Tema 3

Matrices ortogonales y sus aplicaciones

En este Tema trabajaremos únicamente en el espacio vectorial $\mathbb{R}^n$. Además de considerar en él las operaciones de suma de vectores y producto de un vector por un número real, nos fijaremos en el producto escalar entre vectores de $\mathbb{R}^n$.

3.1. Espacio euclídeo

Se llama **espacio euclídeo** al espacio vectorial $\mathbb{R}^n$ sobre el cuerpo $\mathbb{R}$ de los números reales, donde se tienen definidas las siguientes operaciones:

- Suma de vectores de $\mathbb{R}^n$, para dar como resultado un vector de $\mathbb{R}^n$:

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n).$$

- Producto de un vector de $\mathbb{R}^n$ por un número $t \in \mathbb{R}$, para dar como resultado un vector de $\mathbb{R}^n$:

$$t(x_1, \ldots, x_n) = (tx_1, \ldots, tx_n).$$

- Producto escalar de dos vectores $\mathbb{R}^n$ para dar como resultado un número real:

$$u \cdot v = (x_1, \ldots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \cdots x_n y_n \in \mathbb{R}.$$

En forma abreviada, denotaremos con mayúscula las columnas

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

y escribiremos $u \cdot v = X'Y$.

Recuerda que la notación *prima* se utiliza para denotar la **traspuesta de cualquier matriz**. En este caso, la fila X' es la traspuesta de la columna X. En general, si A es una matriz $m \times n$ sobre un cuerpo K, denotamos por A' a la **traspuesta de** A, matriz $n \times m$ que tiene en cada lugar (i, j) el elemento (j, i) de la matriz A:

$$A = \left(\begin{array}{c|c|c} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{array}\right) \mapsto A' = \left(\begin{array}{ccc} a_{11} & \dots & a_{m1} \\ \hline \vdots & & \vdots \\ \hline a_{1n} & \dots & a_{mn} \end{array}\right)$$

El producto escalar en $\mathbb{R}^n$ verifica las siguientes propiedades, que se deducen fácilmente de las propiedades del producto de matrices:

Propiedades del producto escalar

Dados los vectores $u, v, u_1, u_2 \in \mathbb{R}^n$, entonces:

(I) $u \cdot v = v \cdot u$.

(II) $(u_1 + u_2) \cdot v = u_1 \cdot v + u_2 \cdot v$.

(III) $(tu) \cdot v = t(u \cdot v) = u \cdot (tv)$ para todo $t \in \mathbb{R}$.

(IV) $u^2 \geq 0$. Además, $u^2 = 0 \Longleftrightarrow u = 0$.

El hecho de tener definido un producto en el espacio vectorial $\mathbb{R}^n$ nos permite hablar de conceptos como: Vectores ortogonales o perpendiculares, o norma o módulo de un vector. Notar que estos conceptos no pueden definirse en cualquier espacio vectorial.

Definición 3.1.1. (I) Dado un vector de $\mathbb{R}^n$, $u = (x_1, \dots, x_n)$, llamamos **norma** o **módulo** de u al valor

$$\| u \| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

(II) Diremos que los vectores $u, v \in \mathbb{R}^n$ son **ortogonales** si $u \cdot v = 0$. Escribiremos en ese caso $u \perp v$.

(III) Se dice que una base $\{u_1, \ldots, u_r\}$ de un subespacio vectorial S de $\mathbb{R}^n$, es una **base ortogonal** si $u_i \cdot u_j = 0$ siempre que $i \neq j$. Si, además, se cumple que $\| u_i \| = 1$ para cada $1 \leq i \leq r$, entonces la base se denomina **ortonormada**.

Ejemplo 3.1.2. La base canónica de $\mathbb{R}^n$, $\{e_1, \ldots, e_n\}$, es una base ortonormada.

Para reconocer una base ortonormada de $\mathbb{R}^n$, es suficiente escribir sus vectores en forma de columna,

$$P = \left(\begin{array}{c|c|c|c} P^1 & P^2 & \cdots & P^n \end{array} \right)$$

y estudiar P.

