QR Decompose and Matrix Inverse

CE40282-1: Linear Algebra
Hamid R. Rabiee and Maryam Ramezani
Sharif University of Technology

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

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

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

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

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.

$$n \geq k$$

$$
\begin{aligned}
A^T A &= \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}^T \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \\
&= \begin{bmatrix} a_1^T a_1 & a_1^T a_2 & \cdots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \cdots & a_2^T a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^T a_1 & a_n^T a_2 & \cdots & a_n^T a_n \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}
\end{aligned}
$$

"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
  - $Q^{\mathrm{T}} Q = Q Q^{\mathrm{T}} = I$
  - 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 $\boxed{n \geq k}$

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

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

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

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

$$\vdots$$

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

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

Gram–Schmidt in matrix notation

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

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

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

Gram–Schmidt in matrix notation

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

$$A_{4\times6} = [\underline{a_1 \quad a_2} \quad a_3 \quad \underline{a_4} \quad a_5 \quad a_6]$$

Linear Independent

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

block upper triangular matrix

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

$$A_{4\times6} = Q_{4\times\boxed{3}} \times R_{3\times6}$$

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 $\det A \neq 0$.

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*
- example: the matrix

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

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

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

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