Pontificia Universidad Católica de Chile
Facultad de Matemáticas
Departamento de Matemática
Álgebra Lineal (mat1203)

Apuntes-Resumen del Curso

Por Sebastián Soto Rojas (spsoto@uc.cl)

Probablemente sea extraño que un “resumen” tenga aproximadamente 65 páginas, pero Álgebra Lineal es un curso muy denso en contenidos, definiciones y teoremas. Como diría un profesor de la Facultad, podría resultar incluso más denso conceptualmente que cursos tan temidos como los Cálculos en Física e Ingeniería. Sin embargo, históricamente ha habido un dejo de mecanización en los problemas y ejercicios típicos de este curso, y esto ha llevado a una poca valoración del contenido conceptual y de la amplia visión que puede dar este curso de toda la Matemática y su lenguaje. Esto está cambiando en los semestres próximos, y se está intentando que el curso tenga un enfoque cada vez más conceptual, lo que ha incrementado su dificultad.

Por esta razón, urge cada vez más la comprensión de los conceptos y contenidos explicados, los cuales no siempre, por no decir nunca, suelen ser entendibles a la primera, y luego se puede proceder a su sistematización, sin nunca abandonar el concepto y el razonamiento detrás.

Estos apuntes-resumen son el reflejo de una buena cantidad de horas invertidas (y no dormidas) en compendiar de forma lo mejor posible todos los apuntes de todos los profesores que hacen y han hecho Álgebra Lineal alguna vez, señalando además, los razonamientos fundamentales, cuando estos sean necesarios.

En ningún caso estos apuntes reemplazan a los de un profesor del curso, y menos a los que debieran ser tomados en clase, sin embargo, son una referencia rápida a los contenidos del curso **una vez asimilados los conceptos**, o cuando haya algo particular que recordar. No sólo sucederá esto durante el transcurso del semestre, si no que durante el transcurso de toda la carrera, donde en casi todas las especialidades requerirán el uso de ciertos conceptos fundamentales vistos en el curso.

Éxito a todos en el curso, y en el examen. Es posible pasarlo con buena nota, pero requiere el esfuerzo de comprender cada concepto, por muy tedioso que pueda lucir a veces. Cualquier feedback, comentario, aporte, sugerencia y/o error detectado en este documento es siempre bien recibido.

Muchas gracias por las colaboraciones de: Pilar Jadue A. (pijadue@uc.cl) y Francisco Carrasco C. (ftcarrasco@uc.cl) en el desarrollo y/o revisión de estos apuntes.

Basado en los apuntes de los profesores:

1. Carla Barrios
2. Carolina Becerra
3. Iván Huerta

Índice

1. Vectores en $\mathbb{R}^n$

Cabe preguntarse, ¿qué es $\mathbb{R}^n$? Es una pregunta para la que se requieren elementos teóricos antes de poder responder, pero por ahora, se puede responder diciendo que es un *espacio vectorial*. Es decir, un espacio donde están contenidos vectores, y se verifican algunas operaciones y axiomas básicos. Por ahora, basta con conformarse con la *pertenencia* de un *vector* a $\mathbb{R}^n$.

1.1. Definición de vector

Definición: Sea $n \in \mathbb{N}$, un vector en $\mathbb{R}^n$ es conjunto ordenado de n elementos en $\mathbb{R}$. La notación es la siguiente:

$$v = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = (\alpha_1, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & \ldots & \alpha_n \end{pmatrix}^T \quad \text{donde cada } \alpha_i \in \mathbb{R} \text{ y } v \in \mathbb{R}^n$$

A veces se anota $\overrightarrow{v}$, pero no es condición necesaria. Los vectores se pueden escribir usando simplemente letras minúsculas.

Observación: Para $\mathbb{R}^2$ y $\mathbb{R}^3$ se tienen representaciones en el plano cartesiano. A partir de estas se pueden hacer comprensiones de fenómenos de la geometría en general de $\mathbb{R}^n$. Sin embargo, para $n \geq 4$, vectores que **sí** seran necesarios, ya se carece de representaciones gráficas, por lo que todo el trabajo con vectores es puramente algebraico, esa es una de las ventajas del álgebra lineal.

1.2. Operaciones vectoriales

Existen muchísimas operaciones en Álgebra Lineal, sin embargo, las más básicas, y a partir de las que se pueden construir las demás, son la suma y la ponderación por escalar. Algo se dice *lineal*, si verifica los axiomas de estas dos operaciones, que serán definidas a continuación. (¿Se puede hacer alguna referencia de lineal a línea, recta?)

1.2.1. Suma

Definición: Sean $x, y \in \mathbb{R}^n$, entonces se define la *suma vectorial* como $+ : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n}_{\text{dos vectores de } \mathbb{R}^n \text{ a } \mathbb{R}}$:

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix}$$

1.2.2. Ponderación

Definición: Sea $x \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, entonces se define la *ponderación* como una operación $\mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ tal que:

$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$$

No confundir: cada x_i es un número real que define una coordenada del vector, y x es el vector completo, que pertenece a $\mathbb{R}^n$.

Nota: Prescindiremos de los signos $\cdot$ y $\times$, pues estos serán representaciones de otras dos operaciones particulares en el álgebra lineal. Todos las demás operaciones, además de estas, deben evitar usar esta notación.

1.2.3. Propiedades de la suma y la ponderación

Propiedades: Para todo $x, y, z \in \mathbb{R}^n$ y $\alpha, \beta \in \mathbb{R}$ se tiene que:

1. *Conmutatividad*: $x + y = y + x$
2. *Asociatividad*: $(x + y) + z = x + (y + z)$
3. *Elemento neutro*: $x + \overrightarrow{0} = x$. Cero ahora es un vector, y no un escalar.
4. *Inverso aditivo*: $x + (-x) = \overrightarrow{0}$. De esta forma, se puede definir la resta vectorial a partir del inverso aditivo.
5. $(\alpha\beta)\, x = \alpha\,(\beta x)$.
6. $(\alpha + \beta)x = \alpha x + \beta x$.
7. $\alpha(x + y) = \alpha x + \alpha y$.
8. $1x = x$.
9. $\alpha \frac{1}{\alpha} x = x$, con $\alpha \neq 0$.

Notar que $0x = \overrightarrow{0}$.

1.3. Combinaciones lineales y conjunto generado

Definición: Un vector $x \in \mathbb{R}^n$ se dice *combinación lineal* (c.l.) de un conjunto $S = \{x_1, \ldots, x_k\}$ si existen escalares $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ (pudiendo ser uno, e incluso todos cero) tales que:

$$x = \alpha_1 x_1 + \ldots + \alpha_k x_k$$

Además, una combinación lineal se dice *convexa* si $\sum_{i=1}^{k} \alpha_k = 1$.

Definición: Sea $S = \{x_1, \ldots, x_k\}$ un conjunto de vectores en los que cada $x_i \in \mathbb{R}^n$. Entonces el conjunto generado de S, anotado como $\langle S \rangle$ es aquel conjunto que satisface:

$$\langle S \rangle = \{x \in \mathbb{R}^n \,:\, x \text{ es cl. de } S\} = \alpha_1 x_1 + \ldots + \alpha_k x_k \quad \forall \alpha_i \in \mathbb{R}^n$$

Evidentemente, para $\mathbb{R}^n$ y con un conjunto no vacío distinto de $\{0\}$, la cardinalidad (cantidad de elementos) de este conjunto es **infinita**. Sin embargo, este infinito tiene condiciones según la cantidad de vectores y los vectores determinados.

Observación: $\overrightarrow{0} \in \langle S \rangle$ para cualquier conjunto S no vacío.

1.4. Geometría en $\mathbb{R}^n$

Incluso para $n \geq 4$ (e incluso para espacios vectoriales más extraños aún) se puede desarrollar una geometría en que se incluyan los dos conceptos fundamentales: ángulo y distancia, con la definición adecuada de ciertas operaciones, convirtiendo todo el proceso geométrico en uno algebraico.

1.4.1. Producto punto

Definición: Sean $x, y \in \mathbb{R}^n$, entonces se define el producto punto como aquella operación $\cdot\, : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ (convierte dos vectores a un escalar) tal que:

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = x_1 y_1 + \ldots + x_n y_n$$

Es decir, se multiplica componente a componente y se suma.

