  - a right inverse of $\mathbf{A}$, say $\mathbf{AB} = \mathbf{I}$. Then $\mathbf{Bb}$ is a solution, as is verified by noting $\mathbf{A}(\mathbf{Bb}) = (\mathbf{AB})\mathbf{b} = \mathbf{Ib} = \mathbf{b}$.
  - a left inverse of $\mathbf{A}$, say $\mathbf{CA} = \mathbf{I}$, then we can only conclude that $\mathbf{Cb}$ is the *sole candidate* for a solution ; however, it must be checked by substitution to determine whether, in fact, it *is* a solution.

Invertible (Nonsingular) matrices

the following are equivalent for a square matrix A:

- A is invertible
- columns of A are linearly independent
- rows of A are linearly independent
- A has a left inverse
- A has a right inverse

$$\text{row rank}(A) = \text{col rank}(A) = n$$

if any of these hold, all others do

Invertible matrices

Examples

- $I^{-1} = I$
- if Q is orthogonal, *i.e.*, square with $Q^T Q = I$, then $Q^{-1} = Q^T$
- 2×2 matrix A is invertible if and only $A_{11}A_{22} \neq A_{12}A_{21}$

$$A^{-1} = \frac{1}{A_{11}A_{22} - A_{12}A_{21}} \begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$$

  - you need to know this formula
  - there are similar but *much* more complicated formulas for larger matrices (and no, you do not need to know them)

Invertible matrices

Examples

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

- A is invertible, with inverse

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

- verified by checking $AA^{-1} = I$ (or $A^{-1}A = I$)
- we'll soon see how to compute the inverse

Properties

- $(AB)^{-1} = B^{-1}A^{-1}$ (provided inverses exist)
- if A is nonsingular, then A^T are nonsingular

  $(A^T)^{-1} = (A^{-1})^T$ (sometimes denoted A^{-T})
- negative matrix powers: $(A^{-1})^k$ is denoted A^{-k}
- with $A^0 = I$, identity $A^k A^l = A^{k+l}$ holds for any integers k, l

Triangular matrices

- Lower triangular L with nonzero diagonal entries is invertible
  - Proof?

- Upper triangular R with nonzero diagonal entries is invertible
  - Proof?

Gram Matrix

Consider an $n \times m$ matrix A over $\mathbb{R}$, where

$$A = \begin{bmatrix} x_1 & \cdots & x_m \end{bmatrix}$$

The $m \times m$ matrix $A^T A$ is :

$$A^T A = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 & \cdots & x_1^T x_m \\ x_2^T x_1 & x_2^T x_2 & \cdots & x_2^T x_m \\ \vdots & \vdots & \ddots & \vdots \\ x_m^T x_1 & x_m^T x_2 & \cdots & x_m^T x_m \end{bmatrix}$$

Note that $A\colon \mathbb{R}^m \to \mathbb{R}^n$ and $A^T A\colon \mathbb{R}^m \to \mathbb{R}^m$. We've already seen that:

1. $\operatorname{rank} A = \operatorname{rank} A^T A$ and $\operatorname{nullity} A = \operatorname{nullity} A^T A$ (in fact, $N_A = N_{A^T A}$)
2. $A^T A \geq 0$
3. If $N_A = 0$, then the projection matrix onto $\operatorname{Span}(x_1, \ldots, x_m)$ is $A(A^T A)^{-1} A^T$.

This is an example of a Gram matrix.

Gram Matrix

- A Gram matrix is Positive Definite and Symmetric
- $G = AA^T$ is left Gram matrix
- Gram Matrix and Left Gram Matrix are symmetric

- Null space: $N(A^T A) = N(A)$

- Rank: $\text{rank}(A^T A) = rank(A) = n - nullity(A)$

$$C(A^T A) = R(A^T A) = R(A)$$
$$C(AA^T) = R(AA^T) = C(A)$$

Review: Orthonormal Vectors

- a collection of real m-vectors $a_1, a_2, \ldots, a_n$ is *orthonormal* if
  - the vectors have unit norm: $\|a_i\| = 1$
  - they are mutually orthogonal: $a_i^T a_j = 0$ if $i \neq j$

Example

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

Orthogonal Matrix

- Columns of $A_{n\times k} = [a_1 \quad \cdots \quad a_k]$ are orthonormal.

$$n \geq k$$

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

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

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

"matrix with orthonormal columns"

