ROB 101 - Fall 2021

The Vector Space $\mathbb{R}^n$: Part 2

(QR Factorization)

October 20, 2021

UNIVERSITY OF MICHIGAN

Learning Objectives

- A second encounter with some of the essential concepts in Linear Algebra.
- A more abstract view of $\mathbb{R}^n$ as a vector space.

Outcomes

- The QR Factorization is the most numerically robust method for solving systems of linear equations.

- Our recommended “pipeline” for solving $Ax = b$.

Suppose that that the set of vectors $\{u_1, u_2, \ldots, u_m\}$ is linearly independent and you generate a new set of vectors by

$$
\begin{aligned}
v_1 &= u_1, \\
v_2 &= u_2 - \left(\frac{u_2 \bullet v_1}{v_1 \bullet v_1}\right) v_1, \\
v_3 &= u_3 - \left(\frac{u_3 \bullet v_1}{v_1 \bullet v_1}\right) v_1 - \left(\frac{u_3 \bullet v_2}{v_2 \bullet v_2}\right) v_2, \\
&\ \vdots \\
v_k &= u_k - \sum_{i=1}^{k-1} \left(\frac{u_k \bullet v_i}{v_i \bullet v_i}\right) v_i, \quad \text{(General Step)}
\end{aligned}
$$

Then the set of vectors $\{v_1, v_2, \ldots, v_m\}$ is

- orthogonal, meaning, $i \neq j \implies v_i \bullet v_j = 0$
- span preserving, meaning that, for all $1 \leq k \leq m$,

$$\operatorname{span}\{v_1, v_2, \ldots, v_k\} = \operatorname{span}\{u_1, u_2, \ldots, u_k\},$$

- and linearly independent.

Remark

The unit vectors $\{e_1 = \frac{v_1}{\|v_1\|}, e_2 = \frac{v_2}{\|v_2\|}, \ldots, e_m = \frac{v_m}{\|v_m\|}\}$ *form an orthonormal set.*

Suppose that A is an $n \times m$ matrix with linearly independent columns.

Suppose that A is an $n \times m$ matrix with linearly independent columns.

Fact

Then there exists an $n \times m$ matrix Q with orthonormal columns and an upper triangular, $m \times m$, invertible matrix R such that $A = Q \cdot R$.

Q and R are constructed as follows:

- Let $\{u_1, \ldots, u_m\}$ be the columns of A with their order preserved so that

$$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}$$

Q and R are constructed as follows:

- Let $\{u_1, \ldots, u_m\}$ be the columns of A with their order preserved so that

$$A = \begin{bmatrix} u_1 & u_2 & \cdots & u_m \end{bmatrix}$$

- Q is constructed by applying the Gram-Schmidt Process to the columns of A and normalizing their lengths to one,

$$\{u_1, u_2, \ldots, u_m\} \xrightarrow[\text{Process}]{\text{Gram-Schmidt}} \{v_1, v_2, \ldots, v_m\}$$

$$Q := \begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} & \cdots & \frac{v_m}{\|v_m\|} \end{bmatrix}$$

- Because $Q^\mathsf{T} Q = I_m$, it follows that
  $A = Q \cdot R \iff R := Q^\mathsf{T} \cdot A$.

- Because $Q^{\mathsf{T}}Q = I_m$, it follows that $A = Q \cdot R \iff R := Q^{\mathsf{T}} \cdot A$.

- Recalling that the columns of A are linearly independent, if, and only if $x = 0$ is the unique solution to $Ax = 0$, we have that

$$x = 0 \iff Ax = 0 \iff Q \cdot Rx = 0 \iff$$
$$Q^{\mathsf{T}} \cdot Q \cdot Rx = Q^{\mathsf{T}} \cdot 0 \iff Rx = 0 \iff \det(R) \neq 0,$$

where the last step follows because R is square.

Compute the QR Factorization of $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix}$.

