1. Sistemas de Ecuaciones Lineales

Consideraremos sistemas de n ecuaciones y n incógnitas,

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1\\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2\\
\vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots\\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n
\end{aligned}$$

En forma matricial:

$$\mathsf{A}\mathbf{x} = \mathbf{b}$$

con

$$\mathsf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n}\\ a_{21} & a_{22} & \dots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nn}\end{bmatrix} \in \mathbb{R}^{n\times n}, \qquad \mathbf{b} = \begin{bmatrix} b_1\\ b_2\\ \vdots\\ b_n\end{bmatrix} \in \mathbb{R}^n, \qquad \mathbf{x} = \begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix} \in \mathbb{R}^n,$$

La interpretación geométrica para un sistema de 2×2 puede visualizarse en la Figura 1. El caso de una matriz singular, o una matriz mal condicionada, se representa en la Figura 2.

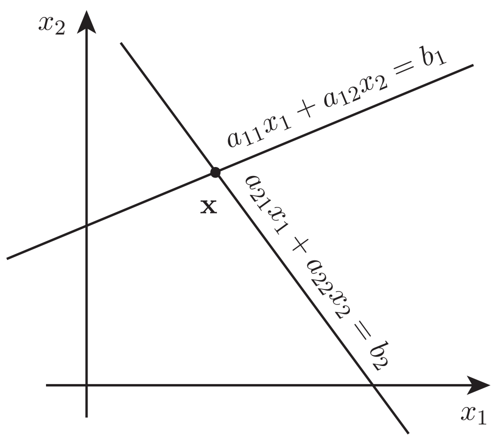

Figura 1: Sistema de dos ecuaciones lineales.

El almacenamiento de una matriz de $n \times n$ en doble precisión requiere $8n^2$ bytes. Para matrices muy grandes esto no es despreciable. Por ejemplo, si $n = 8000$ se requieren 512 MB.

Matriz plena: La mayoría de sus elementos son no nulos.

Matriz rala: La mayoría de sus elementos son nulos. Un ejemplo de matriz rala son las matrices banda. En general, no hace falta almacenar todos los ceros de una matriz rala. Algunos programas, como Matlab y Scilab, permiten trabajar con matrices ralas.

Definición 1 *Decimos que* $\mathsf{A} \in \mathbb{R}^{n\times n}$ *es una matriz* p*-banda si* $\exists\ p \in \mathbb{Z}$*,* $n > p \geq 1$*, tal que*

$$|i - j| \geq p \Rightarrow a_{ij} = 0$$

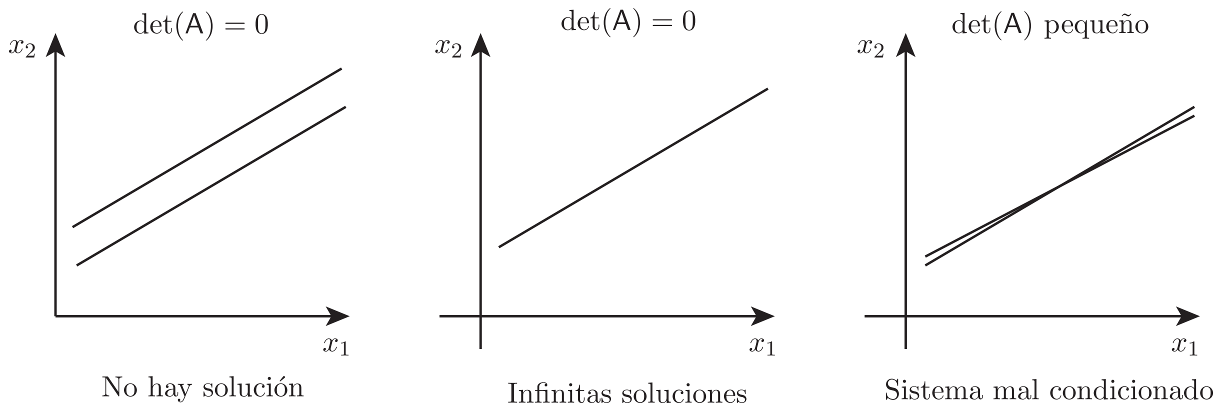

Figura 2: Ejemplo de sistemas singulares y mal condicionado. Un sistema mal condicionado es sensible a errores de redondeo.