Propiedades: Sean $x, y, z \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, entonces:

1. $x \cdot y = y \cdot x$.
2. $x \cdot (y + z) = x \cdot y + x \cdot z$.
3. $\alpha(x \cdot y) = (\alpha x) \cdot y = x \cdot (\alpha y)$.
4. $\overrightarrow{0} \cdot x = \overrightarrow{0}$.
5. $x \cdot x \geq 0$.
6. $x \cdot x = 0 \Longleftrightarrow x = 0$.

Notar que $x \cdot y = 0$ **no** implica que $x = 0$ ó $y = 0$ necesariamente, basta notar que:

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

Observación: Esta observación es importante, pues tiene implicancia en los desarrollos posteriores:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} + \beta \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

Se tiene que:

$$x_1 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ z_1 \end{pmatrix}$$

Y así simultáneamente. Se tiene o mismo de forma análoga para cualquier c.l.

1.4.2. Norma

Definición: Sea $x \in \mathbb{R}^n$. Se define la *norma* o *longitud* (distancia al origen) de un vector como:

$$\|x\| = \sqrt{x \cdot x}$$

De aquí se desprende que:

$$\|x\| = \sqrt{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Para $\mathbb{R}^2$ y $\mathbb{R}^3$ se puede hacer la comprobación. Para espacios de orden superior, basta con hacer uso de la definición.

Propiedades: Sea $x, y \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, entonces:

1. $\|x\|^2 = x \cdot x \geq 0$.
2. $\|x\| = 0 \iff x = 0$.
3. $\|\alpha x\| = |\alpha| \, \|x\|$.
4. $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2x \cdot y$.

Teorema: *(Desigualdad de Cauchy-Schwarz)* Sean $x, y \in \mathbb{R}^n$, entonces:

$$|x \cdot y| \leq \|x\| \, \|y\|$$

Y $|x \cdot y| = \|x\| \, \|y\|$ si y sólo si $x = \alpha y$ para algún $\alpha \in \mathbb{R}$.

Corolario: *(Desigualdad Triangular)* Sean $x, y \in \mathbb{R}^n$, entonces:

$$\|x+y\| \leq \|x\| + \|y\|$$

Definición: Para $x, y \in \mathbb{R}^n$ se define el *ángulo entre vectores* como:

$$\alpha = \arccos\left(\frac{x \cdot y}{\|x\|\,\|y\|}\right)$$

Observaciones:

1. $x \parallel y \Longleftrightarrow x = \alpha y$ para $\alpha \neq 0$.
2. $x \perp y \Longleftrightarrow x \cdot y = 0$.

Definición: $x \in \mathbb{R}^n$ se dice *unitario* si $\|x\| = 1$.

Observación: Para todo vector $x \in \mathbb{R}^n$ se puede obtener su forma unitaria $\widehat{x}$ considerando:

$$\widehat{x} = \frac{x}{\|x\|}$$

1.4.3. Distancia

Definición:

1. Un punto P en $\mathbb{R}^n$ está representado por su *vector posición*, el cual parte desde el orígen y llega hasta el punto P.
2. Sean $x, y \in \mathbb{R}^n$, estos representan puntos del espacio. Luego, la distancia está determinada por:
$$d(x,y) = \|y - x\|$$

1.5. Rectas

Definición: Una recta $\mathcal{L}$ en $\mathbb{R}^n$ está definida como:

$$\mathcal{L} \,:\, p + \langle d \rangle \quad \text{con } p, d \in \mathbb{R}^n$$

Donde p representa el *vector posición*, un vector contenido en la recta que indica su posición con respecto al origen, y d representa el *vector dirección*, el cual se ve determinado por la resta de dos puntos pertenecientes a esta.

Observación: Sean:

$$\mathcal{L}_1 \,:\, p_1 + \langle d_1 \rangle$$
$$\mathcal{L}_2 \,:\, p_2 + \langle d_2 \rangle$$

Se tiene que:

- $\mathcal{L}_1 \parallel \mathcal{L}_2 \iff d_1 = \alpha d_2$.
- $\mathcal{L}_1 \perp \mathcal{L}_2 \iff d_1 \perp d_2 \iff d_1 \cdot d_2 = 0$.
- $\mathcal{L}_1 = \mathcal{L}_2$ si y sólo si $d_1 \parallel d_2$ y $p_1 - p_2 \parallel d_1$.

1.5.1. Ecuaciones de la recta

Definición: Sea $\mathcal{L}_1$ una recta en $\mathbb{R}^n$, entonces:

- La *ecuación vectorial* de $\mathcal{L}_1$ se representa como:

$$p + \lambda d \quad \text{con } \lambda \in \mathbb{R}$$

- La *ecuación paramétrica* de $\mathcal{L}_1$ se representa como:

$$\begin{cases} x_1 = p_1 + \lambda d_1 \\ \vdots \\ x_n = p_n + \lambda d_n \end{cases} \quad \text{con } \lambda \in \mathbb{R}$$

- Las *ecuaciones cartesianas* de $\mathcal{L}_1$ se representan como:

$$\begin{cases} \alpha_1 x_1 + \ldots + \alpha_n x_n \\ \\ \mu_1 x_1 + \ldots + \mu_n x_n \end{cases}$$

  El número de ecuaciones necesario depende de varios parámetros, pero deben permitir llegar a los otros tipos de ecuaciones.

1.6. Hiperplanos

Definición: Para $\mathbb{R}^n$ y dados un vector fijo $n \in \mathbb{R}^n$ y un escalar $b \in \mathbb{R}^n$ se define un *hiperplano* $\mathcal{H}$ como un subconjunto de $\mathbb{R}^n$ para el que:

$$x \in \mathcal{H} \iff n \cdot x = b$$

Un plano es un caso particular de hiperplano. La diferencia fundamental es la imposibilidad de ser graficado de este último. El hiperplano es una generalización de plano para $\mathbb{R}^n$.

Para poder hacer un estudio más exhaustivo de los hiperplanos, se requiere el uso de sistemas de ecuaciones.

2. Sistemas de Ecuaciones

2.1. Conceptos básicos

Definición: Una *ecuación lineal* con n incógnitas es una expresión de la forma:

$$\alpha_1 x_1 + \ldots + \alpha_n x_n = b \quad \text{con } a_i, x_i, b \in \mathbb{R}$$

Cada α_i corresponde a un *coeficiente*, x_i a una *incógnita* y b se conoce como *coeficiente libre*.

Observación: Toda ecuación lineal de n incógnitas representa a un hiperplano de $\mathbb{R}^n$. En particular,

$$\underbrace{\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}}_{n} \cdot \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{x} = b$$

Definición: Se define un *sistema de ecuaciones* de m ecuaciones y n incognitas como un conjunto de ecuaciones lineales que se cumplen simultáneamente:

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = b_1 \\ \vdots \\ \alpha_{n1}x_1 + \ldots + \alpha_{nn}x_n = b_n \end{cases}$$

La primera interpretación es una *intersección* de interplanos, es decir, se buscan los vectores x que pertenezcan simultaneamente a varios hiperplanos dados. Se buscarán los vectores $x = \begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}^T$ que pertenezcan a todos los hiperplanos dados simultáneamente.

Definición: Dos sistemas de ecuaciones se dicen *equivalentes* si las soluciones de cada sistema coinciden.

2.2. Operaciones elementales

Definición: Se conocen como operaciones elementales a aquellas operaciones que aplicadas a un sistema de ecuaciones, obtienen uno equivalente. Estas son:

1. *Permutación*: Se cambia intercambia el orden en que se muestran los sistemas de ecuaciones. Se representa como
$$E_i \longleftrightarrow E_j$$
2. *Escalamiento*: Se amplifica toda una ecuación lineal de un sistema de ecuaciones por un escalar $\alpha \in \mathbb{R}$ distinto de 0.
$$\alpha E_i \longrightarrow E_i$$
3. *Eliminación*: Se suma un múltiplo distinto de 0 de una ecuación lineal a otra:
$$E_i + \alpha E_j \longrightarrow E_i$$

2.3. Notación matricial

La notación matricial permite simplificar la resolución de un sistema de ecuaciones, pues sólo se consideran sus coeficientes numéricos.