Orthogonal Matrix

- Columns of A are orthonormal $\leftrightarrow$ $\boxed{A^T A = I}$
- Square matrix with orthonormal columns is a orthogonal matrix
  - columns and rows are orthonormal vectors
  - $Q^{\mathrm{T}}Q = QQ^{\mathrm{T}} = I$
  - is necessarily invertible with inverse $Q^{\mathrm{T}} = Q^{-1}$

Orthogonal Matrix

- Examples
  - Identity matrix $I^T I = I$
  - Rotation matrix

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

$$= \begin{bmatrix} \cos^2\theta + \sin^2\theta & -\cos\theta\sin\theta + \sin\theta\cos\theta \\ -\sin\theta\cos\theta + \cos\theta\sin\theta & \sin^2\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

  - Reflection matrix

$$\begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}^T \begin{bmatrix} \cos(2\theta) & \sin(2\theta) \\ \sin(2\theta) & -\cos(2\theta) \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2(2\theta) + \sin^2(2\theta) & \cos(2\theta)\sin(2\theta) - \sin(2\theta)\cos(2\theta) \\ \sin(2\theta)\cos(2\theta) - \cos(2\theta)\sin(2\theta) & \sin^2(2\theta) + \cos^2(2\theta) \end{bmatrix}$$

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

Orthogonal Matrix

- All 2x2 orthogonal matrices can be expressed as Rotation or Reflection

Orthonormal Columns Properties

if $A \in \mathbf{R}^{m \times n}$ has orthonormal columns, then the linear function $f(x) = Ax$

This is a mapping with preserving properties of input

- preserves inner products:

$$(Ax)^T(Ay) =$$

- preserves norms:

$$\|Ax\| =$$

- preserves distances: $\|Ax - Ay\| = \|x - y\|$

- preserves angles:

$$\angle(Ax, Ay) = \arccos\left(\frac{(Ax)^T(Ay)}{\|Ax\|\|Ay\|}\right) = \arccos\left(\frac{x^T y}{\|x\|\|y\|}\right) = \angle(x, y)$$

Gram–Schmidt in matrix notation

- run Gram–Schmidt on columns $a_1, \ldots, a_k$ of $n \times k$ matrix A $\quad n \geq k$

$$\tilde{q}_1 = a_1, \quad q_1 = \frac{\tilde{q}_1}{\|\tilde{q}_1\|}$$

$$\Longrightarrow \; a_1 = \|\tilde{q}_1\| q_1$$

$$\tilde{q}_2 = a_2 - (q_1^T a_2) q_1, \quad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|}$$

$$\Longrightarrow \; a_2 = (q_1^T a_2) q_1 + \|\tilde{q}_2\| q_2$$

$$\vdots$$

$$\tilde{q}_i = a_i - (q_1^T a_i) q_1 - \cdots - (q_{i-1}^T a_i) q_{i-1}, \quad q_i = \frac{\tilde{q}_i}{\|\tilde{q}_i\|}$$

$$a_i = (q_1^T a_i) q_1 + \cdots + (q_{i-1}^T a_i) q_{i-1} + \|\tilde{q}_i\| q_i$$

Gram–Schmidt in matrix notation

$$
\begin{aligned}
a_1 &= \|\tilde{q}_1\| q_1 \\
a_2 &= (q_1^T a_2) q_1 + \|\tilde{q}_2\| q_2 \\
&\vdots \\
a_k &= (q_1^T a_k) q_1 + \cdots + (q_{k-1}^T a_k) q_{k-1} + \|\tilde{q}_k\| q_k
\end{aligned}
$$

$$
\begin{bmatrix} a_1 & a_2 & \dots & a_k \end{bmatrix} = \begin{bmatrix} q_1 & q_2 & \dots & q_k \end{bmatrix} \begin{bmatrix} \|\tilde{q}_1\| & q_1^T a_2 & \dots & q_1^T a_k \\ 0 & \|\tilde{q}_2\| & \dots & q_2^T a_k \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & q_{k-1}^T a_k \\ 0 & 0 & \dots & \|\tilde{q}_k\|_1 \end{bmatrix}
$$

$$
A_{n \times k} = Q_{n \times k} \times R_{k \times k}
$$

Gram–Schmidt in matrix notation

- run Gram–Schmidt on columns $a_1, \ldots, a_k$ of $n \times k$ matrix A $\quad n \geq k$
- if columns are linearly independent, get orthonormal $q_1, \ldots, q_k$
- define $n \times k$ matrix Q with columns $q_1, \ldots, q_k$
- $Q^T Q = I$
- from Gram–Schmidt algorithm

$$\begin{aligned} a_i &= (q_1^T a_i) q_1 + \cdots + (q_{i-1}^T a_i) q_{i-1} + \|\tilde{q}_i\| q_i \\ &= R_{1i} q_1 + \cdots + R_{ii} q_i \end{aligned}$$

  with $R_{ij} = q_i^T a_j$ for $i < j$ and $R_{ii} = \|\tilde{q}_i\|$

- defining $R_{ij} = 0$ for $i > j$ we have $A = QR$
- R is upper triangular, with positive diagonal entries