- *Si* $p = 1 \rightarrow$ *matriz diagonal*
- *Si* $p = 2 \rightarrow$ *matriz tridiagonal*

La siguiente matriz es tridiagonal:

$$\mathsf{A} = \begin{bmatrix} a_1 & c_1 & 0 & \dots & \dots & 0 \\ b_2 & a_2 & c_2 & 0 & \dots & 0 \\ 0 & b_3 & a_3 & c_3 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & & \ddots & \ddots & c_{n-1} \\ 0 & \dots & \dots & 0 & b_n & a_n \end{bmatrix}$$

Si podemos reducir un sistema de ecuaciones $\mathsf{A}\mathbf{x} = \mathbf{b}$ a un sistema más simple $\mathsf{B}\mathbf{x} = \mathbf{d}$ que sea más sencillo de resolver, podemos resolver este último, y obtendremos de este modo las soluciones del sistema original.

Definición 2 *Sean* $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{n\times n}$*. Decimos que dos sistemas de ecuaciones lineales* $\mathsf{A}\mathbf{x} = \mathbf{b}$ *y* $\mathsf{B}\mathbf{x} = \mathbf{d}$ *son equivalentes si poseen exactamente las mismas soluciones.*

Se obtienen sistemas de ecuaciones lineales equivalentes realizando operaciones elementales por filas sobre una matriz.

Definición 3 (Operaciones elementales por filas) *Sea* $\mathsf{A} \in \mathbb{R}^{n\times n}$*. Llamaremos operaciones elementales por filas sobre* A *a cada una de las siguientes operaciones*

1) *La ecuación* E_i *puede multiplicarse por una constante* $\lambda \neq 0$ *y la ecuación resultante se emplea en vez de* E_i*. Esta operación se denota por* $(\lambda E_i) \rightarrow (E_i)$*.*

2) *La ecuación* E_j *puede multiplicarse por cualquier constante* λ *y sumarse a la ecuación* E_i*, la ecuación resultante se emplea en vez de* E_i*. Esta operación se denota por* $(E_i + \lambda E_j) \rightarrow (E_i)$*.*

3) *El orden de las ecuaciones* E_i *y* E_j *puede intercambiarse. Esta operación se denota por* $(E_i) \leftrightarrow (E_j)$*.*

2. Determinantes

El determinante es un número real asociado a una matriz cuadrada, que permite calcular varias propiedades como la singularidad, el polinomio característico, y los autovalores de la matriz.

- El determinante de una matriz de 1×1 es: $\det(a) = a$.
- El determinante de una matriz de 2×2 es:

$$\det(\mathsf{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

- El determinante de una matriz de 3×3 es:

$$\det(\mathsf{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Veamos ahora como calcular el determinante de una matriz cuadrada cualquiera. Para ello, empleamos el concepto de menor adjunto de una matriz. Sea $\mathsf{A} = (a_{ij})_{n\times n}$ una matriz de $n \times n$. Dado un par de índices (i, j) representamos por A_{ij} a la matriz que resulta al eliminar la i-ésima fila y la j-ésima columna de A. El *menor adjunto* δ_{ij} de A es el número real dado por:

$$\delta_{ij} = (-1)^{i+j}|\mathsf{A}_{ij}|$$

Teorema 1 (Teorema de Laplace) *Dada una matriz $\mathsf{A} \in \mathbb{R}^{n\times n}$, su determinante se puede calcular mediante el desarrollo de Laplace por una fila (o columna) cualquiera, de la siguiente forma:*

- *Si elegimos la i-ésima fila,* $|\mathsf{A}| = a_{i1}\delta_{i1} + \cdots + a_{in}\delta_{in}$
- *Si elegimos la j-ésima columna,* $|\mathsf{A}| = a_{1j}\delta_{1j} + \cdots + a_{nj}\delta_{nj}$

El desarrollo de Laplace no es práctico para matrices de órdenes superiores ya que su grado de complejidad es $O(n!)$. Resulta más práctico convertir la matriz a triangular mediante la Eliminación de Gauss y multiplicar la diagonal, ya que en ese caso la complejidad se reduce a $O(n^3)$.