Definición: Una matriz A de $m \times n$ es un conjunto ordenado de m filas y n columnas.

$$A = \begin{bmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & & & \vdots \\ a_{m1} & \dots & \dots & a_{mn} \end{bmatrix} \longleftarrow \text{fila 1}$$

Una forma de notar la matriz es:

$$A = [a_{ij}]_{m \times n}$$

Donde:

1. m representa el número de filas.
2. n representa el número de columnas.
3. i representa el indicador de la fila.
4. j representa el indicador de la columna.

Proceso: Se tiene un sistema de ecuaciones dado por:

$$\begin{cases} \alpha_{11}x_1 + \ldots + \alpha_{1n}x_n = b_1 \\ \\ \alpha_{m1}x_1 + \ldots + \alpha_{mn}x_n = b_n \end{cases}$$

La notación es directamente:

$$\begin{bmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{bmatrix}$$

Definición: Todo el sistema de ecuaciones puede resumirse en:

$$Ax = b$$

Donde se considerará:

- La *matriz de los coeficientes*, que es aquella que contiene los coeficientes dependientes del sistema de ecuaciones.
- La *columna de resultados*, que contiene los coeficientes libres. Es un vector de $\mathbb{R}^m$.

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

- El *vector solución*, que contiene los x_i que serán solución del sistema. Pertenece a $\mathbb{R}^n$.

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- La *matriz ampliada*, $[A|b]$, que contiene toda la información esencial del sistema. Es decir, se toma:

$$\left[\begin{array}{cccc} \alpha_{11} & \dots & \alpha_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} & b_m \end{array}\right]$$

En esta última notación, es posible hacer el uso rápido de las operaciones elementales.

2.4. Algoritmo de eliminación de Gauss

Definición: Una matriz está en *forma escalonada* (F.E.) si:

1. Las filas que contienen únicamente ceros (nulas) están bajo las filas no nulas.
2. En cada fila no nula, avanzando de izquierda a derecha hay un primer elemento distinto de 0 que se llama pivote. El pivote de cada fila siempre está como menos una columna más a la derecha que el pivote de la fila anterior.
3. Cada columna que contiene un pivote tiene ceros abajo de él.

Definición: Una matriz está en *forma escalonada reducida* (F.E.R.) si:

1. Está en forma escalonada.
2. Todos los pivotes son iguales a 1.
3. En cada columna con pivote, el único elemento distinto de cero es el pivote.

Algoritmo: Para una matriz de $A_{m\times n}$, el algoritmo de eliminación Gaussiana lleva a la matriz A a su $F.E.$ en primera instancia:

1. Se asume que la primera columna no es nula. Si lo fuera, el proceso comienza en la primera que no sea nula.
2. Buscar el primer elemento de la columna distinto de 0. Si no está en la primera fila, se hace una permutación. Este será el pivote de la primera fila.
3. Eliminar hacia abajo la columna, dejando ceros bajo el pivote. Para la fila i-ésima y la columna $j - ésima$ se realiza considerando la operación de eliminación:

$$F_i - \frac{a_{ij}}{a_{jj}} F_j \longrightarrow F_i$$

4. Se obtiene la matriz con el pivote. Se considera la submatriz comenzando desde la siguiente fila y la siguiente columna y se sigue repitiendo el proceso.

Este proceso se conoce como *pivotear* la matriz. No es necesario que todas las columnas tengan pivote.

Para llevarla a su F.E.R, se toma cada pivote y se aplica la operación:

$$F_i - \frac{a_{ij}}{a_{ii}} F_j \longrightarrow F_j$$

Donde i es una fila superior al pivote, por lo que $i < j$.

2.5. Solución general

Definición: Sea A una matriz de $m \times n$ y $\mathrm{FE}(A)$ su forma escalonada, equivalente a la matriz A. Se define el *rango gaussiano*, $r(A)$, como el número de pivotes de $\mathrm{FE}(A)$.

Teorema: Si S es el conjunto de soluciones de un sistema de ecuaciones lineales, entonces:

$$S = \emptyset \quad \text{ó} \quad \#S = 1 \quad \text{ó} \quad \#S = \infty$$

En el tercer caso, la solución se puede expresar como un conjunto generado.

Definición: Para un sistema lineal de ecuaciones de m ecuaciones y n incógnitas

- Un sistema lineal sin solución se dirá *incompatible* o *inconsistente*. En este caso, $r(A) \neq r(A|b)$. Aparece un pivote en la columna de resultados.
- Un sistema lineal con solución única se dirá *compatible determinado*. Se tiene que $r(A) = r(A|b) = n$.
- Un sistema lineal con infinitas soluciones se dirá *compatible indeterminado*. Algunas variables no quedarán determinadas por las ecuaciones, estas se dirán *variables libres*. Se cumple que $r(A) = r(A|b) < n$. Además:

$$\text{número de variables libres} = \text{número de variables} - r(A)$$

Las variables libres satisfacen el sistema para cualquier valor que se les de, y siempre se corresponden (no exclusivamente) con las filas con pivotes. Por lo tanto, estas pueden formar parte de un conjunto generado en la solución.

2.6. Aplicaciones

2.6.1. Sistemas simultáneos

Se pueden resolver varios sistemas simultáneos de la forma:

$$[A|b_1] \ldots [A|b_2]$$

Pivoteando una sola vez la matriz:

$$[A|b_1| \ldots |b_k]$$

El sistema será equivalente, y los resultados los mismos.

2.6.2. El producto Ax

Definición: Sea $x \in \mathbb{R}^n$ y sea A una matriz de n columnas, las cuales son $\{v_1, \ldots, v_n\}$, entonces se define el producto de una matriz por un vector:

$$\underbrace{[v_1| \ldots |v_n]}_{A} \underbrace{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}}_{x} = \underbrace{x_1 v_1 + \ldots + x_n v_n}_{b}$$

Es decir, al resolver el sistema $Ax = b$ uno también puede responder a la pregunta de cuáles son los coeficientes que acompañan a $v_1, \ldots, v_n$ y que determinan al vector b.

Teorema: Un sistema de ecuaciones $[A|b]$ es compatible si y solo si el vector b es c.l. de las columnas de A.

Propiedades:

1. $A(x + y) = Ax + Ay$.
2. $\alpha Ax = A(\alpha x)$, con $\alpha \in \mathbb{R}$.

2.6.3. Sistemas homogéneos e independencia lineal

Definición: Un sistema se dirá *homogéneo* si su columna de resultados es el vector 0.

Teorema: Para $x \in \mathbb{R}^n$, el sistema homogeneo Ax es compatible si y solo si $r(A) = n$.

Teorema: Si S_0 es el conjunto de soluciones del sistema $Ax = 0$ y x_p es solución particular conocida del sistema $Ax = b$, entonces el conjunto de soluciones de este último esta dado por:

$$S = x_0 + S_0$$

Definición:

- El conjunto de vectores $\{u_1, \ldots, u_n\}$ será *linealmente independiente* (l.i.) cuando la única forma de escribir 0 como combinación lineal de ellos sea la trivial, es decir, multiplicando cada vector por cero.
$$\alpha_1 u_1 + \ldots + \alpha_n u_n = 0 \Longrightarrow \alpha_i = 0 \quad \forall i \leq n$$
- El conjunto de vectores $\{u_1, \ldots, u_n\}$ será *linealmente dependiente* (l.d.) cuando haya otra forma distinta de la trivial para escribir el 0 como c.l. de ellos. Esto implica que al menos uno de ellos se puede escribir como c.l. del resto de los vectores.

Teorema: Para el conjunto $\{v_1, \ldots, v_m\}$ donde cada $v_i \in \mathbb{R}^n$. Si $m > n$, entonces el conjunto es l.d.

2.6.4. Ecuaciones cartesianas a conjunto generado, obtención de la solución de un sistema de ecuaciones lineales

Las ecuaciones lineales y cartesianas que representan a un hiperplano pueden llevar a su transcripción a la ecuación como conjunto generado, y de igual forma la ecuación derivada de su definición.