QR factorization

- $A = QR$ is called *QR factorization* of A
- factors satisfy $Q^T Q = I$, R upper triangular with positive diagonal entries
- can be computed using Gram–Schmidt algorithm (or some variations)
- has a *huge* number of uses, which we'll see soon

QR Decomposition (QU) (Factorization)

- To find QR decomposition:

  1.) Q: Use Gram-Schmidt to find orthonormal basis for column space of A

  2.) Let $R = Q^T A$

- If A is a square matrix, then Q is ……

QR Decomposition (QU) (Factorization)

- if $A \in \mathbf{R}^{m \times n}$ has linearly independent columns then it can be factored as

$$A = QR$$

Q-factor

- Q is $m \times n$ with orthonormal columns ($Q^T Q = I$)
- if A is square ($m = n$), then Q is orthogonal ($Q^T Q = QQ^T = I$)

R-factor

- R is $n \times n$, upper triangular, with nonzero diagonal elements
- R is nonsingular (diagonal elements are nonzero)

QR Decomposition

- Example

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}$$

$$q_1 = \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}, q_2 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, q_3 = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \|\tilde{q}_1\| = 2, \|\tilde{q}_2\| = 2, \|\tilde{q}_3\| = 4$$

- QR:

$$\begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 4 & 2 \\ 0 & 2 & 8 \\ 0 & 0 & 4 \end{bmatrix}$$

Generalization of QR Decompose

$$A_{4\times 6} = [a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6]$$

Linear Independent

$$\begin{cases} a_1 = \alpha_{11} q_1 \\ a_2 = \alpha_{21} q_1 + \alpha_{22} q_2 \\ a_3 = \alpha_{31} q_1 + \alpha_{32} q_2 \\ a_4 = \alpha_{41} q_1 + \alpha_{42} q_2 + \alpha_{43} q_3 \\ a_5 = \alpha_{51} q_1 + \alpha_{52} q_2 + \alpha_{53} q_3 \\ a_6 = \alpha_{61} q_1 + \alpha_{62} q_2 + \alpha_{63} q_3 \end{cases}$$

block upper triangular matrix

$$[a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6] = [q_1 \quad q_2 \quad q_3] \begin{bmatrix} \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} & \alpha_{51} & \alpha_{61} \\ 0 & \alpha_{22} & \alpha_{32} & \alpha_{42} & \alpha_{52} & \alpha_{62} \\ 0 & 0 & 0 & \alpha_{43} & \alpha_{53} & \alpha_{63} \end{bmatrix}$$

$$A_{4\times 6} = Q_{4\times \boxed{3}} \times R_{3\times 6}$$

Inverse via QR factorization (square matrix)

- suppose A is square and invertible
- so its columns are linearly independent
- so Gram–Schmidt gives QR factorization
  - $A = QR$
  - Q is orthogonal: $Q^T Q = I$
  - R is upper triangular with positive diagonal entries, hence invertible
- so we have

$$A^{-1} = (QR)^{-1} = R^{-1}Q^{-1} = R^{-1}Q^T$$

Inverse via QR factorization (square matrix)

Algorithm: Computing Matrix Inverse

Input: $A_{n\times n}$ invertible

Output: $A^{-1}_{n\times n}$

Find QR factorization $A = QR$

$$\begin{bmatrix} \bar{q}_1 & \cdots & \bar{q}_n \end{bmatrix} = Q^T$$

for $i = 1, \ldots, n$ **do**

Solve $Rx_i = \bar{q}_i$ using back substituition

end

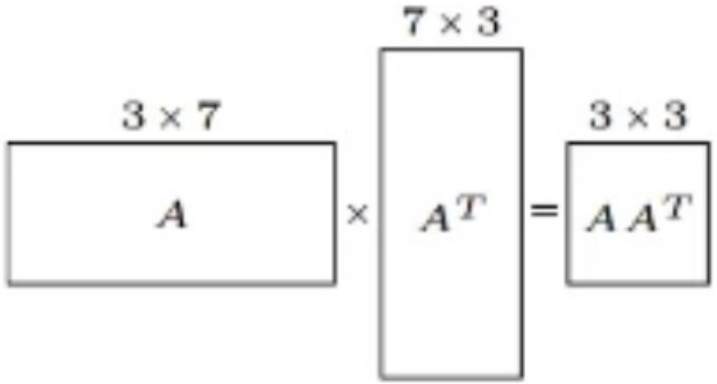
$$A^{-1} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

