Ejercicios proyección ortogonal, cuadrados mínimos y descomposición QR

1. Sea $A \in \mathbb{R}^{3\times 3}$ una matriz de rango 2 que verifica $(A^TA)v = 0$ para $v = (1\ 2\ -1)^T$

   a) Calcular la proyección ortogonal de $\mathbb{R}^3$ sobre $Nul(A)$.

   b) Calcular la distancia del vector $(3\ 1\ 1)^T$ a $Fil(A)$.

   c) Si se sabe que $\left\|b - A(111)^T\right\| \leq \|b - Ax\|$ para todo $x \in \mathbb{R}^3$, hallar todas las soluciones por cuadrados mínimos de $Ax = b$. ¿Cuál de ellas tiene norma mínima?

a) El dato que nos da el ejercicio es que $v = (1\ 2\ -1)^T \in Nul(A^TA)$.

Vamos a probar que

$$Nul(A) = Nul(A^TA)$$

- $Nul(A) \subseteq Nul(A^TA)$ ya que:

$$v \in Nul(A) \quad \Rightarrow \quad Av = 0 \quad \Rightarrow \quad A^TAv = A^T0 = 0 \quad \Rightarrow \quad v \in Nul(A^TA)$$

- $Nul(A^TA) \subseteq Nul(A)$ ya que:

$$v \in Nul(A^TA) \quad \Rightarrow \quad A^TAv = 0 \quad \Rightarrow \quad v^TA^TAv = 0 \quad \Rightarrow \quad (Av)^TAv = 0$$

$$\Rightarrow \quad \|Av\|^2 = 0 \quad \Rightarrow \quad Av = 0 \quad \Rightarrow \quad v \in Nul(A)$$

Entonces $v = (1\ 2\ -1)^T \in Nul(A)$.

Por otro lado, podemos conocer la dimensión de $Nul(A)$ ya que el rango de A es 2.

$$3 = rango(A) + dim(Nul(A)) = 2 + dim(Nul(A)) \quad \Leftrightarrow \quad dim(Nul(A)) = 1$$

Concluimos que

$$Nul(A) = gen\{(1\ 2\ -1)^T\}$$

Entonces la proyección sobre $Nul(A)$ es

$$P_{Nul(A)}(x_1\ x_2\ x_3) = \frac{\left\langle (x_1\ x_2\ x_3)^T, (1\ 2\ -1)^T \right\rangle}{\|(1\ 2\ -1)^T\|^2}(1\ 2\ -1)^T = \frac{x_1 + 2x_2 - x_3}{6}(1\ 2\ -1)^T$$

b) Calculemos la distancia del vector $(3\ 1\ 1)^T$ a $Fil(A)$. Esto es,

$$d((3\ 1\ 1)^T, Fil(A)) = \left\|(3\ 1\ 1)^T - P_{Fil(A)}((3\ 1\ 1)^T)\right\| = \left\|P_{(Fil(A))^\perp}(3\ 1\ 1)^T\right\|$$

Recordemos que

$$Nul(A) = (Fil(A))^\perp$$

Esto vale ya que:

- $dim(Nul(A)) = dim((Fil(A))^\perp)$ pues si $A \in \mathbb{R}^{m\times n}$

$$dim(Nul(A)) = n - rango(A) = n - dim(Fil(A)) = dim((Fil(A))^\perp)$$

- $Nul(A) \subseteq (Fil(A))^{\perp}$ ya que si $A = \begin{pmatrix} F_1^T \\ F_2^T \\ \vdots \\ F_m^T \end{pmatrix}$ tenemos

$$v \in Nul(A) \;\Rightarrow\; Av = 0 \;\Rightarrow\; \begin{pmatrix} F_1^T \\ F_2^T \\ \vdots \\ F_m^T \end{pmatrix} v = \begin{pmatrix} F_1^T v \\ F_2^T v \\ \vdots \\ F_m^T v \end{pmatrix} = \begin{pmatrix} \langle F_1, v\rangle \\ \langle F_2, v\rangle \\ \vdots \\ \langle F_m, v\rangle \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow \quad \langle F_1, v\rangle = 0, \langle F_2, v\rangle = 0, \cdots, \langle F_m, v\rangle = 0 \quad \Rightarrow \quad v \in (Fil(A))^{\perp}$$

Entonces

$$d((3\ 1\ 1)^T, Fil(A)) = \left\| P_{Nul(A)}(3\ 1\ 1)^T \right\| = \left\| \frac{4}{6}(1\ 2\ -1)^T \right\| = \frac{2}{3}\sqrt{6}$$

c) El dato que $\left\| b - A(1\ 1\ 1)^T \right\| \leq \| b - Ax \|$ nos dice que $(1\ 1\ 1)^T$ es una solución por cuadrados mínimos de $Ax = b$.