Si realizamos ahora el producto $P'P$, veremos el resultado de multiplicar cada vector por sí mismo y por los demás:

$$P'P = \left(\begin{array}{c} P_1 \\ \hline P_2 \\ \hline \vdots \\ \hline P_n \end{array} \right) \left(\begin{array}{c|c|c|c} P^1 & P^2 & \cdots & P^n \end{array} \right) = \left(\begin{array}{c|c|c|c} P_1P^1 & P_1P^2 & \cdots & P_1P^n \\ \hline P_2P^1 & P_2P^2 & \cdots & P_2P^n \\ \hline \vdots & \vdots & & \vdots \\ \hline P_nP^1 & P_nP^2 & \cdots & P_nP^n \end{array} \right)$$

> Una base de $\mathbb{R}^n$, $\{P^1, \cdots, P^n\}$, será ortonormada solamente si la matriz P verifica que $P'P = I_n$.

Definición 3.1.3. Se dice que P, matriz $n \times n$ sobre K, es una matriz **ortogonal** si verifica que $P'P = I_n$.

Nota 3.1.4. Si P es una matriz ortogonal, la matriz P' verifica que $P'P = I_n$. Por tanto, la matriz P es regular y su inversa es precisamente P'.

Esta es una de las ventajas de trabajar con bases de $\mathbb{R}^n$ ortonormadas. La matriz de cambio P de la base canónica a una base ortonormada tiene una inversa muy fácil de calcular, tan fácil como escribir la traspuesta de la matriz.

Veamos ahora cómo encontrar bases ortonormadas no solo en $\mathbb{R}^n$, sino en cualquier subespacio vectorial S de $\mathbb{R}^n$:

Método de ortogonalización de Gram-Schmidt:

Comenzamos tomando una base cualquiera de S, $\{u_1, \ldots, u_r\}$, con $r \geq 2$, y nos fijamos, en primer lugar, en $\{u_1, u_2\}$:

Vamos a sustituir $\{u_1, u_2\}$ por una nueva familia, $\{v_1, v_2\}$, de forma que

(I) $\mathbb{R}\langle v_1, v_2\rangle = \mathbb{R}\langle u_1, u_2\rangle$.

(II) $\{v_1, v_2\}$ es una familia ortogonal.

Para ello, mantenemos $v_1 = u_1$ y sustituimos u_2 por $v_2 = u_2 - t_1 v_1$, eligiendo t_1 de forma que $v_1 \cdot v_2 = 0$, es decir, $0 = v_1 \cdot u_2 - t_1(v_1 \cdot v_1)$. El único valor posible es

$$t_1 = \frac{u_2 \cdot v_1}{v_1 \cdot v_1}.$$

Como solo hemos realizado una operación elemental de tipo II, sabemos que el subespacio que generaban $\{u_1, u_2\}$ es el mismo que el que generan los nuevos vectores $\{v_1, v_2\}$.

Supongamos que, tras sucesivos pasos, ya hemos obtenido $\{v_1, \ldots, v_{r-1}\}$, una familia ortogonal de vectores de $\mathbb{R}^n$, con

$$\mathbb{R}\langle v_1, \ldots, v_{r-1}\rangle = \mathbb{R}\langle u_1, \ldots, u_{r-1}\rangle.$$

Consideramos entonces la familia $\{v_1, \ldots, v_{r-1}, u_r\}$ y sustituimos u_r por un nuevo vector, $v_r = u_r - t_1 v_1 - t_2 v_2 - \cdots - t_r v_{r-1}$, de forma que se verifiquen

(I) $\mathbb{R}\langle v_1, \ldots, v_{r-1}, v_r\rangle = \mathbb{R}\langle v_1, \ldots, v_{r-1}, u_r\rangle$.

(II) $\{v_1, \ldots, v_{r-1}, v_r\}$ es una familia ortogonal.

Como los vectores $v_1, \ldots, v_{r-1}$ ya son ortogonales dos a dos, bastará asegurar que

$$\begin{aligned} v_1 \cdot v_r &= 0 \\ v_2 \cdot v_r &= 0 \\ &\vdots \\ v_{r-1} \cdot v_r &= 0. \end{aligned}$$

El sistema anterior tiene como solución única:

$$\begin{aligned} t_1 &= \tfrac{u_r \cdot v_1}{v_1 \cdot v_1} \\ t_2 &= \tfrac{u_r \cdot v_2}{v_2 \cdot v_2} \\ &\vdots \\ t_r &= \tfrac{u_r \cdot v_{r-1}}{v_r \cdot v_r}. \end{aligned}$$

Una vez conseguida la base ortogonal de S, $\{v_1, \ldots, v_r\}$, sustituimos cada v_i por el vector $\dfrac{v_i}{\| v_i \|}$ para obtener una base ortonormada de S.