Dado el sistema, este se resuelve pivoteando la matriz respectiva y se obtendrán ecuaciones de la forma. Se revisará un caso concreto. Por ejemplo, se tiene el sistema ya resuelto:

$$\begin{array}{ccccccccccc} x_1 & + & 0 & + & \alpha x_3 & + & 0 & + & 0 & = & b_1 \\ & & x_2 & + & \beta x_3 & + & 0 & + & \gamma x_5 & = & b_2 \\ & & & & & & x_4 & + & \delta x_5 & = & b_3 \end{array}$$

Luego, las variables libres siempre (pero no exclusivamente) corresponden a las columnas sin pivote. Se despejan las "columnas" con pivote, es decir, se despeja en términos de x_1, x_2, x_4:

$$\begin{array}{ccc} x_1 & = & b_1 - \alpha x_3 \\ x_2 & = & b_2 - \beta x_3 - \gamma x_5 \\ x_4 & = & b_3 - \delta x_5 \end{array}$$

Es decir, nuestro conjunto generado tendrá $5 - 3 = 2$ vectores.

Sabemos que un vector solución es de la forma $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix}$.

Reemplazando con las igualdades anteriormente obtenidas, se tiene que:

$$x = \begin{pmatrix} b_1 - \alpha x_3 \\ b_2 - \beta x_3 - \gamma x_5 \\ x_3 \\ b_3 - \delta x_5 \\ x_5 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ 0 \\ b_3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -\alpha \\ -\beta \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} 0 \\ -\gamma \\ 0 \\ -\delta \\ 1 \end{pmatrix}$$

Como x_3 y x_5 son variables libres, e independientes una de otra, pueden ser cualquier valor real. Por lo tanto, la solución, o el hiperplano según corresponda a la interpretación, queda expreado como:

$$S = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ 0 \\ b_3 \\ 0 \end{pmatrix}}_{\text{posición}} + \left\langle \begin{pmatrix} -\alpha \\ -\beta \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -\gamma \\ 0 \\ -\delta \\ 1 \end{pmatrix} \right\rangle$$

Con la práctica adecuada, este método es generalizable a cualquier solución de sistema de ecuaciones, e incluso puede ser desprendido el conjunto generado directamente de la matriz.

3. Matrices

3.1. Transformaciones lineales en $\mathbb{R}^n$

Definición: Una transformación lineal (t.l.) de $\mathbb{R}^n$ en $\mathbb{R}^m$ es una función $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ que cumple:

1. $T(x+y) = T(x) + T(y)$.
2. $T(\alpha x) = \alpha T(x)$.

Donde, $\alpha \in \mathbb{R}$ y $x, y \in \mathbb{R}^n$.

Teorema: Sean $T_1 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ y $T_2 : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dos t.l., entonces $T_1 + T_2$ también es una t.l. de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Teorema: Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una t.l. y $\lambda \in \mathbb{R}$, entonces λT también es una t.l. de $\mathbb{R}^n$ en $\mathbb{R}^m$.

Teorema: Sean: $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $S : \mathbb{R}^m \longrightarrow \mathbb{R}^p$ dos t.l. Entonces, $S \circ T$ es una t.l. de $\mathbb{R}^n$ en $\mathbb{R}^p$.

3.2. Matrices canónicas de una transformación lineal

Definición: Sea $n \in \mathbb{N}$ fijo, entonces se definen los vectores canónicos como:

$$\begin{gathered} e_1 = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \end{pmatrix}^T \\ \vdots \\ e_i = \underbrace{\begin{pmatrix} 0 & \dots & 1 & \dots & 0 \end{pmatrix}^T}_{1 \text{ en la posición } i\text{-ésima}} \\ \vdots \\ e_n = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \end{pmatrix}^T \end{gathered}$$

Todos estos vectores e_i pertenencen al $\mathbb{R}^n$ que se esté usando.

Notar que $\|e_i\| = 1$ y $e_i \cdot e_j = 0$ para $i \neq j$.

Observación: Entonces, para todo $x \in \mathbb{R}^n$ se tiene que:

$$x = x_1 e_1 + \ldots + x_n e_n = [e_1|\ldots|e_n] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{con } x_i \in \mathbb{R}$$

Es decir **todo** vector se puede escribir como combinación lineal de sus vectores canónicos.

Método: Toda transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ puede ser representada por una matriz canónica. Es decir, se tiene que:

$$T(x) = Ax$$

Con $A_{m \times n}$, y $x \in \mathbb{R}^n$.

Notar que con la observación anterior, se tiene que:

$$T(x) = T(x_1 e_1 + \ldots + x_n e_n)$$

$$\begin{aligned} T(x) &= T(x_1 e_1 + \ldots + x_n e_n) \\ &= x_1 T(e_1) + \ldots + x_n T(e_n) \\ &= [T(e_1)|\ldots|T(e_n)] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

Sabiendo la definición de la transformación, se puede calcular fácilmente cada $T(e_i)$. Por lo tanto, la obtención de la matriz es directa. Se tiene que:

$$A = [T(e_1)|\ldots|T(e_n)]$$

Definición: Dada una t.l. de $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, la matriz de $m \times n$:

$$A = [T(e_1)|\ldots|T(e_n)]$$

es la *matriz canónica* de la transformación lineal T.

3.3. Operaciones matriciales

3.3.1. Suma

Definición: Sean A, B matrices de $m \times n$, se sabe por teorema que su suma también es una transformación lineal, considerando que:

$$A = [a_1|\dots|a_n] \quad \text{y} \quad B = [b_1|\dots|b_n]$$

Donde cada a_i y b_i es la columna respectiva de A ó B.

$$A + B = [a_1 + b_1|\dots|a_n + b_n]$$

Propiedades: Sean A y B matrices de $m \times n$:

1. $A + B = B + A$.
2. $A + (B + C) = (A + B) + C$.
3. Existe un neutro aditivo, la matriz de $m \times n$ denominada $\mathbb{O}$ y que está compuesta únicamente por ceros.
4. Existe el inverso aditivo, $-A$, que cumple que:

$$A + (-A) = \mathbb{O}$$

3.3.2. Ponderación por escalar

Definición: Sea A una matriz de $m \times n$ y $\alpha \in \mathbb{R}$, entonces si A está definida como:

$$A = [a_1|\dots|a_n]$$

Entonces, la ponderación de una matriz por un escalar α está definida por:

$$\alpha A = [\alpha a_1|\dots|\alpha a_n]$$

Propiedades: Sean A y B matrices de $m \times n$ y $\alpha, \beta \in \mathbb{R}$, entonces:

1. $(\alpha + \beta)A = \alpha A + \beta B$.
2. $\alpha(A + B) = \alpha A + \alpha B$.
3. $(\alpha + \beta)(A + B) = \alpha A + \beta A + \alpha B + \beta B$.
4. $(\alpha\beta)A = \alpha(\beta A) = \beta\,(\alpha A)$.
5. $1A = A$, $(-1)A = -A$ y $0A = \mathbb{O}$.

3.3.3. Multiplicación de matrices

Definición: Sean A una matriz de $m \times n$ y B una matriz de $n \times p$, entonces la matriz AB es una matriz de $m \times p$ es la multiplicación de ambas matrices y representa a la composición de las transformaciones lineales respectivas asociadas a ellas. Si

$$B = [b_1 | \dots | b_n]$$

Entonces:

$$AB = A\,[b_1 | \dots | b_n] = [Ab_1 | \dots | Ab_n]$$

Es decir, si se considera a A por filas:

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix}$$

Entonces, se tiene que:

$$AB = \begin{bmatrix} a_1^T b_1 & \dots & a_1^T b_n \\ \vdots & \ddots & \vdots \\ a_n^T b_1 & \dots & a_n^T b_n \end{bmatrix} = \begin{bmatrix} a_1 \cdot b_1 & \dots & a_1 \cdot b_n \\ \vdots & \ddots & \vdots \\ a_n \cdot b_1 & \dots & a_n \cdot b_n \end{bmatrix}$$

Esta es la definición que puede resultar más rápida: producto punto de filas por columnas. Para que dos matrices $A_{m\times n}$ y $B_{n\times p}$ puedan multiplicarse. Debe cumplirse que las dimensiones de los extremos interiores coincidan.