Veamos ahora algunas propiedades de los determinantes:

- En el teorema de Laplace vemos que el valor del determinante no cambia cuando se intercambian las filas por columnas, es decir, $|\mathsf{A}| = |\mathsf{A}^\mathsf{T}|$.
- Si $\mathsf{A} \in \mathbb{R}^{n\times n}$ es triangular, su determinante es el producto de los elementos de la diagonal principal,

$$\mathsf{A} \in \mathbb{R}^{n\times n} \text{ triangular } \Rightarrow \det(\mathsf{A}) = a_{11}a_{22}\dots a_{nn}$$

- Teniendo en cuenta que la matriz identidad es triangular, se deduce que $\det(\mathbf{I}) = 1$.

3. Rango de una Matriz

El rango de una matriz se define como (a) el máximo número de columnas linealmente independientes de la matriz, o (b) el máximo número de filas linealmente independientes de la matriz. Ambas definiciones son equivalentes.

Cotas en el rango de un producto de matrices

Si $\mathsf{A} \in \mathbb{R}^{m\times n}$ y $\mathsf{B} \in \mathbb{R}^{n\times p}$, luego

$$\text{rango}(\mathsf{A}) + \text{rango}(\mathsf{B}) - n \leq \text{rango}(\mathsf{AB}) \leq \min\{\text{rango}(\mathsf{A}), \text{rango}(\mathsf{B})\}$$

En particular, si $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{n\times n}$ son no singulares, luego

$$n \leq \text{rango}(\mathsf{AB}) \leq n,$$

es decir, $\text{rango}(\mathsf{AB}) = n$, y el producto AB es también no singular.

Productos $\mathsf{A}^\mathsf{T}\mathsf{A}$ y $\mathsf{A}\mathsf{A}^\mathsf{T}$

Para $\mathsf{A} \in \mathbb{R}^{m\times n}$ se cumple

$$\text{rango}(\mathsf{A}^\mathsf{T}\mathsf{A}) = \text{rango}(\mathsf{A}) = \text{rango}(\mathsf{A}\mathsf{A}^\mathsf{T})$$

En particular, si $n \leq m$ y $\text{rango}(\mathsf{A}) = n$, luego $\mathsf{A}^\mathsf{T}\mathsf{A}$ es no singular.

4. Matriz Inversa

Sea $\mathsf{A} \in \mathbb{R}^{n\times n}$. Decimos que A^{-1} es la inversa de A si

$$\mathsf{A}\mathsf{A}^{-1} = \mathsf{A}^{-1}\mathsf{A} = \mathbf{I}$$

Si la inversa de A existe, luego es única.
Para una matriz de 2×2, la matriz inversa esta dada por:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}, \qquad \text{si} \;\; ad - bc \neq 0$$

El siguiente teorema resume las condiciones para que una matriz tenga inversa.

Teorema 2 *Sea $\mathsf{A} \in \mathbb{R}^{n\times n}$. Los siguientes enunciados son equivalentes:*

1) A *tiene inversa.*

2) $\det(\mathsf{A}) \neq 0$.

3) *Las filas de A son linealmente independientes (forman una base de $\mathbb{R}^n$).*

4) *Las columnas de A son linealmente independientes (forman una base de $\mathbb{R}^n$).*

5) *El sistema $\mathsf{A}\mathbf{x} = \mathbf{0}$, donde $\mathbf{x} \in \mathbb{R}^n$, tiene como única solución $\mathbf{x} = \mathbf{0}$.*

6) *Para cada $\mathbf{b} \in \mathbb{R}^n$, el sistema $\mathsf{A}\mathbf{x} = \mathbf{b}$, posee una única solución $\mathbf{x} \in \mathbb{R}^n$.*

7) $\text{rango}(\mathsf{A}) = n$.