3.2. Diagonalización ortogonal de matrices simétricas reales

Vamos a estudiar ahora un tipo de matrices reales que siempre son diagonalizables.

Definición 3.2.1. Una matriz cuadrada A se denomina **simétrica** cuando es igual a su traspuesta: $A = A'$.

Esto significa que, para cualquier elemento a_{ij} de la matriz A, se cumple que $a_{ij} = a_{ji}$, o lo que es lo mismo, que la matriz es simétrica respecto de su diagonal principal, formada por $[a_{11}, \ldots, a_{nn}]$.

En esta sección comprobaremos que toda matriz simétrica real es diagonalizable, es decir, que existe una base de $\mathbb{R}^n$ formada por vectores propios de A. Pero además, esta base puede transformarse en una base ortonormada de $\mathbb{R}^n$ sin que deje de estar formada por vectores propios de A.

Veamos, paso por paso, todo lo que ocurrirá cuando intentemos diagonalizar una matriz simétrica real A, $n \times n$:

(i) Todas las raíces de su polinomio característico $|\lambda I_n - A|$ serán reales, $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$, por lo que su polinomio característico se podrá descomponer totalmente, de la forma
$$p(\lambda) = (\lambda - \lambda_1)^{m_1} \cdots (\lambda - \lambda_r)^{m_r},$$
con $m_1 + \cdots + m_r = n$.

(ii) Para cada valor propio $\lambda_i \in \mathbb{R}$, sabíamos que $\dim S(\lambda_i) \leq m_i$. Por ser A simétrica, se puede asegurar que
$$\dim S(\lambda_i) = m_i \text{ para cada } i = 1, \ldots, r.$$

(iii) Por tanto, en cada $S(\lambda_i)$, podremos obtener m_i vectores propios linealmente independientes de valor propio λ_i.

(iv) Aplicando el método de Gram-Schmidt a los m_i vectores, transformaremos los vectores anteriores en una base ortonormada de cada $S(\lambda_i)$.

(v) Al unir las bases ortonormadas de los distintos subespacios fundamentales, observaremos que los vectores correspondientes a distintos valores propios son ortogonales entre sí, lo que sucede solo por ser la matriz A simétrica.

(vi) Escribiendo dichos vectores como columnas, construiremos una matriz P, ortogonal, que verifica $P^{-1}AP = D$, matriz diagonal.

(vii) Dicha matriz P verifica que $P^{-1} = P'$, por lo que también cumplirá que $P'AP = D$, matriz diagonal.

Ejemplo 3.2.2. Consideremos $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$, matriz simétrica real 3×3.

(i)

$$|\lambda I_3 - A| = \begin{vmatrix} \lambda - 1 & -1 & -1 \\ -1 & \lambda - 1 & -1 \\ -1 & -1 & \lambda - 1 \end{vmatrix} = (\lambda - 3)\lambda^2,$$

que, como esperábamos, tiene solo raíces reales: $\lambda_1 = 3$, de multiplicidad $m_1 = 1$ y $\lambda_2 = 0$, de multiplicidad $m_2 = 2$.

(ii)
- Para $\lambda_1 = 3$, resolvemos el sistema $(3I_3 - A)X = 0$, que da como soluciones $\begin{cases} x = z \\ y = z \\ z = z \end{cases}$ ó $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.
- Para $\lambda = 0$, resolvemos el sistema $(0I_3 - A)X = 0$, ó simplemente $AX = 0$, que da como soluciones

$$\begin{cases} x = -y - z \\ y = y \\ z = z \end{cases} \quad \text{ó} \quad \begin{pmatrix} x \\ y \\ z \end{pmatrix} = y \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

(iii) Los subespacios fundamentales de la matriz A son:

- $S(3) = \mathbb{R}\langle(1, 1, 1)\rangle$, con $\dim S(3) = 1$ $(= m_1)$.
- $S(0) = \mathbb{R}\langle(-1, 1, 0),\ (-1, 0, 1)\rangle$, con $\dim S(0) = 2$ $(= m_2)$.