Definición: Sea $n \in \mathbb{N}$, la matriz $\mathbb{I}_n$ de $n \times n$ se conoce como *matriz de identidad de orden matriz de identidad de orden* n, y está definida como:

$$\mathbb{I}_n = [e_1 | \dots | e_n] = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

Es decir, es una matriz con unos en la diagonal y ceros en cualquier otra posición.

Propiedades: Asumiendo que A, B y C son matrices que sí pueden ser operadas entre sí, se tiene que:

1. $(AB)C = A(BC)$.
2. $A(B+C) = AB+AC$, y $(A+B)C = AC+BC$. Mucho cuidado con multiplicar por la izquierda o la derecha.
3. $\alpha(AB) = (\alpha A)B = A(\alpha B)$.
4. $\mathbb{O}A = A\mathbb{O} = \mathbb{O} \longleftarrow$con la dimensión que le corresponda.
5. $A_{m\times n}\mathbb{I}_n = A$, $\mathbb{I}_m A = A$.

Tenga cuidado:

1. $AB \neq BA$ en general.
2. **NO** hay leyes de cancelación. $AB = AC$ no implica que $A = C$.
3. $AB = \mathbb{O}$ no implica que $A = \mathbb{O}$ ó $B = \mathbb{O}$.

3.3.4. Transposición: La matriz transpuesta

Definición: Sea $A_{m\times n}$, entonces se define la matriz transpuesta de A, que se anota como A^T, como una matriz de $n \times m$ que se obtiene de intercambiar la fila como columna.

$$A = [a_{ij}]_{m\times n} \Longrightarrow A^T = [a_{ji}]_{n\times m}$$

Observaciones:

1. $(\forall x \in \mathbb{R}^n)\,(\forall y \in \mathbb{R}^m) \quad (Ax)\cdot y = x\cdot(A^T y)$.
2. $x \cdot y = x^T y$.

Propiedades: Para matrices A, B las cuales sí pueden ser multiplicadas y/o sumadas entre ellas según corresponda:

1. $(A^T)^T = A$.
2. $(A+B)^T = A^T + B^T$.
3. $(\alpha A)^T = \alpha A^T$, donde $\alpha \in \mathbb{R}$.
4. $(AB)^T = B^T A^T$.

3.4. Transformaciones lineales inyectivas y sobreyectivas

3.4.1. Conceptos básicos

Definición: Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una transformación lineal. Se definen los conjuntos:

1. El *núcleo* o *kernel* de T como el conjunto:

$$\mathrm{Ker}(T) = \{x \in \mathbb{R}^n \,:\, T(x) = 0\}$$

2. La *imagen* de T como el conjunto:

$$\mathrm{Im}(T) = \{y \in \mathbb{R}^m \,:\, \exists x \in \mathbb{R}^n \,:\, T(x) = y\}$$

Teorema: Sea $A_{m\times n}$ la matriz de una t.l. $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$, se tiene que si

$$A = [a_1|\ldots|a_n] \quad \text{con cada } a_i \in \mathbb{R}^m$$

1. $\mathrm{Ker}(T) = \{x \in \mathbb{R}^n \,:\, Ax = 0\}$.
2. $\mathrm{Im}(T) = \langle a_1, \ldots, a_n \rangle = \langle b_1, \ldots, b_k \rangle$ con $k \leq n$ y donde el conjunto $\{b_1, \ldots, b_k\}$ es l.i.

Definición: Sea $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una t.l. y $A_{m\times n}$ la matriz canónica asociada. Entonces:

1. T y su matriz A se dicen inyectivas si y solo si
$$(\forall x \in \mathbb{R}^n)\,(\forall y \in \mathbb{R}^n) \quad T(x) = T(y) \Longrightarrow x = y$$
2. T y su matriz A se dicen sobreyectivas si y solo si
$$\mathrm{Rec}(T) = \mathbb{R}^m$$

3.4.2. Relación con la matriz canónica

Teorema: Sea $A_{m\times n}$, entonces todas las siguientes afirmaciones son equivalentes entre sí. Si se cumple una, se cumplen todas.

- A es sobre.
- Las filas de A son l.i.
- Los sistemas $Ax = e_1, \ldots, Ax = e_m$ son compatibles.
- $Im(A) = \mathbb{R}^m$.
- El sistema $Ax = b$ tiene solución para todo $b \in \mathbb{R}^m$.
- $r(A) = m \Rightarrow m \leq n$.

Teorema: Análogamente, para la matriz $B_{m\times n}$ las siguientes afirmaciones son equivalentes:

- B es 1-1.
- Las columnas de B son l.i.
- $Ker(B) = \{\overrightarrow{0}\}$ ó $Bx = 0$ tiene solución única.
- Si $b \in Im(B)$, entonces $Bx = b$ tiene solución única.
- $r(B) = n \Rightarrow m \geq n$.

3.5. Matrices inversas

Definición: Sea $A_{m\times n}$, entonces se dirá que:

1. A tiene *inversa por la izquierda* si existe una matriz X de $n \times m$ tal qu
$$XA = \mathbb{I}_n$$
2. A tiene *inversa por la derecha* si existe una matriz X de $n \times m$ tal que:
$$AX = \mathbb{I}_m$$

Siempre que se cumpla esta condición X será la inversa por la izquierda o la derecha según corresponda. En general, estas matrices **no** son únicas.

3.5.1. Inversa por la derecha

Teorema: A tiene inversa por la derecha si y solo si A es sobre.

Método: Sólo en caso de que no se cumpla la particularidad, existirán infintias matrices que cumplen esta condición.

1. Se toma $[A|\mathbb{I}]$ y se lleva a su FER.
2. Se resuelve la solución general para cada sistema.
3. Cada sistema es una columna de la matriz inversa por la derecha. Se asigna cualquier valor a la parte generada, si es que existe, o se deja expresado en términos de variables.

Evidentemente, si las columnas no son l.i., no será única.

3.5.2. Inversa por la izquierda

Teorema: A tiene inversa por la izquierda si y solo si A es 1-1.

Método: Se aprovecha el hecho de que ya sabemos cómo calcular la inversa por la derecha:

1. $XB = \mathbb{I} \Leftrightarrow B^T X^T = \mathbb{I}$.
2. Se encuentra la matriz por el método de búsqueda de inversas por la derecha. (Se pivotea $[B^T|\mathbb{I}]$).
3. Cada columna de X^T pasará a ser la fila respectiva de X.

Corolario: Sea $A_{m\times n}$, entonces:

1. Si A tiene inversa por la derecha, entonces $m \leq n$.
2. Si A tiene inversa por la izquierda, entonces $n \leq m$.
3. Si A tiene inversa por la izqueirda y por la derecha, entonces $m = n$.

3.6. Inversa de matrices cuadradas

Teorema: $A_{n\times n}$ tiene inversa por la izquierda si y solo si A tiene inversa por la derecha.

Corolario:

1. En este caso, la inversa es única.
2. Se verifican simultáneamente todos los teoremas de sobreyectividad e inyectividad para esta matriz: es biyectiva.

Definición: Sea $A_{n\times n}$ con $r(A) = n$, entonces la única matriz que es inversa por la derecha e izquierda se llama *matriz inversa* de A y únicamente en este caso se anota A^{-1}. Entonces, A es la única matriz que cumple que:

$$AA^{-1} = A^{-1}A = \mathbb{I}_n$$

Además, si A tiene inversa, se dice *invertible*.

Método de cálculo: Notar que como la matriz será 1-1 y a la vez cuadrada, cada columna tendrá un pivote, por lo tanto al pivotear $[A|\mathbb{I}]$ se tendrá una solución unica para cada sistema. Por lo tanto, $FER(C|\mathbb{I}) = [\mathbb{I}|A^{-1}]$.

3.6.1. Propiedades de las matrices invertibles

Todas las demostraciones de invertibilidad se realizan bajo el siguiente principio: *Si B es una matriz inversa de A, entonces $AB = \mathbb{I}$ y $BA = \mathbb{I}$.*

- $(A^{-1})^{-1} = A$.
- $(\forall \alpha \in \mathbb{R}\backslash\{0\}) \quad (\alpha A)^{-1} = \dfrac{1}{\alpha}A^{-1}$.
- Si A y B son invertibles, entonces AB también lo es, y $(AB)^{-1} = B^{-1}A^{-1}$.
- A es invertible $\Leftrightarrow A^T$ es invertible. Se sigue que $(A^T)^{-1} = (A^{-1})^T$.

