Descomposición QR por medio del método modificado de Gram y Schmidt (un tema del curso “Álgebra Lineal Numérica”)

Egor Maximenko

Instituto Politécnico Nacional
Escuela Superior de Física y Matemáticas
México

7 de mayo de 2021

Problema: descomposición QR delgada

Está dada una matriz $A \in \mathbb{R}^{n \times m}$ con $m \leq n$ y $\mathrm{r}(A) = m$.

Encontrar $Q \in \mathbb{R}^{n \times m}$ y $R \in \mathbb{R}^{m \times m}$ tales que:

- las columnas de Q formen una lista ortonormal, esto es, $Q^\top Q = I_m$,
- la matriz R sea triangular superior,
- $A = QR$.

Problema: descomposición QR delgada

Está dada una matriz $A \in \mathbb{R}^{n\times m}$ con $m \le n$ y $\mathrm{r}(A) = m$.

Encontrar $Q \in \mathbb{R}^{n\times m}$ y $R \in \mathbb{R}^{m\times m}$ tales que:

- las columnas de Q formen una lista ortonormal, esto es, $Q^\top Q = I_m$,
- la matriz R sea triangular superior,
- $A = QR$.

$$\underbrace{\begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \\ A_{4,1} & A_{4,2} & A_{4,3} \\ A_{5,1} & A_{5,2} & A_{5,3} \end{bmatrix}}_{\text{columnas lin. indep.}} = \underbrace{\begin{bmatrix} Q_{1,1} & Q_{1,2} & Q_{1,3} \\ Q_{2,1} & Q_{2,2} & Q_{2,3} \\ Q_{3,1} & Q_{3,2} & Q_{3,3} \\ Q_{4,1} & Q_{4,2} & Q_{4,3} \\ Q_{5,1} & Q_{5,2} & Q_{5,3} \end{bmatrix}}_{\text{columnas ortonormales}} \begin{bmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ 0 & R_{2,2} & R_{2,3} \\ 0 & 0 & R_{3,3} \end{bmatrix}.$$

Hay varios métodos para resolver el problema de descomposición QR:

- método de ortogonalización de Gram y Schmidt,
- método modificado de ortogonalización de Gram y Schmidt,
- reflexiones ortogonales (reflexiones de Householder),
- rotaciones ortogonales (rotaciones de Givens).

Hay varios métodos para resolver el problema de descomposición QR:

- método de ortogonalización de Gram y Schmidt,
- método modificado de ortogonalización de Gram y Schmidt,
- reflexiones ortogonales (reflexiones de Householder),
- rotaciones ortogonales (rotaciones de Givens).

En este tema veremos el segundo método.

Prerrequisitos para entender bien este tema.

- La matriz de proyección ortogonal al subespacio generado por un vector.
- La matriz del complemento ortogonal respecto al subespacio generado por un vector.
- Listas de vectores linealmente independientes y sus propiedades.
- Subespacios generados por listas de vectores y sus propiedades.
- Matrices ortogonales y matrices triangulares.
- El producto Ax es una combinación lineal de las columnas de A.
- El renglón de los productos internos $[\langle a_1, q\rangle, \ldots, \langle a_m, q\rangle]$ es $q^\top A$.

El producto de una matriz por un vector es una combinación lineal de las columnas de la matriz (repaso)

Si $A \in \mathbb{R}^{n\times m}$, $x \in \mathbb{R}^m$, entonces

$$Ax =$$

El producto de una matriz por un vector es una combinación lineal de las columnas de la matriz (repaso)

Si $A \in \mathbb{R}^{n \times m}$, $x \in \mathbb{R}^m$, entonces

$$Ax = \sum_{k=1}^{m} A_{*,k} x_k.$$

$$\begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \\ A_{4,1} & A_{4,2} & A_{4,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} =$$

El producto de una matriz por un vector es una combinación lineal de las columnas de la matriz (repaso)