Recordemos que las soluciones por cuadrados mínimos $\hat{x}$ del sistema $Ax = b$ verifican:

- $\hat{x}$ es solución de $A^T Ax = A^T b$ (ecuaciones normales).
- $\hat{x}$ es solución de $Ax = P_{Col(A)}(b)$.

Estas soluciones son de la forma

$$\hat{x} = x_p + x_h$$

donde x_p es una solución particular y x_h es solución del sistema homogéneo asociado (es decir, $x_h \in Nul(A)$).

Tenemos que $x_p = (1\ 1\ 1)^T$ y $x_h = \alpha(1\ 2\ -1)^T$, con $\alpha \in \mathbb{R}$. Entonces todas las soluciones por cuadrados mínimos de la ecuación $Ax = b$ son

$$\hat{x} = (1\ 1\ 1)^T + \alpha(1\ 2\ -1)^T$$

con $\alpha \in \mathbb{R}$.

Veamos cual de estas soluciones tiene norma mínima. Para ello vamos a probar que existe una única solución por cuadrados mínimos del sistema $Ax = b$ que pertenece a $Fil(A)$ y que esta solución tiene norma mínima.

Sea $A \in \mathbb{R}^{m \times n}$. Vimos que $Nul(A) = (Fil(A))^{\perp}$. Entonces, cualquier solución x_p del sistema $Ax = b$ por cuadrados mínimos, puede escribirse como

$$x_p = P_{Fil(A)}(x_p) + P_{(Fil(A))^{\perp}}(x_p) = P_{Fil(A)}(x_p) + P_{Nul(A)}(x_p)$$

Llamamos $x_f = P_{Fil(A)}(x_p)$. Tenemos que:

- $x_f \in Fil(A)$.
- Si x_p' es otra solución $x_p' = x_p + x_p' - x_p$, con $x_p' - x_p \in Nul(A)$. Entonces

$$P_{Fil(A)}(x_p') = P_{Fil(A)}(x_p) + P_{Fil(A)}(x_p' - x_p) = x_f$$

- x_f es solución por cuadrados mínimos de $Ax = b$.
  Sabemos que x_p es solución, así que $Ax_p = P_{Col(A)}(b)$. Entonces

  $$P_{Col(A)}(b) = Ax_p = A(x_f + P_{Nul(A)}(x_p)) = Ax_f + A(P_{Nul(A)}(x_p)) = Ax_f$$

  Luego, x_f es solución por cuadrados mínimos de $Ax = b$.
- Veamos que tiene norma mínima. Las soluciones de la ecuación por cuadrados mínimos se pueden escribir como

  $$\hat{x} = x_f + x_h$$

  Entonces
  $$\|\hat{x}\|^2 = \|x_f + x_h\|^2 = \|x_f\|^2 + \|x_h\|^2 \geq \|x_f\|^2$$
  ya que x_f y x_h son ortogonales.
  O sea,
  $$\|\hat{x}\| \geq \|x_f\|$$
  para toda solución $\hat{x}$ por cuadrados mínimos de $Ax = b$.
- x_f es la única solución por cuadrados mínimos que tiene norma mínima.
  Si x' fuera otra solución por cuadrados mínimos de norma mínima entonces

  $$\|x'\| = \|x_f\|$$

  y además podríamos escribir

  $$x' = x_f + (x' - x_f)$$

  donde $x' - x_f \in Nul(A)$. Entonces

  $$\|x'\|^2 = \|x_f + (x' - x_f)\|^2 = \|x_f\|^2 + \|x' - x_f\|^2$$

  $$\|x' - x_f\|^2 = 0$$

  Por lo tanto,
  $$x' = x_f$$

Busquemos la solución por cuadrados mínimos de norma mínima para el sistema $Ax = b$. Teníamos que la solución general es

