Cálculo Numérico - Elementos de Cálculo Numérico

Descomposición QR - Transformaciones de Householder

Mercedes Pérez Millán

Dto. de Matemática–FCEyN–Universidad de Buenos Aires
IMAS-CONICET

3 de mayo de 2021

Resumen

Recuerdo transformaciones ortogonales

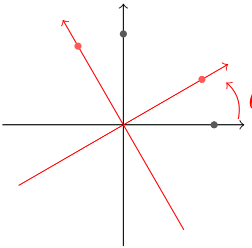
Descomposición QR vía transformaciones de Householder

Ejemplo

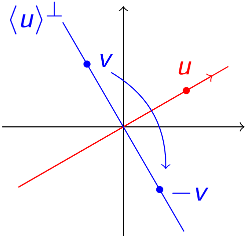
Matrices ortogonales en $\mathbb{R}^{n\times n}$:

- $Q \in \mathbb{R}^{n\times n}$ es ortogonal
  $\Leftrightarrow$ las columnas de Q son una base ortonormal (b.o.n.) de $\mathbb{R}^n$
  $\Leftrightarrow$ Q es inversible y $Q^{-1} = Q^t$.
- $\langle Qv, Qw\rangle = \langle v, w\rangle$. Y por lo tanto preserva ángulos y distancias.
- Si λ es autovalor de Q, entonces $\lambda = \pm 1$.

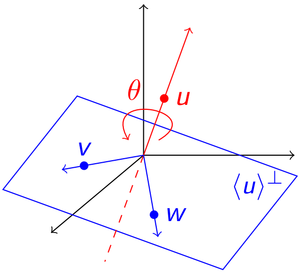
Rotación en $\mathbb{R}^2$:



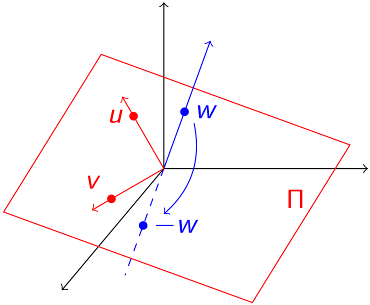
Simetría en $\mathbb{R}^2$:



Rotación en $\mathbb{R}^3$:



Simetría en $\mathbb{R}^3$:



En $\mathbb{R}^n$:

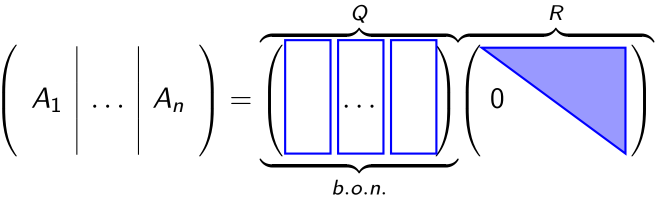
Si $Q \in \mathbb{R}^{n\times n}$ es ortogonal, existe $\{v_1, \ldots, v_n\}$ b.o.n. de $\mathbb{R}^n$ tal que si $C = (v_1|\ldots|v_n)$:

$$C^{-1}QC = \begin{pmatrix} 1 &&&&&&&& \\ & \ddots &&&&&&& \\ && 1 &&&&&& \\ &&& -1 &&&&& \\ &&&& \ddots &&&& \\ &&&&& -1 &&& \\ &&&&&& A_1 && \\ &&&&&&& \ddots & \\ &&&&&&&& A_k \end{pmatrix}$$

con $A_i = \begin{pmatrix} \cos(\theta_i) & -\operatorname{sen}(\theta_i) \\ \operatorname{sen}(\theta_i) & \cos(\theta_i) \end{pmatrix}$

Descomposición QR:

- Conviene mirar Clase 10 del libro: L.N. Trefethen, D. Bau. Numerical Linear Algebra, SIAM 1997.
- Queremos descomponer a la matriz $A \in \mathbb{R}^{n\times n}$ como $A = QR$, con Q y R tales que:

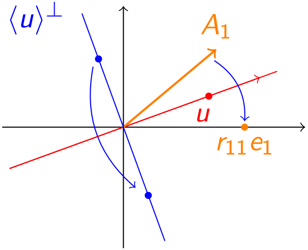


- Si $A = (A_1|\dots|A_n)$, multipliquemos por matrices "convenientes" a izquierda para llevarla a una matriz triangular superior R.