We extract the columns of A and obtain

$$\{u_1, u_2, u_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

We extract the columns of A and obtain

$$\{u_1, u_2, u_3\} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

and

$$\left\{ \tilde{v}_1 = \frac{v_1}{\|v_1\|} = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \tilde{v}_2 = \frac{v_2}{\|v_2\|} = \frac{\sqrt{38}}{38} \begin{bmatrix} -1 \\ 1 \\ 6 \end{bmatrix}, \tilde{v}_3 = \frac{v_3}{\|v_3\|} = \frac{\sqrt{19}}{19} \begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix} \right\}.$$

Therefore,

$$Q \approx \begin{bmatrix} 0.707107 & -0.162221 & -0.688247 \\ 0.707107 & 0.162221 & 0.688247 \\ 0.000000 & 0.973329 & -0.229416 \end{bmatrix}$$

and

$$R = Q^{\mathsf{T}} \cdot A \approx \begin{bmatrix} 1.41421 & 2.12132 & 0.707107 \\ 0.00000 & 3.08221 & 1.13555 \\ 0.00000 & 0.00000 & 0.458831 \end{bmatrix}.$$

Remark

Because R is upper triangular, everything below its diagonal is zero. Hence, the calculation of R can be sped up by extracting its coefficients from the Gram-Schmidt Process instead of doing the indicated matrix multiplication.

If we use r_{ij} to denote the coefficients in the linear combinations, we end up with

$$
\begin{aligned}
u_1 &= r_{11}\frac{v_1}{\|v_1\|} \\
u_2 &= r_{12}\frac{v_1}{\|v_1\|} + r_{22}\frac{v_2}{\|v_2\|} \\
u_3 &= r_{13}\frac{v_1}{\|v_1\|} + r_{23}\frac{v_2}{\|v_2\|} + r_{33}\frac{v_3}{\|v_3\|}.
\end{aligned}
$$

Writing this out in matrix form then gives $A = Q \cdot R$,

$$\underbrace{\begin{bmatrix} u_1 & u_2 & u_3 \end{bmatrix}}_{A} =$$

$$\underbrace{\begin{bmatrix} r_{11}\frac{v_1}{\|v_1\|} & r_{12}\frac{v_1}{\|v_1\|} + r_{22}\frac{v_2}{\|v_2\|} & r_{13}\frac{v_1}{\|v_1\|} + r_{23}\frac{v_2}{\|v_2\|} + r_{33}\frac{v_3}{\|v_3\|} \end{bmatrix}}_{Q \cdot R}$$

$$= \underbrace{\begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} & \frac{v_3}{\|v_3\|} \end{bmatrix}}_{Q} \cdot \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix}}_{R}$$

For the last step, we break it down into Q multiplying the various columns of R,

$$\underbrace{\begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} & \frac{v_3}{\|v_3\|} \end{bmatrix}}_{Q} \cdot \begin{bmatrix} r_{11} \\ 0 \\ 0 \end{bmatrix} = r_{11}\frac{v_1}{\|v_1\|}$$

$$\underbrace{\begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} & \frac{v_3}{\|v_3\|} \end{bmatrix}}_{Q} \cdot \begin{bmatrix} r_{12} \\ r_{22} \\ 0 \end{bmatrix} = r_{12}\frac{v_1}{\|v_1\|} + r_{22}\frac{v_2}{\|v_2\|}$$

$$\underbrace{\begin{bmatrix} \frac{v_1}{\|v_1\|} & \frac{v_2}{\|v_2\|} & \frac{v_3}{\|v_3\|} \end{bmatrix}}_{Q} \cdot \begin{bmatrix} r_{13} \\ r_{23} \\ r_{33} \end{bmatrix} = r_{13}\frac{v_1}{\|v_1\|} + r_{23}\frac{v_2}{\|v_2\|} + r_{33}\frac{v_3}{\|v_3\|}.$$

Solutions of Linear Equations via the QR Factorization

Suppose that A is $n \times n$ and its columns are linearly independent. Let $A = Q \cdot R$ be its QR Factorization. Then

$$\boxed{Ax = b \iff Q \cdot Rx = b \iff Rx = Q^{\mathsf{T}} b}$$

Solutions of Linear Equations via the QR Factorization

Suppose that A is $n \times n$ and its columns are linearly independent. Let $A = Q \cdot R$ be its QR Factorization. Then

$$\boxed{Ax = b \iff Q \cdot Rx = b \iff Rx = Q^{\mathsf{T}} b}$$

Hence, whenever $\det(A) \neq 0$, for solving $Ax = b$:

- factor $A =: Q \cdot R$,
- compute $\bar{b} := Q^{\mathsf{T}} b$, and then
- solve $Rx = \bar{b}$ via back substitution.