(iv) Mediante el método de Gram-Schmidt, transformamos la base de cada uno de estos subespacios en una base ortonormada:

- En $S(3)$ obtenemos así la base $\{(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}$.
- En $S(0)$, obtenemos la base $\{(-1/\sqrt{2}, 1/\sqrt{2}, 0),\ (-1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6})\}$.

(V) Como los vectores de $S(3)$ son ortogonales a los de $S(0)$, la base que resulta al unir las bases de $S(3)$ y $S(0)$, es una base ortonormada de $\mathbb{R}^3$:

$$\{(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}), (-1/\sqrt{2}, 1/\sqrt{2}, 0), (-1/\sqrt{6}, -1/\sqrt{6}, 2/\sqrt{6})\}$$

(VI) La matriz $P = \left(\begin{array}{r|r|r} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{array}\right)$ es ortogonal y verifica que

$$P'AP = P^{-1}AP = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

3.3. Complemento ortogonal

Dado un plano S en $\mathbb{R}^3$, que pase por el origen de coordenadas, existe una recta formada por todos los vectores que son ortogonales a los del plano. La denotaremos $S^\perp$.

Por ejemplo, si S es el plano de ecuación $x + y + z = 0$, la recta $S^\perp$ será la generada por el vector $(1, 1, 1)$.

Análogamente, si nos dan una recta de $\mathbb{R}^3$ que pase por el origen, el conjunto de todos los vectores de $\mathbb{R}^3$ que son ortogonales a los de la recta, formará un plano.

Definición 3.3.1. Si S es un subespacio vectorial de $\mathbb{R}^n$, se denomina **complemento ortogonal de** S a

$$S^\perp = \{v \in \mathbb{R}^n \mid u \cdot v = 0 \text{ para todo } u \in S\}.$$

Observa que cada vector $v \in S^\perp$ debe ser ortogonal a **todos** los vectores de S.

Ejercicio 3.3.2. Estudia si hay algún vector que pertenezca al mismo tiempo al subespacio S y a su complemento ortogonal $S^\perp$.

En la práctica, para hallar el complemento ortogonal de un subespacio S, tomaremos una base cualquiera de S,

$$\{u_1 = (a_{11}, a_{12}, \ldots, a_{1n}), \ldots, u_r = (a_{r1}, a_{r2}, \ldots, a_{rn})\}$$

y construiremos la matriz $r \times n$ que tenga como filas dichos vectores:

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{m2} & \ldots & a_{rn} \end{pmatrix}$$

Los vectores de $S^\perp$ son los vectores

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \text{ tales que } \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \dots & a_{rn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} (*)$$

Deducimos que $S^\perp$ es un subespacio vectorial de $\mathbb{R}^n$ con

$$\dim S^\perp = n - \dim S.$$

Ejemplo 3.3.3. Halla el complemento ortogonal del subespacio

$$S = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - y + t = 0,\ z - t = 0\}$$

y comprueba que se verifica la relación anterior entre $\dim S$ y $\dim S^\perp$.

Ejercicio 3.3.4. El sistema de ecuaciones (*) nos proporciona todos los vectores de $\mathbb{R}^n$ que son ortogonales a los de la base, es decir, a u_1, a u_2,...

Comprueba que, si un vector v es ortogonal a los vectores de la base de S, $\{u_1, \dots, u_r\}$, entonces también será ortogonal a cualquier vector de S, $u = t_1u_1 + \cdots + t_ru_r$ con $t_1, \dots, t_r \in \mathbb{R}$.