Transformaciones de Householder

$$A = (A_1| \dots |A_n)$$

Para llevar A_1 a un múltiplo de e_1, el primer canónico, una opción (para poner ceros debajo de la diagonal) es aplicar una simetría respecto del (hiper)plano que pasa "por el medio" de A_1 y e_1:



$$Q_1 A = \left(\begin{array}{c|c} r_{11} & * \\ \hline 0 & A' \end{array}\right)$$

En realidad:

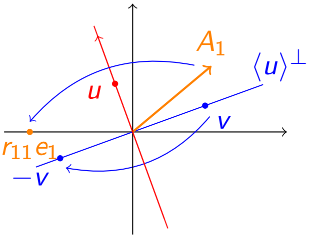
$$Q_1(A_1|\dots|A_n) = \left(\begin{array}{c|c} r_{11} & * \\ \hline 0 & A' \end{array}\right)$$

- Como $\|Q_1 A_1\| = \|A_1\| = |r_{11}|$, entonces $r_{11} = \pm\|A_1\|$.
- Para evitar restar dos números parecidos, conviene que $\mathrm{signo}(r_{11}) = -\mathrm{signo}(a_{11})$, i.e., $r_{11} = -\mathrm{signo}(a_{11})\|A_1\|$.

Si $\tilde{v} = A_1 - r_{11}e_1$, y $v = \frac{\tilde{v}}{\|\tilde{v}\|}$, entonces:

$$Q_1 = Id_n - 2vv^t$$

$$= Id_n - 2\left(\boxed{v}\right)\left(\boxed{\quad v \quad}\right)$$



Obs: $Q_1^t = (Id_n)^t - 2((v^t)^t v^t) = Q_1$

Cuenta para leer detenidamente en casa:

$$
\begin{aligned}
\color{red}{Q_1 A_1} &= A_1 - 2vv^t A_1 = A_1 - 2v\langle v, A_1\rangle \\
&= A_1 - \frac{2\langle A_1 - r_{11}e_1, A_1\rangle}{\|A_1 - r_{11}e_1\|} v \\
&= A_1 - \frac{2(\|A_1\|^2 - r_{11}a_{11})}{\|A_1 - r_{11}e_1\|} \frac{A_1 - r_{11}e_1}{\|A_1 - r_{11}e_1\|} \\
&= A_1 - \frac{2(\|A_1\|^2 - r_{11}a_{11})}{\langle A_1 - r_{11}e_1, A_1 - r_{11}e_1\rangle}(A_1 - r_{11}e_1) \\
&= A_1 - \frac{2(\|A_1\|^2 - r_{11}a_{11})}{\|A_1\|^2 - 2r_{11}a_{11} + \underbrace{\|r_{11}e_1\|^2}_{\|A_1\|^2}}(A_1 - r_{11}e_1) \\
&= A_1 - (A_1 - r_{11}e_1) \color{red}{= r_{11}e_1}. \color{green}{\checkmark}
\end{aligned}
$$

Cómo seguimos:

Por medio de los reflectores de Householder:

- $$Q_1(A_1|\dots|A_n) = \left(\begin{array}{c|c} r_{11} & * \\ \hline 0 & A' \end{array}\right).$$
- $Q_2' = Id_{n-1} - 2v'(v')^t$, con $v' = \frac{A_1' - r_{22}e_1'}{\|A_1' - r_{22}e_1'\|}$ y $Q_2 = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & Q_2' \end{array}\right)$.
- $Q_k' = Id_{n-k+1} - 2v'(v')^t$ y

$$Q_k = \left(\begin{array}{c|c} Id_{k-1} & 0 \\ \hline 0 & Q_k' \end{array}\right).$$

(v' el que corresponda en $\mathbb{R}^{n-k+1}$)

- Finalmente:
  $Q = (Q_{n-1}\dots Q_1)^{-1} = Q_1^t \dots Q_{n-1}^t = Q_1 \dots Q_{n-1}$.

Algoritmo (para leer con paciencia en casa)

Para calcular R escribiendo los resultados sobre A (notar que guarda los vectores v):

para $k = 1$ **hasta** n

$$
\begin{aligned}
x &= A_{k:m,k} \\
v_k &= x + \mathrm{signo}(x_1)\|x\|e_1 \\
v_k &= v_k/\|v_k\| \\
A_{k:m,k} &= A_{k:m,k:n} - 2v_k(v_k^t A_{k:m,k:n})
\end{aligned}
$$

Ejemplo:

Calcular la descomposición QR de la matriz A:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

- Utilizando transformaciones de Householder.
- Aplicando de manera directa el método de ortonormalización de Gramm-Schmidt.