4. Factorizaciones matriciales

4.1. Matrices elementales

Definición: Son matrices cuadradas que se obtienen al realizar una operación elemental (permutación, escalamiento, eliminación) a la matriz identidad $\mathbb{I}$.

- **Permutación:** P_{ij} representa la matriz elemental de permutación, intercambiando la fila i con la fila j, o bien la columa i con la columna j. Ejemplo de $(P_{35})_{5\times 5}$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

- **Escalamiento:** $E_i(\alpha)$ representa la matriz elemental de escalamiento, donde se multiplica la fila i (o columna i) por $\alpha \in \mathbb{R}$. Ejemplo de $E_3(4)_{5\times 5}$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- **Eliminación:** $E_{ij}(c)$ con $c \in \mathbb{R}$ representa la matriz elemental de eliminación, donde se suma a la fila i c veces la fila j. En eliminación gaussiana, tenemos que siempre se cumple que $i > j$. Ejemplo de $E_{23}(-4)$:
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

  Notar que es simplemente ubicar c en la posición (i, j) de la matriz elemental. Se puede pensar haciendo también la misma operación elemental en la matriz que nos piden.

4.2. Inversas y transpuestas de matrices elementales

Matriz Elemental	**Matriz Inversa**	**Matriz Transpuesta**
P_{ij}	P_{ij} (1)	$P_{ji} = P_{ij}$ (2)
$E_i(\alpha)$	$E_i(\frac{1}{\alpha}) \Leftrightarrow \alpha \neq 0$	$E_i(\alpha)$ (3)
$E_{ij}(c)$	$E_{ij}(-c)$ (4)	$E_{ji}(c)$ (5)

(1) Notar que la operación inversa de intercambiar una fila con otra es volver a intercambiar la misma fila de nuevo.

(2) Notar que lo apropiado seria pensar que se intercambian columnas con columnas, pero esto es exactamente lo mismo que intercambiar filas con filas en una matriz de permutación.

(3) Aplicarle la transformación a la fila i o a la columna i es exactamente lo mismo, por eso en la transposición no se altera.

(4) Es lo mismo que restarle lo que anteriormente se habia sumado.

(5) Ver la matriz de ejemplo: el número agregado se traslada a la posición (j, i).

Teorema: Sea $A_{m\times n}$ una matriz cualquiera y $M_{m\times m}$ una matriz elemental, entonces

$$A' = MA$$

es la matriz obtenida al aplicar a A la operación elemental que representa la matriz M. *Las operaciones elementales se aplican por la izquierda a la matriz.*

IMPORTANTE: Las operaciones elementales por columnas se realizan de la siguiente forma: $A' = AM^T$, considerando que M es la operación aplicada a una fila. Notar también que debe coincidir con la multiplicación de la matriz por la derecha.

Recordar: Toda matriz A cuadrada e invertible se puede llevar a su $FER(A) = \mathbb{I}$, por medio de operaciones elementales. Por lo tanto, $O_k \dots O_1 A = \mathbb{I}$. Luego $A = (O_k \dots O_1)^{-1}$, donde O_i representa una operación elemental.

4.3. Factorización A = LU

Definición:

1. $A = (a_{ij})_{m\times n}$ es *triangular superior* si $a_{ij} = 0$ con $i > j$, es decir, si todos los elementos bajo la diagonal son ceros. Esto no quiere decir que los elementos en la diagonal o sobre la diagonal tengan que ser distintos de 0.
2. De forma análoga, $B = (b_{ij})_{m\times n}$ es *triangular inferior* si $b_{ij} = 0$ con $i < j$, es decir, si todos los elementos sobre la diagonal son ceros, sin decir nada sobre los elementos en la diagonal o bajo ella.
3. $C = (c_{ij})_{m\times n}$ es *diagonal* si $c_{ij} = 0$ si $i \neq j$. ¡También pueden aparecer ceros en la diagonal!

Teoremas: Considerando las definiciones anteriores:

- Si A y B son triangulares superiores, entonces AB es triangular superior.
- Si A y B son triangulares inferiores, entonces AB es triangular inferior.
- Si A y B son ambas triangulares superiores o inferiores con 1s en la diagonal, entonces AB es triangular superior o inferior según corresponda, y con 1s en la diagonal.
- Si A es triangular inferior y triangular superior, entonces A es diagonal. Por ejemplo, la matriz 0 es diagonal.
- Si A es triangular superior $\Rightarrow$ A^T es triangular inferior. La expresión recíproca también se verifica.
- Si A es triangular superior e invertible $\Rightarrow$ A^{-1} es triangular superior. Basta con notar que los pivotes ya estarían determinados en este caso. Si presenta unos en la diagonal, la matriz inversa también los presentará.

Definición: Sea $A_{m\times n}$ una matriz que no requiera operaciones de permutación para ser llevada a su FE, es decir, basta con usar operaciones elementales de eliminación, entonces

$$A = LU$$

donde:

- L es una matriz triangular inferior cuadrada de $m \times m$.
- U es una matriz triangular superior que corresponde a una forma escalonada de A.

Obtención de U: Simplemente se lleva A a su F.E. registrando las operaciones elementales efectuadas.

Obtención de L: $A \sim\sim\sim\sim FE(A) = E_k \dots E_1 A$. Donde E_i representa una operación elemental de eliminación. Notar que todas las E_i son triangulares inferiores con unos en la diagonal (pues en la eliminación gaussiana $i > j$).

$$U = \underbrace{E_k \dots E_i}_{L^{-1}} A$$

$$L = (E_k \dots E_1)^{-1} = E_1^{-1} \dots E_k^{-1}$$

Importante: Recordemos que para las matrices de eliminación se cumple que $E_{ij}(c)^{-1} = E_{ij}(-c)$ en L, luego, **y sólo en el caso de la obtención de L** operaciones consecutivas de eliminación verifican que se ubicarán en su respectivo casillero en la matriz final. Por ejemplo:

$$\begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}$$

Esto sólo se verifica debido al orden en que se multiplican las matrices de eliminación debido a la factorización LU, por lo que no es una generalidad para el producto de matrices elementales de eliminación.

Observación: Si describimos $A_{m\times n} = LU$ mediante filas, se tiene que: $\overrightarrow{A_k} = \sum_{i=1}^{m} l_{k,i} \cdot \overrightarrow{U_i}$.

$$\begin{bmatrix} A_1 \\ \vdots \\ A_m \end{bmatrix} = \begin{bmatrix} 1 & & \mathbb{O} \\ \vdots & \ddots & \\ l_{m1} & \cdots & 1 \end{bmatrix} \begin{bmatrix} U_1 \\ \vdots \\ U_m \end{bmatrix}$$

$A_1 = 1 \cdot U_1$

$\vdots$

$A_4 = l_{41}U_1 + l_{42}U_2 + l_{43}U_3 + U_4$

$\vdots$

$A_m = \underbrace{l_{m1}U_1 + l_{m2}U_2 + \ldots + 1 \cdot U_m}_{m\, términos} \longleftarrow$ sólo el ultimo termino va acompañado de un 1.

Se concluye que la fila i-ésima de A es una c.l. de todas las filas de U con la fila i-ésima de L como coeficientes.

4.4. Factorización PA = LU

Extiende la factorización $A = LU$ a matrices que requieran operaciones de permutación para ser llevadas a su F.E.

Definición: Sea $A_{m\times n}$ una matriz que requiera una cantidad mínima de k operaciones de permutación para poder ser llevada a su F.E., entonces:

$$PA = LU$$

donde:

- P es una matriz de permutación con el producto de las k operaciones. Es cuadrada e invertible, por ser producto de matrices elementales.
- L y U son las matrices inferior cuadrada y superior respectivas.

Observaciones: Si bien en general la conmutación del producto de matrices no se cumple, si se cumplen las siguientes propiedades:

- $P_{ij} \cdot E_{kl}(c) = E_{kl}(c) \cdot P_{ij}$, donde $i, j \neq k$ e $i, j \neq l$.
- $P_{ij} \cdot E_{ij}(c) = E_{ji}(c) \cdot P_{ij}$. En la eliminación se intercambian los coeficientes.
- $P_{ij} \cdot E_{jk}(c) = E_{ik}(c) \cdot P_{ij}$.
- $P_{ij} \cdot E_{ki}(c) = E_{kj}(c) \cdot P_{ij}$. Concluimos que en estos dos últimos casos se intercambia la letra coincidente con la otra de la permutación.