Si $A \in \mathbb{R}^{n\times m}$, $x \in \mathbb{R}^m$, entonces

$$Ax = \sum_{k=1}^{m} A_{*,k} x_k.$$

$$\begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{3,1} & A_{3,2} & A_{3,3} \\ A_{4,1} & A_{4,2} & A_{4,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1 \begin{bmatrix} A_{1,1} \\ A_{2,1} \\ A_{3,1} \\ A_{4,1} \end{bmatrix} + x_2 \begin{bmatrix} A_{1,2} \\ A_{2,2} \\ A_{3,2} \\ A_{4,2} \end{bmatrix} + x_3 \begin{bmatrix} A_{1,3} \\ A_{2,3} \\ A_{3,3} \\ A_{4,3} \end{bmatrix}.$$

El renglón de los productos internos de una lista de vectores por un vector (repaso)

Sea $q \in \mathbb{R}^n$ y sea $V = [v_1, \ldots, v_m] \in \mathbb{R}^{n \times m}$.

Entonces

$$[\langle v_1, q\rangle,\ \ldots,\ \langle v_m, q\rangle] = [q^\top v_1,\ \ldots,\ q^\top v_m] = q^\top V.$$

Ejemplo ($m = 2,\ n = 3$):

$$[q_1, q_2, q_3] \begin{bmatrix} V_{1,1} & V_{1,2} \\ V_{2,1} & V_{2,2} \\ V_{3,1} & V_{3,2} \end{bmatrix} = \Big[V_{1,1}q_1 + V_{2,1}q_2 + V_{3,1}q_3,\ \ V_{1,2}q_1 + V_{2,2}q_2 + V_{3,2}q_3\Big].$$

Proyección ortogonal de un vector al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$ *y sea* $v \in \mathbb{R}^n$.

Denotemos por S *el subespacio generado por* q: $\quad S := \ell(q)$.

Pongamos

$$u := \langle v, q \rangle \, q, \qquad w := v - u.$$

Entonces

$$u \in S, \quad w \in S^{\perp}, \quad v = u + w.$$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$*. Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 =$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 = P_q$,
- $P_q^\top =$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$*. Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 = P_q$,
- $P_q^\top = P_q$,
- $P_q q =$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 = P_q$,
- $P_q^\top = P_q$,
- $P_q q = q$,
- *para cada* x *en* S, $\quad P_q x =$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 = P_q$,
- $P_q^\top = P_q$,
- $P_q q = q$,
- *para cada* x *en* S, $\quad P_q x = x$,
- *para cada* y *en* $S^\perp$, $\quad P_q y =$

La matriz de la proyección ortogonal al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(q), \qquad \boxed{P_q := qq^\top}.$$

Entonces

- $P_q^2 = P_q$,
- $P_q^\top = P_q$,
- $P_q q = q$,
- *para cada* x *en* S, $\quad P_q x = x$,
- *para cada* y *en* $S^\perp$, $\quad P_q y = 0_n$.

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 =$

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 = I_n - P_q$,
- $(I_n - P_q)^\top =$

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 = I_n - P_q$,
- $(I_n - P_q)^\top = I_n - P_q$,
- $(I_n - P_q)q =$

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 = I_n - P_q$,
- $(I_n - P_q)^\top = I_n - P_q$,
- $(I_n - P_q)q = 0_n$,
- *para cada* x *en* S, $\quad (I_n - P_q)x =$

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 = I_n - P_q$,
- $(I_n - P_q)^\top = I_n - P_q$,
- $(I_n - P_q)q = 0_n$,
- *para cada* x *en* S, $\quad (I_n - P_q)x = 0_n$,
- *para cada* y *en* $S^\perp$, $\quad (I_n - P_q)y =$

La matriz del complemento ortogonal respecto al subespacio generado por un vector normalizado (repaso)

Proposición

Sea $q \in \mathbb{R}^n$ *tal que* $\|q\| = 1$. *Pongamos*

$$S := \ell(a), \qquad P_a := aa^\top.$$

Entonces la matriz $I_n - P_q$ *tiene las siguientes propiedades:*

- $(I_n - P_q)^2 = I_n - P_q$,
- $(I_n - P_q)^\top = I_n - P_q$,
- $(I_n - P_q)q = 0_n$,
- *para cada* x *en* S, $\quad (I_n - P_q)x = 0_n$,
- *para cada* y *en* $S^\perp$, $\quad (I_n - P_q)y = y$.