Solve the system of linear equations

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix}}_{b}.$$

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 3 & 1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 0.707 & -0.162 & -0.688 \\ 0.707 & 0.162 & 0.688 \\ 0.000 & 0.973 & -0.229 \end{bmatrix}}_{Q} \cdot \underbrace{\begin{bmatrix} 1.414 & 2.121 & 0.707 \\ 0.000 & 3.082 & 1.135 \\ 0.000 & 0.000 & 0.458 \end{bmatrix}}_{R}.$$

We form

$$\bar{b} := Q^{\mathsf{T}} b = \begin{bmatrix} 3.535 \\ 7.299 \\ 0.458 \end{bmatrix}.$$

Then use back substitution to solve

$$\underbrace{\begin{bmatrix} 1.414 & 2.121 & 0.707 \\ 0.000 & 3.082 & 1.135 \\ 0.000 & 0.000 & 0.458 \end{bmatrix}}_{R} x = \underbrace{\begin{bmatrix} 3.535 \\ 7.299 \\ 0.458 \end{bmatrix}}_{\bar{b}},$$

which yields

$$x = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Suppose that A is $n \times m$ and its columns are linearly independent (tall matrix). Let $A = Q \cdot R$ be its QR Factorization.

Suppose that A is $n \times m$ and its columns are linearly independent (tall matrix). Let $A = Q \cdot R$ be its QR Factorization.

Then $A^\mathsf{T} A = R^\mathsf{T} \cdot Q^\mathsf{T} \cdot Q \cdot R = R^\mathsf{T} \cdot R$ and thus

$$A^\mathsf{T} \cdot Ax = A^\mathsf{T} b \iff R^\mathsf{T} \cdot Rx = R^\mathsf{T} \cdot Q^\mathsf{T} b \iff Rx = Q^\mathsf{T} b,$$

where we have used the fact that R is invertible.

Hence, whenever the columns of A are linearly independent, a least squared error solution to $Ax = b$ is computed as

- factor $A =: Q \cdot R$,
- compute $\bar{b} := Q^{\mathsf{T}} b$, and then
- solve $Rx = \bar{b}$ via back substitution.

- The columns of A will be linearly dependent when $Ax = b$ has fewer equations than unknowns.

Underdetermined Equations

- The columns of A will be linearly dependent when $Ax = b$ has fewer equations than unknowns.

- A is $n \times m$ and $m > n$; more columns than rows (wide matrices).

- The columns of A will be linearly dependent when $Ax = b$ has fewer equations than unknowns.
- A is $n \times m$ and $m > n$; more columns than rows (wide matrices).
- When $Ax = b$ with fewer equations than unknowns, one says that it is *underdetermined*.

- The columns of A will be linearly dependent when $Ax = b$ has fewer equations than unknowns.
- A is $n \times m$ and $m > n$; more columns than rows (wide matrices).
- When $Ax = b$ with fewer equations than unknowns, one says that it is *underdetermined*.
- Because, to determine x uniquely, at a minimum, we need as many equations as unknowns.

Remark

Is there a difference between being underdetermined and having an infinite number of solutions?

Remark

Is there a difference between being underdetermined and having an infinite number of solutions?

Yes. It's possible to be underdetermined and have no solution at all when $b \notin \operatorname{colspan}\{A\}$.

Remark

Is there a difference between being underdetermined and having an infinite number of solutions?

Yes. It's possible to be underdetermined and have no solution at all when $b \notin \mathrm{colspan}\{A\}$.