Nota 3.3.5. Si consideramos la matriz A que aparece en el sistema de ecuaciones (*), queda claro que, al resolver el sistema, hemos hallado el complemento ortogonal del subespacio generado por las filas de A. Pero ese mismo sistema se utilizaba para hallar el núcleo de la matriz A, por lo que

$$(\mathrm{Fil}\, A)^\perp = \ker A,$$

3.4. Proyección ortogonal

Si consideramos de nuevo el plano S de ecuación $x + y + z = 0$ en $\mathbb{R}^3$ y su complemento ortogonal, $S^\perp = \mathbb{R}\langle(1, 1, 1)\rangle$, es fácil ver en un dibujo que cualquier vector $v \in \mathbb{R}^3$ se puede descomponer como suma de un vector de S y otro de $S^\perp$. Por ejemplo, el vector $(1, 2, 3)$ se escribirá como

$$(1, 2, 3) = (-y - z, y, z) + (t, t, t), \text{ para algunos } (-y - z, y, z) \in S \text{ y } (t, t, t) \in S^\perp.$$

Ejercicio 3.4.1. Encuentra los valores y, z y t resolviendo el sistema de ecuaciones que se plantea. Observa que la solución es única.

En general, se cumple el siguiente resultado.

Teorema 3.4.2. *Dado un subespacio vectorial S del espacio $\mathbb{R}^n$, para cada vector $v \in \mathbb{R}^n$, existe un único $u \in S$ y un único $w \in S^\perp$ de forma que $v = u + w$.*

Definición 3.4.3. Para cada $v \in \mathbb{R}^n$, el vector u que aparece en la expresión anterior se llama **proyección ortogonal** de v sobre el subespacio S.

Si u es la proyección ortogonal de v sobre S, la otra componente de la descomposición, $w = v - u$, es ortogonal a u. Obviamente, hay otras formas de descomponer v como $v = u' + (v - u')$ con $u' \in S$, pero en estos casos la componente $v - u'$ no será ortogonal a u'. Además, se tiene el siguiente resultado.

Teorema 3.4.4. *Si u es la proyección ortogonal de $v \in \mathbb{R}^n$ sobre un subespacio vectorial S, entonces*

$$\| v - u \| = min\{\| v - u' \|; u' \in S\}.$$

3.5. Soluciones aproximadas de un sistema de ecuaciones

En esta sección, consideraremos sistemas de ecuaciones $AX = B$ del tipo

$$\left.\begin{array}{r} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array}\right\},$$

en los que los coeficientes a_{ij} y los términos b_i son números reales.

Estos sistemas pueden también escribirse también como

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \cdots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Serán compatibles si existen valores $x_1, \ldots, x_n \in \mathbb{R}$ que hagan cierta la igualdad anterior, es decir, si el vector columna B se puede escribir como combinación lineal de los vectores columna $\{A^1, \cdots, A^n\}$.

$$\boxed{\text{El sistema } AX = B \text{ es compatible solo si el vector } B \in \mathrm{Col}(A).}$$

En la práctica, podemos encontrarnos con sistemas de ecuaciones reales del tipo $AX = B$, donde los datos a_{ij} y b_i han sido tomados experimentalmente.

Debido a los errores que aparecen en las mediciones, es posible que, aunque esperemos obtener un sistema compatible, lo que obtengamos en la práctica no lo sea.

Como no podemos hablar de soluciones del sistema incompatible obtenido, hallaremos **soluciones aproximadas del sistema**, de la siguiente forma.:

(I) Si el sistema dado $AX = B$ es incompatible, sabemos que $B \notin \mathrm{Col}(A)$.

(II) Buscamos un vector columna $C \in \mathbb{R}^m$ que sea *lo más parecido posible* a B y que sí pertenezca al subespacio $\mathrm{Col}(A)$, para que el sistema $AX = C$ tenga alguna solución:

(III) Tal vector $C \in \mathbb{R}^m$ será la proyección ortogonal del vector B sobre el subespacio $\mathrm{Col}(A)$, ya que, como sabemos

$$\| B - C \| = \min \{ \| B - D \| \mid D \in \mathrm{Col}(A) \}.$$

(IV) En lugar de resolver el sistema $AX = B$, resolveremos el sistema $AX = C$, que será compatible por ser $C \in \mathrm{Col}(A)$.

Definición 3.5.1. Dado un sistema de ecuaciones sobre $\mathbb{R}$, $AX = B$, con A matriz $m \times n$, llamamos **solución aproximada del sistema** a cualquier vector columna $X \in \mathbb{R}^n$ que sea solución (ordinaria) del sistema $AX = C$, donde $C \in \mathbb{R}^m$ es la proyección ortogonal del vector B sobre el subespacio $\mathrm{Col}(A)$, contenido en $\mathbb{R}^m$.