8) *Todos los autovalores (o valores propios) de A son distintos de cero.*

Un sistema de ecuaciones lineales $\mathsf{A}\mathbf{x} = \mathbf{b}$, con $\mathsf{A} \in \mathbb{R}^{n\times n}$ y $\det(\mathsf{A}) \neq 0$ se puede resolver invirtiendo la matriz A como sigue:

$$\mathbf{x} = \mathsf{A}^{-1}\mathbf{b}$$

Sin embargo, mediante la Eliminación de Gauss se puede resolver el sistema de forma más eficiente sin calcular la inversa.

5. Matriz Simétrica

$\mathsf{A} \in \mathbb{R}^{n\times n}$ es simétrica si y solo si

$$a_{ij} = a_{ji}, \qquad 1 \le i, j \le n$$

A simétrica significa que $\mathsf{A}^\mathsf{T} = \mathsf{A}$.

Teorema 3 *Sea $\mathsf{A} \in \mathbb{R}^{n\times n}$ una matriz simétrica real. Luego, existe un conjunto de n pares autovalor-autovector $\{\lambda_i, \mathbf{v}_i\}$, $i = 1, \ldots, n$, que satisfacen las siguientes propiedades:*

a) Los autovalores $\lambda_1, \ldots \lambda_n$, son las raíces del polinomio característico $f(\lambda)$ de A. Todos los autovalores λ_i son números reales, y pueden repetirse de acuerdo a su multiplicidad.

b) Los autovectores $\mathbf{v}_1, \ldots, \mathbf{v}_n$ son ortogonales entre sí, y pueden elegirse de longitud unitaria:

$$\begin{aligned} \mathbf{v}_i^\mathsf{T}\mathbf{v}_j &= 0, \qquad 1 \le i, j \le n, \quad i \ne j \\ \mathbf{v}_i^\mathsf{T}\mathbf{v}_i &= 1, \qquad 1 \le i \le n \end{aligned}$$

c) Para cada vector $\mathbf{x} = [x_1, x_2, \ldots, x_n]^\mathsf{T}$ existe un único vector $\mathbf{c} = [c_1, c_2, \ldots, c_n]^\mathsf{T}$ tal que:

$$\mathbf{x} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Si los autovectores son de longitud unitaria, las constantes están dadas por

$$c_i = \mathbf{x}^\mathsf{T}\mathbf{v}_i, \quad i = 1, \ldots, n$$

d) Sea la matriz de autovectores

$$\mathsf{P} = \left[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\right]$$

Luego

$$\mathsf{P}^\mathsf{T}\mathsf{A}\mathsf{P} = \mathsf{D} = \begin{bmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{bmatrix}$$

Además,

$$\begin{aligned} &\mathsf{P}\mathsf{P}^\mathsf{T} = \mathsf{P}^\mathsf{T}\mathsf{P} = \mathbf{I} \\ &\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^\mathsf{T} \end{aligned}$$

6. Matriz Definida Positiva

Definición 4 *Sea $\mathsf{A} \in \mathbb{R}^{n\times n}$ una matriz simétrica.*

- A *es* **definida positiva** *si sus autovalores son positivos, es decir, $\lambda_i > 0$, $i = 1, \ldots, n$.*
- A *es* **semidefinida positiva** *si sus autovalores son no negativos, es decir, $\lambda_i \ge 0$, $i = 1, \ldots, n$.*
- A *es* **definida negativa** *si sus autovalores son negativos, es decir, $\lambda_i < 0$, $i = 1, \ldots, n$.*