$$\hat{x} = (1\ 1\ 1)^T + \alpha(1\ 2\ -1)^T$$

Entonces

$$x_f = P_{Fil(A)}(\hat{x}) = P_{Fil(A)}((1\ 1\ 1)^T + \alpha(1\ 2\ -1)^T) = P_{Fil(A)}((1\ 1\ 1)^T)$$

$$x_f = (1\ 1\ 1)^T - P_{Nul(A)}((1\ 1\ 1)^T) = (1\ 1\ 1)^T - \frac{1}{3}(1\ 2\ -1)^T = \left(\frac{2}{3}\ \frac{1}{3}\ \frac{4}{3}\right)^T$$

2. Dada $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

   a) Hallar la descomposición QR de A.

   b) Utilizar la descomposición QR para resolver por cuadrados mínimos el sistema $Ax = b$ con $b = (0\ 1\ 2\ \ -1)^T$.

   c) Calcular la matriz de la proyección de $\mathbb{R}^4$ sobre $Col(A)$ en la base canónica.

Recordemos que, dada una matriz $A \in \mathbb{R}^{n\times m}$ con $rg(A) = m$, una descomposición QR de A es una factorización

$$A = QR$$

con $Q \in \mathbb{R}^{n\times m}$ y $R \in \mathbb{R}^{m\times m}$ tales que $Q^TQ = I$ y R es una matriz triangular superior con números positivos en la diagonal principal. Además

- Las columnas de Q forman una base ortonormal de $Col(A)$.
- $P = QQ^T$ es la matriz de la proyección de $\mathbb{R}^n$ sobre $Col(A)$ en la base canónica.

a) Hallemos una descomposición QR de A. Para ello buscaremos una base ortonormal de $Col(A)$. Comencemos buscando una base ortogonal.

- $u_1 = (1\ 0\ 0\ 0)^T$
- $u_2 = (1\ 0\ \ -1\ 0)^T - \dfrac{\left\langle (1\ 0\ \ -1\ 0)^T, (1\ 0\ 0\ 0)^T \right\rangle}{\|(1\ 0\ 0\ 0)^T\|^2}(1\ 0\ 0\ 0)^T = (0\ 0\ \ -1\ 0)^T$
- $$u_3 = (1\ 1\ 0\ 1)^T - \frac{\left\langle (1\ 1\ 0\ 1)^T, (1\ 0\ 0\ 0)^T \right\rangle}{\|(1\ 0\ 0\ 0)^T\|^2}(1\ 0\ 0\ 0)^T - \frac{\left\langle (1\ 1\ 0\ 1)^T, (0\ 0\ \ -1\ 0)^T \right\rangle}{\|(0\ 0\ \ -1\ 0)^T\|^2}(0\ 0\ -1\ 0)^T$$

  $u_3 = (0\ 1\ 0\ 1)^T$

Una base ortogonal de $Col(A)$ es $\{(1\ 0\ 0\ 0)^T, (0\ 0\ \ -1\ 0)^T, (0\ 1\ 0\ 1)^T\}$.

Una base ortonormal de $Col(A)$ es $\{(1\ \ 0\ \ 0\ \ 0)^T, (0\ \ 0\ \ -1\ \ 0)^T, (0\ \ \frac{1}{\sqrt{2}}\ \ 0\ \ \frac{1}{\sqrt{2}})^T\}$.

Entonces

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

Para calcular R, la despejamos

$$QR = A$$
$$Q^TQR = Q^TA$$
$$R = Q^TA$$

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

b) Utilicemos la descomposición QR para resolver $Ax = b$ por cuadrados mínimos.

$$A^T Ax = A^T b$$

$$(QR)^T QRx = (QR)^T b$$
$$R^T Q^T QRx = R^T Q^T b$$
$$R^T Rx = R^T Q^T b$$

Como R es inversible, R^T también lo será y, multiplicando a ambos lados por la inversa, nos queda

$$Rx = Q^T b$$

Resolvemos entonces el sistema

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}$$

cuya solución es $x_1 = 2$, $x_2 = -2$, $x_3 = 0$. Entonces

$$\hat{x} = \begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$$

c) La matriz de la proyección sobre $Col(A)$ es

$$P = QQ^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}$$