If the rows of A *are linearly independent, then*

$Ax = b$ *(underdetermined)* $\iff$ $Ax = b$ *(infinite number of solutions).*

Minimum Norm Solution to Underdetermined Equations

Consider an underdetermined system of linear equations $Ax = b$. If the rows of A are linearly independent (equivalently, the columns of A^{T} are linearly independent), then

$$x^* = \arg\min_{Ax=b} \|x\| \iff x^* = A^{\mathsf{T}} \cdot (A \cdot A^{\mathsf{T}})^{-1} b$$

$$\iff x^* = A^{\mathsf{T}}\alpha \;\text{ and }\; A \cdot A^{\mathsf{T}}\alpha = b.$$

Minimum Norm Solution to Underdetermined Equations

- Suppose we do the QR Factorization of A^T instead of A itself, so that
$$A^\mathsf{T} = Q \cdot R.$$

Minimum Norm Solution to Underdetermined Equations

- Suppose we do the QR Factorization of A^T instead of A itself, so that
$$A^\mathsf{T} = Q \cdot R.$$

- Because the columns of A^T are linearly independent, R is square and invertible.

Minimum Norm Solution to Underdetermined Equations

- Suppose we do the QR Factorization of A^{T} instead of A itself, so that
$$A^{\mathsf{T}} = Q \cdot R.$$

- Because the columns of A^{T} are linearly independent, R is square and invertible.

- It follows that $A = R^{\mathsf{T}} \cdot Q^{\mathsf{T}}$ and $A \cdot A^{\mathsf{T}} = R^{\mathsf{T}} \cdot R$ because $Q^{\mathsf{T}} \cdot Q = I$.

Minimum Norm Solution to Underdetermined Equations

Then

$$
\begin{aligned}
x^* = \arg\min_{Ax=b} \|x\| &\iff x^* = Q \cdot \left(R^\mathsf{T}\right)^{-1} b \\
&\iff x^* = Q\beta \quad \text{and} \quad R^\mathsf{T}\beta = b.
\end{aligned}
$$

Minimum Norm Solution to Underdetermined Equations

We note that R^T is lower triangular, and thus $R^\mathsf{T}\beta = b$ can be solved via forward substitution.

Minimum Norm Solution to Underdetermined Equations

We note that R^{T} is lower triangular, and thus $R^{\mathsf{T}}\beta = b$ can be solved via forward substitution.

Hence, for solving underdetermined problems $Ax = b$:

- Check that the columns of A^{T} are linearly independent and compute $A^{\mathsf{T}} = Q \cdot R$.

Minimum Norm Solution to Underdetermined Equations

We note that R^T is lower triangular, and thus $R^\mathsf{T}\beta = b$ can be solved via forward substitution.

Hence, for solving underdetermined problems $Ax = b$:

- Check that the columns of A^T are linearly independent and compute $A^\mathsf{T} = Q \cdot R$.
- Solve $R^\mathsf{T}\beta = b$ by forward substitution.

Minimum Norm Solution to Underdetermined Equations

We note that R^{T} is lower triangular, and thus $R^{\mathsf{T}}\beta = b$ can be solved via forward substitution.

Hence, for solving underdetermined problems $Ax = b$:

- Check that the columns of A^{T} are linearly independent and compute $A^{\mathsf{T}} = Q \cdot R$.
- Solve $R^{\mathsf{T}}\beta = b$ by forward substitution.
- $x^* = Q\beta$.

Find a minimum norm solution to the system of underdetermined equations

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}}_{A} x = \underbrace{\begin{bmatrix} 1 \\ 4 \end{bmatrix}}_{b}.$$

Example

The columns of A^T are linearly independent and we compute the QR Factorization to be

$$A^\mathsf{T} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0.707 & -0.408 \\ 0.707 & 0.408 \\ 0.000 & 0.816 \end{bmatrix}}_{Q} \cdot \underbrace{\begin{bmatrix} 1.414 & 2.121 \\ 0.000 & 1.224 \end{bmatrix}}_{R}.$$

We solve $R^\mathsf{T}\beta = b$ and obtain

$$\beta = \begin{bmatrix} 0.707 \\ 2.041 \end{bmatrix},$$

and then $x^* = Q\beta$ to arrive at the final answer

$$x^* = \begin{bmatrix} -0.333 \\ 1.333 \\ 1.667 \end{bmatrix}.$$

To verify that x^* is indeed a solution, we substitute it into $Ax - b$ and obtain

$$Ax^* - b = \left[\begin{array}{c} 0.000000 \\ -4.441e\text{-}16 \end{array} \right],$$

which looks like a pretty good solution!

Next Time

- The Vector Space $\mathbb{R}^n$: Part 3
- Read Chapter 10 of ROB 101 Book