- A *es* **semidefinida negativa** *si sus autovalores son no positivos, es decir,* $\lambda_i \leq 0$, $i = 1, \ldots, n$.
- A *es* **indefinida** *si no cumple ninguna de las definiciones anteriores.*

Teorema 4 *Una matriz* $\mathsf{A} \in \mathbb{R}^{n\times n}$ *simétrica es semidefinida positiva si y solo si* A *se puede factorizar como* $\mathsf{A} = \mathsf{B}^\mathsf{T}\mathsf{B}$, *y* A *es definida positiva si y solo si* B *es no singular.*

Demostración. Siendo $\mathsf{A} \in \mathbb{R}^{n\times n}$ simétrica, existe una matriz ortogonal P tal que $\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^\mathsf{T}$, donde $\mathsf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ es real. Si $\lambda_i \geq 0$, para $i = 1, \ldots, n$, luego $\mathsf{D}^{1/2}$ existe,

$$\mathsf{A} = \mathsf{P}\mathsf{D}\mathsf{P}^\mathsf{T} = \mathsf{P}\mathsf{D}^{1/2}\mathsf{D}^{1/2}\mathsf{P}^\mathsf{T} = \mathsf{B}^\mathsf{T}\mathsf{B}, \quad \text{con} \;\; \mathsf{B} = \mathsf{D}^{1/2}\mathsf{P}^\mathsf{T}$$

Siendo $\mathrm{rango}(\mathsf{P}) = n$, y $\mathrm{rango}(\mathsf{D}^{1/2}) = n$ si y solo si $\lambda_i > 0$, para $i = 1, \ldots, n$, se cumple que B es no singular si y solo si $\lambda_i > 0$, para $i = 1, \ldots, n$.

A la inversa, si A se puede factorizar como $\mathsf{A} = \mathsf{B}^\mathsf{T}\mathsf{B}$, luego todos los autovalores de A son no negativos ya que para cualquier par autovalor-autovector $(\lambda_i, \mathbf{v}_i)$, $i = 1, \ldots, n$, tenemos:

$$\lambda_i = \frac{\mathbf{v}_i^\mathsf{T}\mathsf{A}\mathbf{v}_i}{\mathbf{v}_i^\mathsf{T}\mathbf{v}_i} = \frac{\mathbf{v}_i^\mathsf{T}\mathsf{B}^\mathsf{T}\mathsf{B}\mathbf{v}_i}{\mathbf{v}_i^\mathsf{T}\mathbf{v}_i} = \frac{\|\mathsf{B}\mathbf{v}_i\|^2}{\|\mathbf{v}_i\|^2} \geq 0$$

donde $\|\cdot\|$ es la norma vectorial euclídea. Si B es no singular, luego $\mathsf{B}\mathbf{v}_i \neq \mathbf{0}$ y $\lambda_i > 0$. □

Teorema 5 *Una matriz simétrica* $\mathsf{A} \in \mathbb{R}^{n\times n}$ *es definida positiva si y solo si* $\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} > 0$, *para todo* $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.

Demostración. Si A es definida positiva, luego $\mathsf{A} = \mathsf{B}^\mathsf{T}\mathsf{B}$, para B no singular, luego

$$\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathsf{B}^\mathsf{T}\mathsf{B}\mathbf{x} = \|\mathsf{B}\mathbf{x}\|^2 \geq 0$$

con igualdad si y solo si $\mathbf{x} = \mathbf{0}$.

A la inversa, si $\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} > 0$ para todo $\mathbf{x} \neq \mathbf{0}$, luego para cada par autovalor-autovector $(\lambda_i, \mathbf{v}_i)$, $i = 1, \ldots, n$, tenemos

$$\lambda_i = \frac{\mathbf{v}_i^\mathsf{T}\mathsf{A}\mathbf{v}_i}{\mathbf{v}_i^\mathsf{T}\mathbf{v}_i} > 0$$

lo cual completa la demostración. □

Teorema 6 *Para matrices* A *reales simétricas, los siguientes enunciados son equivalentes y sirven como definición de matriz definida positiva.*