Ejemplo 3.5.2. Vamos a hallar todas las soluciones aproximadas del sistema de ecuaciones $\left\{ \begin{array}{l} x - y + z = 0'1 \\ x - y + z = 0'2 \end{array} \right.$

El sistema escrito en forma matricial es $AX = B$,

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0'1 \\ 0'2 \end{pmatrix},$$

por lo que el subespacio generado por las columnas de A es

$$\mathrm{Col}\, A = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \rangle = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle \text{ y claramente } B \notin \mathrm{Col}\, A$$

Hallamos la proyección ortogonal de B sobre el subespacio $\operatorname{Col} A$. Para ello, podemos calcular

$$(\operatorname{Col} A)^{\perp} = \{(x, y) \mid (1, 1) \begin{pmatrix} x \\ y \end{pmatrix} = 0\} = \{(-y, y) \mid y \in \mathbb{R}\}$$

La proyección ortogonal de B sobre $\operatorname{Col} A$ será un vector C de la forma (t, t) para algún $t \in \mathbb{R}$, que verifique:

$$(0'1,\ 0'2) = (t, t) + (-y, y) \text{ para algún } y \in \mathbb{R}$$

Resolviendo el sistema, obtenemos $t = 0'15$ e $y = 0'05$, de donde $C = (0'15,\ 0'15)$.

Por tanto, las soluciones aproximadas del sistema $AX = B$ son las soluciones ordinarias del sistema $AX = C$,

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0'15 \\ 0'15 \end{pmatrix},$$

es decir,

$$\left\{ \begin{array}{l} x = y - z + 0'15 \\ y = y \\ z = z \end{array} \right. \mapsto y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0'15 \\ 0 \\ 0 \end{pmatrix} = S + u$$

Como ocurre en el ejemplo anterior, las soluciones aproximadas de cualquier sistema $AX = B$ formarán un subespacio trasladado $S + u$. De entre todas ellas, nos interesará en muchos caso encontrar la que tenga la menor norma posible.

Definición 3.5.3. (I) Se llama **solución de norma mínima** de un sistema de ecuaciones compatible a la que tiene la menor norma entre todas las soluciones obtenidas.

(II) Se llama **solución aproximada de norma mínima** de un sistema de ecuaciones incompatible $AX = B$ a la que tiene la menor norma posible entre todas las soluciones del sistema $AX = C$, siendo C la proyección ortogonal de B sobre el subespacio de columnas de A.

En un subespacio trasladado $S + u$, el vector que tiene menor norma es el único vector de $S + u$ que es ortogonal al subespacio vectorial S

Ejemplo 3.5.4. Halla la solución aproximada de noma mínima del sistema

$$\begin{cases} x - y + z = 0'1 \\ x - y + z = 0'2 \end{cases}$$

De todas las soluciones aproximadas que hemos obtenido arriba, las del subespacio trasladado $S + u$, buscamos ahora la que de menor norma.

Esta será un vector de la forma $(y - z + 0'15, y, z)$, como todos los del subespacio trasladado, que, además, sea ortogonal al subespacio

$$S = \mathbb{R}\langle(1, 1, 0),\ (-1, 0, 1)\rangle$$

Hallaremos esta solución resolviendo el sistema:

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} y - z + 0'15 \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Longleftrightarrow \begin{cases} 2y - z + 0'15 = 0 \\ -y + 2z - 0'15 = 0 \end{cases},$$

que da como solución única $y = -0'05$, $z = 0'05$.

Sustituyendo ahora estos valores en $(y - z + 0'15, y, z)$, obtenemos la única solución aproximada de norma mínima del sistema $AX = B$:

$$(x, y, z) = (0'05,\ -0'05,\ 0'05)$$

Nota 3.5.5. Si el sistema $AX = B$ dado fuera compatible, el vector B ya estaría en el subespacio $\mathrm{Col}(A)$. Por tanto, en este caso, la proyección ortogonal de B sobre este subespacio sería el propio B y el concepto de solución aproximada del sistema $AX = B$ coincidiría con el de solución ordinaria de dicho sistema.