Método: Con estas observaciones, se puede realizar la factorización. Notemos que:

$A \sim\sim\sim\sim FE(A) = U = E_{ij}(c_1)\dots P_{uv}E_{mn}(c_i)\dots E_{jk}(c_g)A$.

Se pueden efectuar las conmutaciones para que U quede como:

$$U = \overbrace{E_i \dots E_1}^{L^{-1}} \cdot \underbrace{P_k \dots P_1}_{P} A$$

Y de esta forma se puede finalmente efectuar la factorización $PA = LU$.

4.5. Aplicaciones de la factorización PA = LU

Sabiendo P, L y U se pueden resolver todos estos tipos de problemas sin la necesidad de encontrar A.

4.5.1. Resolución de Ax = b

$$\begin{array}{rcl} Ax & = & b \\ PAx & = & Pb \\ L\underbrace{Ux}_{y} & = & Pb \end{array}$$

(1) Se resuelve $Ly = Pb$. Pb se obtiene rápidamente, pues b es multiplicado por una matriz de permutación. El sistema se obtiene por sustitución directa, pues L es triangular y ya está pivoteada.

(2) Se resuelve $Ux = y$. También se obtiene rápidamente, pues U ya está pivoteada.

4.5.2. Encontrar $A^{-1}b$

$$\begin{array}{rcl} A^{-1}b & = & x \\ b & = & Ax \end{array}$$

Este sistema se resuelve como en el caso anterior.

4.5.3. Resolver AX = B

$$\begin{array}{rrcl} P\cdot\backslash & AX & = & B \\ & PAX & = & PB \\ \Leftrightarrow & L\underbrace{UX}_{Y} & = & PB \end{array}$$

(1) Se resuelve $LY = PB \Leftrightarrow L[y_1|\ldots|y_m] = [Pb_1|\ldots|Pb_m]$. Entonces, se resuelve pivoteando $[L|PB]$, ¡pero L es invertible! Entonces $FER(L|PB) = [\mathbb{I}|Y]$.

(2) Se resuelve $UX = Y$. Se resuelve de forma análoga: se obtiene $FER(U|Y)$ y se obtiene el valor de cada columna de X.

4.5.4. Resolver $A^T x = b$

$$PA = LU \;/()^T \iff A^T P^T = U^T L^T \iff A^T = U^T L^T P$$

Entonces:

$$A^T x = b \iff U^T \underbrace{L^T \underbrace{Px}_{y}}_{z} = b$$

(1) Se resuelve $U^T z = b$. U^T es triangular inferior, luego bastará pivotear "hacia abajo".

(2) Se resuelve $L^T y = z$. L^T es triangular superior y 1-1, luego se resuelve rápidamente.

(3) Se resuelve $Px = y \Leftrightarrow x = P^{-1}y = P^T y$.

4.6. Factorización de Cholesky

4.6.1. Factorización LU de matrices simétricas

Definición: $A_{n\times n}$ es simétrica si y sólo si $A = A^T$, lo que viene a significar que cada término se ve reflejado por la diagonal de la matriz.

Primera Factorización de Cholesky: Si A es simétrica y factorizable como $A = LU$, es decir, no se requieren operaciones de permutación para factorizarla, entonces:

$$A = LDL^T$$

donde:

- L ya lo conocemos.
- D es una matriz diagonal, que contiene a la diagonal de U.

4.6.2. Formas cuadráticas

Definición: Una forma cuadrática es una función $Q:\ \mathbb{R}^n \longrightarrow \mathbb{R}$ si para todo $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ se cumple que todo término no nulo del recorrido es de grado 2. Formalmente: $Q(\overrightarrow{x}) = \sum_{i=1}^{n} m_i x_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{i-1} r_{ij} x_i x_j$.

Definición: La matriz A representa a una forma cuadrática si y sólo si:

$$(\forall x \in \mathbb{R}^n)\ Q(x) = x^T A x$$

Existen, eso sí, infinitas matrices que representan a una f.c. $Q(x)$.

Teorema: Existe una única matriz simétrica S que representa a $Q(x)$.

Notar: $Q(\overrightarrow{0}) = 0 \in \mathbb{R}$.

Clasificación de Formas Cuadráticas:

- *Definida positiva*: $(\forall x \neq \overrightarrow{0})\ Q(x) > 0$.
- *Semidefinida positiva*: $(\forall x \in \mathbb{R}^n)\ Q(x) \geq 0$.
- *Definida negativa*: $(\forall x \neq \overrightarrow{0})\ Q(x) < 0$.
- *Semidefinida negativa*: $(\forall x \in \mathbb{R}^n)\ Q(x) \leq 0$.
- *No definida*: $(\exists x_1, x_2 \in \mathbb{R}^n)\ Q(x_1) > 0 \wedge Q(x_2) < 0$.

Recordar: Para hacer demostraciones afirmando este tipo de expresiones, es aconsejable siempre evaluar $x^T A x$ y hacer la demostración.

Para matrices D diagonales, quiere decir que sólo apareceran términos del tipo x_{ii} en la forma cuadrática. Son los más fáciles de clasificar:

$$\begin{bmatrix} d_{11} & & \mathbb{O} \\ & \ddots & \\ \mathbb{O} & & d_{nn} \end{bmatrix}$$

- Q y su matriz D son definidas positivas si $d_{ii} > 0$.
- Q y su matriz D son semidefinidas positivas si $d_{ii} \geq 0$.

Ocurre de forma análoga para definido negativo y semidefinido negativo.

Teorema: Sea $Q(x) = x^T S x$ una f.c. y sea S una matriz simétrica tal que $S = LDL^T$ tal que $D = (d_{ij})_{n \times n}$. Entonces:

- Q y su matriz S son definidas positivas si $d_{ii} > 0$.
- Q y su matriz S son semidefinidas positivas si $d_{ii} \geq 0$.
- Q y su matriz S son definidas negativas si $d_{ii} < 0$.
- Q y su matriz S son semidefinidas positivas si $d_{ii} \geq 0$.
- En caso contrario, son no definidas.

Definición: Se define como *diagonalización de una matriz cuadrada* al proceso por en el que una matriz S simétrica se lleva a su forma LDL^T, de modo que $Q(x)$ se escribe como la exclusiva suma de términos al cuadrado. Notar que:

$$Q(x) = x^T S x = \underbrace{x^T L}_{u^T} D \underbrace{L^T x}_{u}$$

4.6.3. Segunda factorización de Cholesky

IMPORTANTE: Las siguientes demostraciones también se extienden a definidas negativas.

Definición: Sea $A = (a_{ij})_{n \times n}$, definimos se define como submatriz A_k con $k = \{1, \ldots, n\}$ de la forma $A_k = (a_{ij})_{k \times k}$ donde el elemento (i, j) de A_k concide con el mismo elemento de A.

Teoremas:

1. Si una matriz A es definida positiva, entonces A es invertible.
2. Si A es definida positiva, entonces A_k es definida positiva para todo $k < n$.
3. Si A es definida positiva, entonces A tiene factorización LU.
4. Una matriz cuadrada A admite descomposición $A = LU$ si sus $n-1$ submatrices principales son invertibles, pues así se garantiza que no se deben efectuar operaciones de permutación.

Corolario: Todas las submatrices de una matriz definida positiva son invertibles.

Método de factorización: Sea A definida positiva, tenemos que $A = LDL^T$. Se sigue que:

$$D = \begin{bmatrix} d_{11} & & \mathbb{O} \\ & \ddots & \\ \mathbb{O} & & d_{nn} \end{bmatrix}$$

Se tiene que $(\forall i \leq n)\, d_{ii} > 0$. Por lo tanto,

$$\sqrt{D} = \sqrt{D}^T = \begin{bmatrix} \sqrt{d_{11}} & & \mathbb{O} \\ & \ddots & \\ \mathbb{O} & & \sqrt{d_{nn}} \end{bmatrix}$$

Luego $A = L\sqrt{D}^T\sqrt{D}L^T = \underbrace{(\sqrt{D}L^T)^T}_{R^T}\underbrace{\sqrt{D}L^T}_{R}$.

