QR Decomposition and Pseudo Inverse

CE282: Linear Algebra

Computer Engineering Department

Sharif University of Technology

Hamid R. Rabiee

Maryam Ramezani

Definition

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

$$A = [x_1 \ \cdots x_m]$$

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

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

This is a **Gram matrix**

Gram Matrix

Important

- A Gram matrix is Symmetric
- Gram Matrix and Left Gram Matrix are symmetric
- Null space: $N\left(A^T A\right) = N(A)$
- Rank: $\text{rank}\left(A^T A\right) = rank(A) = n - nullity(A)$

$$C{:}column\ space \qquad R{:}row\ space$$
$$C\left(A^T A\right) = R\left(A^T A\right) = R(A)$$
$$C\left(AA^T\right) = R\left(AA^T\right) = C(A)$$

Review: Orthonormal Vectors

Note

A collection of real m-vectors $a_1, a_2, \dots, a_n$ 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

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

Review: Orthogonal Matrix

Important

If the columns of $A_{n\times k} = [a_1, \ \dots, \ a_k\ ]$ are orthonormal, for $n \geq k$. Then :

$$A^T A = [a_1, a_2, \dots, a_n]^T [a_1, a_2, \dots, a_n] = \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \dots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \dots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \dots & a_n^T a_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

"matrix with orthonormal columns"

Orthogonal Matrix

Note

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

Orthogonal Matrix

Example

- ❑ Identity matrix $I^T I = I$

- ❑Rotation matrix

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

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

Orthogonal Matrix

Example

- Reflection matrix

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

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

Important

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

Orthonormal Columns Properties

Note

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

- Preserves inner product:

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

- Preserves norm:

$$\|Ax\| = \|x\|$$

This is a mapping with preserving properties of input

- Preserves distances:

$$\|Ax - Ay\| = \|x - y\|$$

- Preserves angels:

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

Gram–Schmidt in matrix notation

Important

Run Gram-Schmidt on columns $a_1, \dots, a_k$ of $n \times k$ matrix A:

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

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

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

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

$$\vdots$$

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

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

Gram–Schmidt in matrix notation

Important

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

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

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

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

Gram–Schmidt in matrix notation

Important

1. Run Gram-Schmidt on columns $a_1, \dots, a_k$ of n $\times$ k matrix A
2. If columns are linearly independent, get orthonormal $q_1, \dots, q_k$
3. Define n $\times$ k matrix Q with columns $q_1, \dots, q_k$
4. $Q^T Q = I$
5. 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 $R_{1j} = q_i^T a_j$ for $i < j$ and $R_{ii} = \|\tilde{q}_i\|$

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

QR factorization

Definition

A factorization of a matrix A as

$$A = QR$$

where Factors satisfy $Q^T Q = I$, R upper triangular with positive diagonal entries, is called a **QR factorization** of A.

Note

the QR factorization of a matrix :

- ❑ 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)

Important

To find QR decomposition:

- Q: Use Gram-Schmidt to find orthonormal basis for column space of A
- Let $R = Q^T A$

- If A is a square matrix, then Q is square and orthonormal (orthogonal)

QR Decomposition (QU) (Factorization)

Theorem

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

$$A = QR$$

Q-factor

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

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

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

Linear Independent

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

Block upper triangular matrix

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

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

Inverse via QR factorization (square matrix)

Note

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 $A = QR_1$

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

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

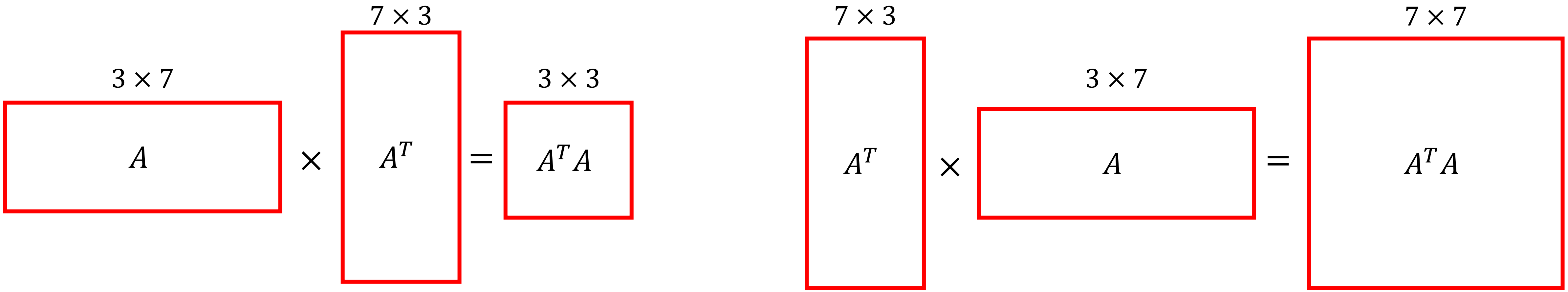
  Solve $Rx_i = \bar{q}_i$ using back substituition

end

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

Review

$$rank(A) = rank\left(A^T\right) = rank\left(A^T A\right) = rank(AA^T)$$



Pseudo-inverse of tall matrix

Theorem

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

Definition

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

Theorem

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

Definition

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

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

Pseudo-inverse of square matrix

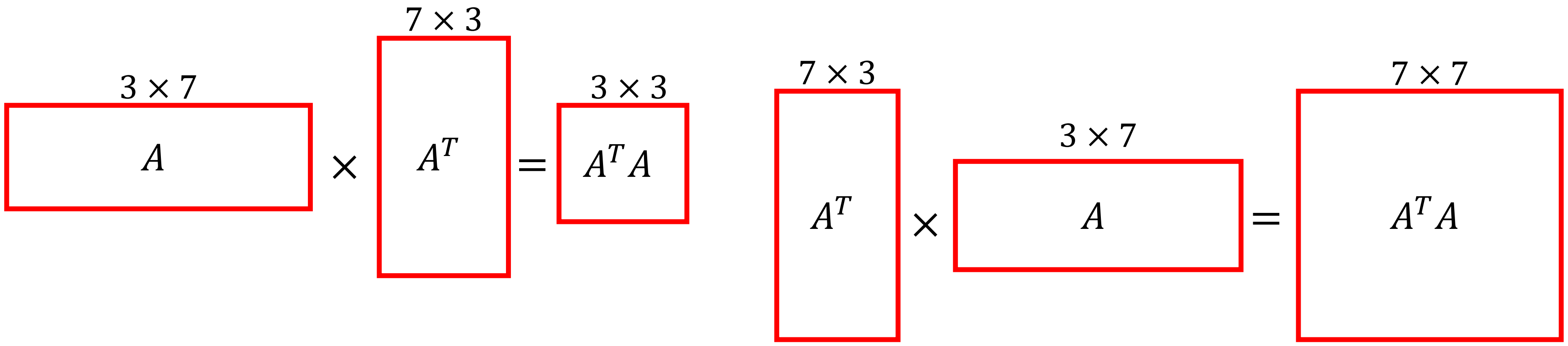
- For $A_{n\times n}$ Pseudo-inverse is the inverse

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

Pseudo-inverse

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

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



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

Pseudo-inverse via QR factorization

Important

- ❑Tall or square matrix: linearly independent columns
- ❑Wide or square matrix: linearly independent rows

Inverse of Special Matrices

Note

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

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

- Inverse of orthogonal matrix is its transpose.

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

Inverse of block matrix

- ❑ A is a block upper triangular

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

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

Theorem

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

Permutation matrix

Note

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

$$P^{-1} = P^T$$