La acción del complemento ortogonal a una lista de vectores (repaso)

Sea $q \in \mathbb{R}^n$ tal que $\|q\| = 1$ y sea

$$V = [v_1, \ldots, v_m] \in \mathbb{R}^{n \times m}.$$

Consideremos la siguiente matriz:

$$W := \left[v_1 - P_q v_1, \ldots, v_m - P_q v_m\right] = (I_n - P_q)V = V - (qq^\top)V = V - q(q^\top V).$$

Pongamos

$$\boxed{\lambda := q^\top V} =$$

La acción del complemento ortogonal a una lista de vectores (repaso)

Sea $q \in \mathbb{R}^n$ tal que $\|q\| = 1$ y sea

$$V = [v_1, \ldots, v_m] \in \mathbb{R}^{n \times m}.$$

Consideremos la siguiente matriz:

$$W \coloneqq \left[v_1 - P_q v_1, \ldots, v_m - P_q v_m\right] = (I_n - P_q)V = V - (qq^\top)V = V - q(q^\top V).$$

Pongamos

$$\lambda \coloneqq q^\top V = \left[\langle v_1, q\rangle, \; \ldots, \; \langle v_m, q\rangle\right].$$

Entonces

La acción del complemento ortogonal a una lista de vectores (repaso)

Sea $q \in \mathbb{R}^n$ tal que $\|q\| = 1$ y sea

$$V = [v_1, \ldots, v_m] \in \mathbb{R}^{n \times m}.$$

Consideremos la siguiente matriz:

$$W := \left[v_1 - P_q v_1, \ldots, v_m - P_q v_m\right] = (I_n - P_q)V = V - (qq^\top)V = V - q(q^\top V).$$

Pongamos

$$\lambda := q^\top V = \left[\langle v_1, q\rangle,\ \ldots,\ \langle v_m, q\rangle\right].$$

Entonces

$$W = V - q\,\lambda.$$

La acción del complemento ortogonal a una lista de vectores (repaso)

$q \in \mathbb{R}^n$, $V \in \mathbb{R}^{n \times m}$,

$$\lambda = q^\top V, \qquad W = V - q\,\lambda.$$

Programación en Octave:

```
la = q' * V;
W = V - q * la;
```

Programación en NumPy, una de varias maneras:

```
la = atleast_2d(q) @ V
W = V - outer(q, la)
```

El método clásico de Gram–Schmidt ($m = 4$)

Están dados vectores linealmente independientes $a_1, a_2, a_3, a_4 \in \mathbb{R}^n$.

Paso 1:

$$w_1 := a_1,$$

$$\nu_1 := \|w_1\|, \qquad q_1 := \frac{1}{\nu_1}\, w_1.$$

El método clásico de Gram–Schmidt ($m = 4$)

Paso 2:

$$\lambda_{1,2} := q_1^\top a_2,$$

$$w_2 := a_2 - \lambda_{1,2} q_1.$$

En otras palabras, $w_2 := a_2 - P_{q_1} a_2$.

$$\nu_2 := \|w_2\|, \qquad q_2 := \frac{1}{\nu_2} w_2.$$

El método clásico de Gram–Schmidt ($m = 4$)

Paso 3:

$$\lambda_{1,3} := q_1^\top a_3, \qquad \lambda_{2,3} := q_2^\top a_3,$$

$$w_3 := a_3 - \lambda_{1,3} q_1 - \lambda_{2,3} q_2.$$

En otras palabras, $w_3 := a_3 - P_{q_1,q_2} a_3$.

$$\nu_3 := \|w_3\|, \qquad q_3 := \frac{1}{\nu_3} w_3.$$

El método clásico de Gram–Schmidt ($m = 4$)

Paso 4:

$$\lambda_{1,4} := q_1^\top a_4, \qquad \lambda_{2,4} := q_2^\top a_4, \qquad \lambda_{3,4} := q_3^\top a_4,$$

$$w_4 := a_4 - \lambda_{1,4} q_1 - \lambda_{2,4} q_2 - \lambda_{3,4} q_3.$$

En otras palabras, $w_4 := a_4 - P_{q_1, q_2, q_3} a_4$.