Con Householder:

- $a_{11} = 0$ y $\|A_1\| = 1 \;\Rightarrow\; \tilde{v}_1 = A_1 + e_1$, y $v_1 = \frac{\tilde{v}_1}{\|\tilde{v}_1\|}$, entonces:

$$
\begin{aligned}
Q_1 = Id_3 - 2v_1 v_1^t &= Id_3 - \frac{2}{\|\tilde{v}_1\|^2}\tilde{v}_1\tilde{v}_1^t \\
&= Id_3 - \frac{2}{2}\begin{pmatrix}1\\0\\1\end{pmatrix}(1,0,1) \\
&= Id_3 - \begin{pmatrix}1 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0 & -1\\ 0 & 1 & 0\\ -1 & 0 & 0\end{pmatrix}.
\end{aligned}
$$

$$
Q_1 A = \left(\begin{array}{c|cc} -1 & 0 & -1\\ \hline 0 & 1 & 0\\ 0 & -1 & -1\end{array}\right) \Rightarrow A' = \begin{pmatrix}1 & 0\\ -1 & -1\end{pmatrix}.
$$

- $a'_{11} = 1 > 0$ y $\|A'_1\| = \sqrt{2} \;\Rightarrow\; \tilde{v}_2 = A'_1 + \sqrt{2}e_1$, y $v_2 = \frac{\tilde{v}_2}{\|\tilde{v}_2\|}$, entonces:

$$
\begin{aligned}
Q'_2 = Id_2 - 2v_2v_2^t &= Id_2 - \frac{2}{\|\tilde{v}_2\|^2}\tilde{v}_2\tilde{v}_2^t \\
&= Id_2 - \frac{2}{4+2\sqrt{2}}\begin{pmatrix} 1+\sqrt{2} \\ -1 \end{pmatrix}(1+\sqrt{2}, -1) \\
&= Id_2 - \frac{1}{2+\sqrt{2}}\begin{pmatrix} 3+2\sqrt{2} & -1-\sqrt{2} \\ -1-\sqrt{2} & 1 \end{pmatrix} \\
&= \begin{pmatrix} -\frac{1+\sqrt{2}}{2+\sqrt{2}} & \frac{1+\sqrt{2}}{2+\sqrt{2}} \\ \frac{1+\sqrt{2}}{2+\sqrt{2}} & \frac{1+\sqrt{2}}{2+\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}
\end{aligned}
$$

$$
Q_2 = \left(\begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{array}\right)
$$

Así, $R = Q_2 Q_1 A = \begin{pmatrix} -1 & 0 & -1 \\ 0 & -\sqrt{2} & -\frac{1}{\sqrt{2}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} \end{pmatrix}$.

¿Y quién es Q?

$$Q = (Q_2 Q_1)^{-1} = Q_1^{-1} Q_2^{-1} = Q_1^t Q_2^t = Q_1 Q_2$$

$$= \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -1 & 0 & 0 \end{pmatrix}$$

Con Gram-Schmidt:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \left(\begin{array}{c|c|c} A_1 & A_2 & A_3 \end{array} \right)$$

- $\tilde{q}_1 = A_1, \|\tilde{q}_1\| = 1 \;\Rightarrow\; q_1 = (0, 0, 1)^t,$
- $\tilde{q}_2 = A_2 - \langle A_2, q_1 \rangle q_1 = (1, 1, 0)^t - 0 = (1, 1, 0)^t \;\Rightarrow$ $q_2 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)^t,$
- $\tilde{q}_3 = A_3 - \langle A_3, q_1 \rangle q_1 - \langle A_3, q_2 \rangle q_2 = (1, 0, 1)^t - (0, 0, 1)^t - (\frac{1}{2}, \frac{1}{2}, 0)^t = (\frac{1}{2}, -\frac{1}{2}, 0)^t \;\Rightarrow\; q_3 = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0)^t,$

$$\Rightarrow \; Q = \left(\begin{array}{c|c|c} q_1 & q_2 & q_3 \end{array} \right) = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 1 & 0 & 0 \end{pmatrix}$$

¿Y quién es R?

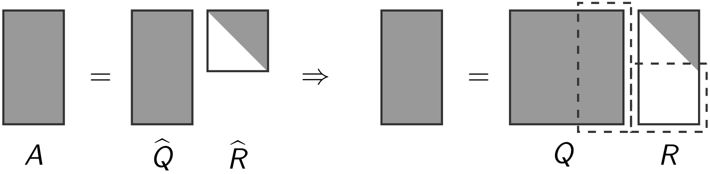
$$R = Q^{-1}A = Q^t A = \begin{pmatrix} q_1 \\ \hline q_2 \\ \hline q_3 \end{pmatrix} \left(\begin{array}{c|c|c} A_1 & A_2 & A_3 \end{array} \right)$$

$$= \begin{pmatrix} \langle q_1, A_1\rangle & \langle q_1, A_2\rangle & \langle q_1, A_3\rangle \\ 0 & \langle q_2, A_2\rangle & \langle q_2, A_3\rangle \\ 0 & 0 & \langle q_3, A_3\rangle \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 1 \\ 0 & \sqrt{2} & \frac{1}{\sqrt{2}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$$

¿Y si $A \in \mathbb{R}^{m\times n}$ no es cuadrada?

- Si $m \geq n$, completo $\widehat{Q}$ a Q para que sus columnas sean una b.o.n., y agrego filas de ceros a $\widehat{R}$:



- Si $m \leq n$, recordar que $r_{ij} = \langle q_i, A_j \rangle$.