- $\mathbf{x}^\mathsf{T}\mathsf{A}\mathbf{x} > 0, \quad \forall\, \mathbf{x} \in \mathbb{R}^n, \;\; \mathbf{x} \neq \mathbf{0}$.
- *Todos los autovalores de* A *son estrictamente positivos.*
- $\mathsf{A} = \mathsf{B}^\mathsf{T}\mathsf{B}$ *para alguna matriz* B *no singular.* B *no es única, pero existe una única matriz triangular superior* R *con elementos diagonales positivos tal que* $\mathsf{A} = \mathsf{R}^\mathsf{T}\mathsf{R}$. *Esta es la factorización de Cholesky de* A.
- *Toda submatriz principal de* A *es definida positiva.*

7. Ortogonalización de Gram-Schmidt

Sea $\beta = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ una base arbitraria (no ortogonal) de un espacio euclídeo n-dimensional, $\mathcal{S}$.

Objetivo: A partir de β construir una base ortogonal normalizada $O = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ del espacio $\mathcal{S}$.

Estrategia: Construir O secuencialmente de modo que $O_k = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$ es una base ortogonal normalizada de $\mathcal{S}_k = \operatorname{span}\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$, para $k = 1, \ldots, n$.

El primer vector $\mathbf{u}_1$ se obtiene normalizando $\mathbf{x}_1$,

$$\mathbf{w}_1 = \mathbf{x}_1, \qquad \mathbf{u}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}$$

Luego definimos

$$\mathbf{w}_2 = \mathbf{x}_2 - \alpha \mathbf{u}_1$$

Debemos elegir α tal que $\mathbf{w}_2$ sea ortogonal a $\mathbf{u}_1$,

$$0 = \mathbf{w}_2^\top \mathbf{u}_1 = (\mathbf{x}_2^\top - \alpha \mathbf{u}_1^\top)\mathbf{u}_1 = \mathbf{x}_2^\top \mathbf{u}_1 - \alpha \|\mathbf{u}_1\|^2$$

de donde

$$\alpha = \mathbf{x}_2^\top \mathbf{u}_1$$
$$\mathbf{w}_2 = \mathbf{x}_2 - (\mathbf{x}_2^\top \mathbf{u}_1)\mathbf{u}_1$$

Luego, definimos $\mathbf{u}_2$ normalizando $\mathbf{w}_2$,

$$\mathbf{u}_2 = \frac{\mathbf{w}_2}{\|\mathbf{w}_2\|}$$

La interpretación gráfica de la construcción de $\mathbf{u}_2$ puede verse en la Figura 3.

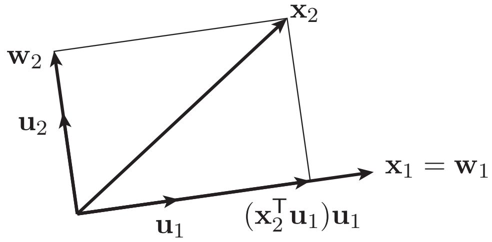

Figura 3: Ortogonalización de Gram-Schmidt.

Generalizando este procedimiento, definimos

$$\mathbf{w}_k = \mathbf{x}_k - \sum_{i=1}^{k-1} (\mathbf{x}_k^\top \mathbf{u}_i)\mathbf{u}_i, \qquad \mathbf{u}_k = \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|}$$

Vemos que para $k = 3$ obtenemos

$$\mathbf{w}_3^\top = \mathbf{x}_3^\top - (\mathbf{x}_3^\top \mathbf{u}_1)\mathbf{u}_1^\top - (\mathbf{x}_3^\top \mathbf{u}_2)\mathbf{u}_2^\top,$$

y se compruba que

$$\mathbf{w}_3^\mathsf{T}\mathbf{u}_1 = 0 \qquad \text{y} \qquad \mathbf{w}_3^\mathsf{T}\mathbf{u}_2 = 0.$$

Bibliografía

1. Carl D. Meyer, *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia PA, United States, 2000.