$$\nu_4 := \|w_4\|, \qquad q_4 := \frac{1}{\nu_4} w_4.$$

El método clásico de Gram–Schmidt, paso general

En el k-ésimo paso, se calcula

$$w_k := a_k - P_{q_1,\ldots,q_{k-1}}\, a_k.$$

De manera detallada,

$$\lambda_{1,k} := q_1^\top a_k, \qquad \lambda_{2,k} := q_2^\top a_k, \qquad \ldots, \qquad \lambda_{k-1,k} := q_{k-1}^\top a_k,$$

$$w_k := a_k - \sum_{j=1}^{k-1} \lambda_{j,k} q_j.$$

Luego se normaliza el vector w_k:

$$\nu_k := \|w_k\|, \qquad q_k := \frac{1}{\nu_k}\, w_k.$$

El método modificado de Gram–Schmidt ($m = 4$)

Están dados vectores linealmente independientes $a_1, a_2, a_3, a_4 \in \mathbb{R}^n$.

Los denotamos como $a_1^{(0)}, a_2^{(0)}, a_3^{(0)}, a_4^{(0)}$.

Paso 1. Normalizamos $a_1^{(0)}$:

$$\nu_1 := \|a_1^{(0)}\|, \qquad q_1 := \frac{1}{\nu_1}\, a_1^{(0)}.$$

Luego restamos de los vectores siguientes sus proyecciones sobre q_1:

$$\lambda_{1,2} := q_1^\top a_2^{(0)}, \quad \lambda_{1,3} := q_1^\top a_3^{(0)}, \quad \lambda_{1,4} := q_1^\top a_4^{(0)},$$

$$a_2^{(1)} := a_2^{(0)} - \lambda_{1,2} q_1, \quad a_3^{(1)} := a_3^{(0)} - \lambda_{1,3} q_1, \quad a_4^{(1)} := a_4^{(0)} - \lambda_{1,4} q_1.$$

El método modificado de Gram–Schmidt ($m = 4$)

Paso 2. Normalizamos $a_2^{(1)}$:

$$\nu_2 := \|a_2^{(1)}\|, \qquad q_2 := \frac{1}{\nu_2}\, a_2^{(1)}.$$

Luego restamos de los vectores siguientes sus proyecciones sobre q_2:

$$\lambda_{2,3} := q_2^\top a_3^{(1)}, \quad \lambda_{2,4} := q_2^\top a_4^{(1)},$$

$$a_3^{(2)} := a_3^{(1)} - \lambda_{2,3} q_2, \quad a_4^{(2)} := a_4^{(1)} - \lambda_{2,4} q_2.$$

El método modificado de Gram–Schmidt ($m = 4$)

Paso 3. Normalizamos $a_3^{(2)}$:

$$\nu_3 := \|a_3^{(2)}\|, \qquad q_3 := \frac{1}{\nu_3}\, a_3^{(2)}.$$

Luego restamos de los vectores siguientes sus proyecciones sobre q_3:

$$\lambda_{3,4} := q_3^\top a_4^{(2)},$$

$$a_4^{(3)} := a_4^{(2)} - \lambda_{3,4} q_3.$$

El método modificado de Gram–Schmidt ($m = 4$)

Paso 4. Normalizamos $a_4^{(3)}$:

$$\nu_4 := \|a_4^{(3)}\|, \qquad q_4 := \frac{1}{\nu_4}\, a_4^{(3)}.$$

El método modificado de Gram–Schmidt, paso general

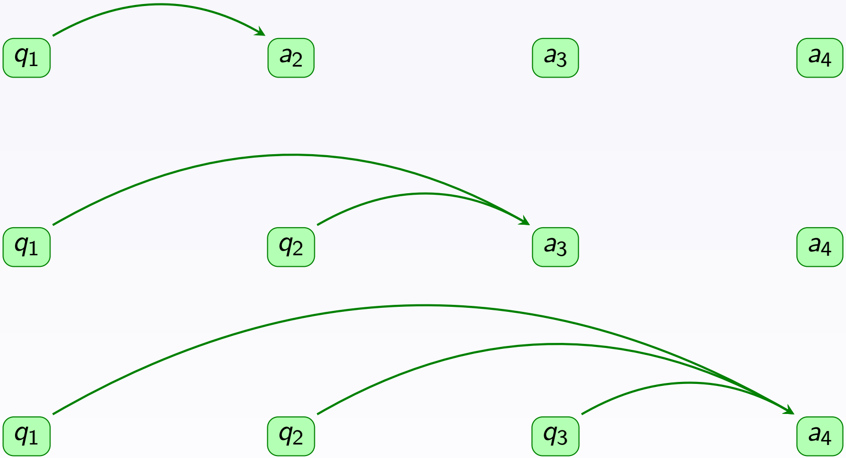
Para m general y j general $(1 \le j \le m)$, en el paso j hacemos las siguientes operaciones.

$$\nu_j := \|a_j^{(j-1)}\|, \qquad q_j := \frac{1}{\nu_j}\, a_j^{(j-1)}.$$

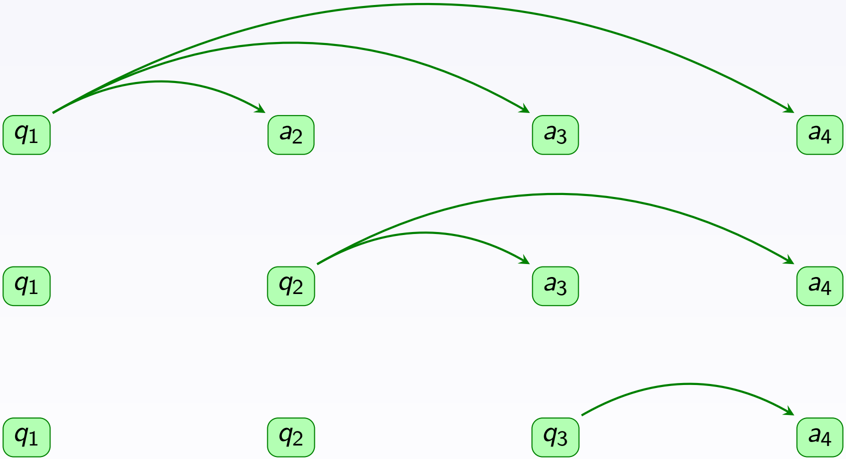
$$\lambda_{j,k} := q_j^\top a_k^{(j-1)} \qquad (j+1 \le k \le m),$$

$$a_k^{(j)} := a_k^{(j-1)} - \lambda_{j,k} q_j \qquad (j+1 \le k \le m).$$

El método clásico de Gram–Schmidt ($m = 4$)



El método modificado de Gram–Schmidt ($m = 4$)



Ejercicio

Demostrar que los coeficientes $\lambda_{j,k}$ en el método modificado de Gram y Schmidt son los mismos que en el método clásico.

Idea:

$$a_4^{(2)} = a_4^{(1)} - \lambda_{2,4} q_2 = a_4 - \lambda_{1,4} q_1 - \lambda_{2,4} q_2,$$

$$q_3^\top a_4^{(2)} = q_3^\top a_4 - \lambda_{1,4} \underbrace{q_3^\top q_1}_{0} - \lambda_{2,4} \underbrace{q_3^\top q_2}_{0} = q_3^\top a_4.$$

Como consequencia, los vectores $q_1, \ldots, q_m$ en el método modificado de Gram–Schmidt son los mismos que en el método clásico.

Ejercicios de programación

Ejercicio. Escribir una función que aplique el método modificado de Gram–Schmidt a una matriz A de tamaño $n \times 4$ y devuelva las matrices Q y R.

Ejercicio. Escribir una función que aplique el método modificado de Gram–Schmidt a una matriz A de tamaño $n \times m$ y devuelva las matrices Q y R. Usar dos ciclos anidados y las operaciones con columnas.

Ejercicio. Escribir una función que aplique el método modificado de Gram–Schmidt a una matriz A de tamaño $n \times m$ y devuelva las matrices Q y R.
Usar un ciclo. Calcular $V - q(q^\top V)$ con operaciones matriciales.

Ejercicio. Hacer pruebas. Comparar los errores con el método clásico de Gram–Schmidt.