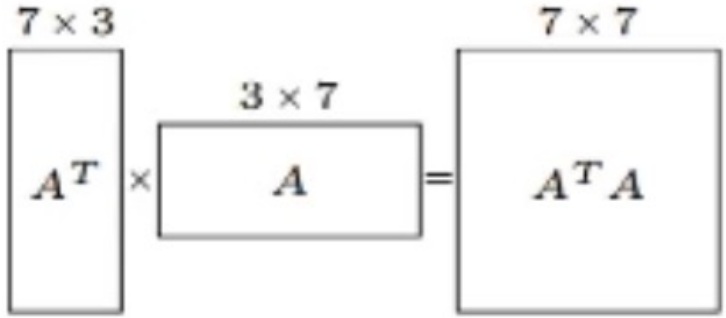
Invertibility of Gram matrix

- Theorem: Invertibility of Gram matrix: $A_{m\times n}$ has linearly independent columns if and only if $A^T A$ is invertible.
  - Proof?

Review

- $\text{rank}(A) = \text{rank}(A^T) = \text{rank}(A^T A) = \text{rank}(A A^T)$

- 



Pseudo-inverse of tall matrix

- $A_{m\times n}$ (tall or square) has linearly independent columns if and only if $A^T A$ is invertible.

- For a tall matrix with linear independent columns is one of its left inverse with this form:

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

Pseudo-inverse of wide matrix

- $A_{m\times n}$ (wide or square) has linearly independent columns if and only if AA^T is invertible.

- For a tall matrix with linear independent columns is one of its right inverse with this form:

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

Pseudo-inverse of square matrix

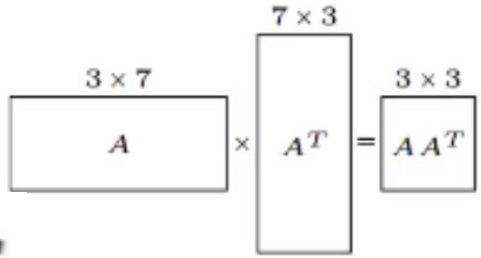
- For $A_{n\times n}$ Pseudo-inverse is the inverse

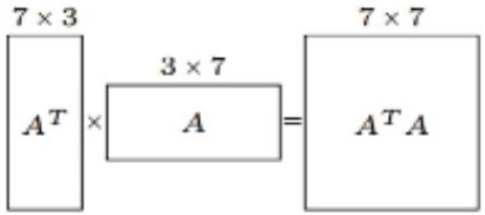
$$\begin{cases} A^{\dagger} = (A^T A)^{-1} A^T \\ A^{\dagger} = A^T (A A^T)^{-1} \end{cases} \qquad A^{\dagger} = A^{-1}$$

Pseudo-inverse

- For FULL RANK $A_{m\times n}$

$$A^{\dagger} = \begin{cases} (A^T A)^{-1} A^T & m \geq n \\ A^{-1} & m = n \\ A^T (A A^T)^{-1} & m \leq n \end{cases}$$





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

Pseudo-inverse via QR factorization

- Tall or square matrix: linearly independent columns

- Wide or square matrix: linearly independent rows

Inverse of Special Matrices

- $II = I \quad \rightarrow \quad I^{-1} = I$
- Diagonal matrix is invertible if and only if diagonal elements are non-zero

$$A = \begin{bmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{bmatrix} \Longrightarrow A^{-1} = \begin{bmatrix} \frac{1}{a_{11}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{a_{nn}} \end{bmatrix}$$

- Inverse of orthogonal matrix is its transpose.

$$A^T A = I \Longrightarrow A^{-1} = A^T$$

Inverse of block matrix

- A is a block upper triangular

$$A = \begin{bmatrix} \boxed{A_{11}} & A_{12} \\ 0 & \boxed{A_{22}} \end{bmatrix}$$

$\boldsymbol{p \times p}$ $\boldsymbol{q \times q}$

$$A^{-1} = ?$$

- A block diagonal matrix is invertible if each block on the diagonal is invertible.

Permutation matrix

- **Proposition** Let P be a $K \times K$ permutation matrix. Then, P is invertible and

$$P^{-1} = P^{\intercal}$$