5. Determinantes

Definición: Un determinante es una función: $\det : M_n(\mathbb{R}) \longrightarrow \mathbb{R}$ que tiene la particularidad de entregar 0 para una matriz no invertible. Se denota $\det(A)$ ó $|A|$[a]. Es la única función que satisface las siguientes tres propiedades fundamentales:

1. $\det(\mathbb{I}) = 1$.
2. *Propiedad alternante*: Siendo $\widehat{A}$ una matriz obtenida al intercambiar dos filas de A, entonces $\det(\widehat{A}) = -\det(A)$.
3. *Propiedad multilineal*:

$$\begin{vmatrix} f_1^T \\ \vdots \\ f_{i-1}^T \\ \alpha u^T + \beta v^T \\ f_{i+1}^T \\ \vdots \\ f_n^T \end{vmatrix} = \alpha \begin{vmatrix} f_1^T \\ \vdots \\ f_{i-1}^T \\ u^T \\ f_{i+1}^T \\ \vdots \\ f_n^T \end{vmatrix} + \beta \begin{vmatrix} f_1^T \\ \vdots \\ f_{i-1}^T \\ v^T \\ f_{i+1}^T \\ \vdots \\ f_n^T \end{vmatrix}$$

[a] ¡No confundir con valor absoluto!

5.1. Propiedades básicas

- $\det(PA) = (-1)^r \det(A)$. Donde P es una matriz producto de r operaciones elementales de permutación.
- Si A tiene dos filas iguales, entonces $\det(A) = 0$. De forma análoga si A tiene una fila con ceros.
- Respecto a matrices elementales:
  - $\det(P_{ij}) = -1$
  - $\det(E_i(\alpha)) = \alpha$
  - $\det(E_{ij}(-c)) = 1$
- $det(MA) = det(M) \cdot det(A)$, donde M es una matriz elemental.
- $det(\alpha A_{n\times n}) = \alpha^n |A|$.
- Sea D matriz diagonal, $det(D) = \displaystyle\prod_{i=1}^{n} d_{ii}$, para todo $d_{ii} \in \mathbb{R}$ y $i \leq n$.
- Si $PA_{n\times n} = LU$, entonces $det(A) = (-1)^r det(U)$.
- Si A es triangular, $det(A) = \displaystyle\prod_{i=1}^{n} l_{ii}$.
- A es invertible $\iff det(A) \neq 0$.
- $det(AB) = det(A) \cdot det(B)$.
- $det(A^{-1}) = \dfrac{1}{det(A)}$.
- $det(A^T) = det(A)$.

5.2. Cálculo de determinantes

Recursión

- $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$
- $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = a_1 \begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix} - a_2 \begin{vmatrix} b_1 & b_3 \\ c_1 & c_3 \end{vmatrix} + a_3 \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}.$

Definición: Para $A_{n\times n}$ se define A_{ij} como la matriz de $(n-1)\times(n-1)$ que se obtiene al eliminar la fila i-ésima y la columna j-ésima.

- Por la i-ésima fila: $det(A_{n\times n}) = \displaystyle\sum_{j=1}^{n}(-1)^{i+j}a_{ij}det(A_{ij})$. De preferencia se desarrolla por la fila 1, es decir, $i=1$.
- Por la j-ésima columna: $det(A_{n\times n}) = \displaystyle\sum_{i=1}^{n}(-1)^{i+j}a_{ij}det(A_{ij})$.

5.3. Clasificación de formas cuadráticas con determinantes

Teorema: Sea A una matriz simétrica y A_i con $i = 1, \ldots n$ las submatrices principales de A, se tiene que:

1. A es definida positiva si y sólo si $det(A_i) > 0\,\forall i$.
2. A es definida negativa si y sólo si $(-1)^i det(A_i) > 0\,\forall i$.

5.4. Cofactores y matriz adjunta

Definición: Se define el cofactor del elemento a_{ij}como:

$$Co(a_{ij}) = (-1)^{i+j}det(A_{ij})$$

Entonces, $det(A) = \displaystyle\sum_{j=1}^{n} a_{ij}Co(a_{ij}) = \sum_{i=1}^{n} a_{ij}Co(a_{ij})$.

Definición: Se define la matriz adjunta de A como:

$$Adj(A) = (Co(a_{ij}))^T_{n\times n}$$

Teorema: Sea $A_{n\times n}$, se tiene que:

$$A \cdot Adj(A) = Adj(A) \cdot A = det(A) \cdot \mathbb{I}_n$$

Consecuencias: Del teorema anterior se sigue que:

- A es invertible $\iff det(A) \neq 0$. Entonces, $\underbrace{\frac{1}{det(A)} Adj(A)}_{¡A^{-1}!} \cdot A = \mathbb{I}_{n \times n}$. $A^{-1} = \frac{1}{det(A)} Adj(A)$.
- $Adj(A) = det(A) \cdot A^{-1}$, entonces si A es invertible: $det(Adj(A)) = det(A)^{n-1}$.
- $Adj(A) \cdot Adj(Ajd(A)) = det(Adj(A)) \cdot \mathbb{I}_n$. Entonces, $det(Adj(Adj(A))) = det(A)^{(n-1)^2}$.
- $Adj(A^T) = Adj(A)^T$.
- $Adj(Adj(A)) = det(A)^{n-2} A$.
- $Adj(\mathbb{I}) = \mathbb{I}$.

6. Espacios vectoriales

6.1. Cuerpos finitos

Definición: Un cuerpo $\mathbb{K}$ es un conjunto de elementos que satisfacen los siguientes axiomas básicos para la suma y el producto:
Suma: Operación cerrada: $\mathbb{K} \times \mathbb{K} \longrightarrow \mathbb{K}$.

1. *Conmutatividad*: $x + y = y + x$.
2. *Asociatividad*: $x + (y + z) = (x + y) + z$.
3. *Existencia de neutro aditivo*: Existe un 0_k tal que $x + 0_k = x$.
4. *Existencia de inverso aditivo*: de modo que el elemento sumado a su inverso den el elemento neutro.
$$x + (-x) = 0_k$$

Producto: Operación cerrada: $\mathbb{K} \times \mathbb{K} \longrightarrow \mathbb{K}$.

1. *Conmutatividad*: $x \cdot y = y \cdot x$.
2. *Asociatividad*: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$.
3. *Existencia de inverso aditivo*: Existe un 1_k tal que $x \cdot 1_k = x$.
4. *Existencia de un inverso aditivo*: Para todo $x \neq 0_k$ existe un x^{-1} tal que $x \cdot x^{-1} = 1$.

Distributividad: $x \cdot (y + z) = x \cdot y + x \cdot z$.

Teorema: $\mathbb{Z}_n$ es un cuerpo si y sólo si $\mathbb{Z}_n = \mathbb{Z}_p$ con p primo.

Definición: Para $n \in \mathbb{Z}$ se define:

$$n \bmod p = i \iff n = k \cdot p + i, i \in \{1, \ldots, p-1\}$$

Observaciones:

- $m \oplus n = (m + n) \bmod p$.

- $m \otimes n = (m+n) \operatorname{mód} p$.
- Para las operaciones de suma y multiplicación basta realizar la operatoria real y luego aplicar módulo.
- $\dfrac{a}{b} = x \iff a = b \cdot x$. Esta forma optimiza el cálculo de fracciones. Sólo es válido si $b \operatorname{mód} p \neq 0$.
- Se pueden hacer todas las operaciones en $\mathbb{R}$ y luego convertir a $\mathbb{Z}_p$ cuidando de no efectuar una división por un múltiplo de p.
- $\#\mathbb{Z}_p^n = p^n$.
- $v_1 \in \mathbb{Z}_p^n \Rightarrow \# < v_1 >= p$.

Proposiciones:

1. $m \operatorname{mód} p + n \operatorname{mód} p = m \oplus n$.
2. $m \operatorname{mód} p \times n \operatorname{mód} p = m \otimes n$.

$\oplus$:

1. *Asociatividad:* $(x \oplus y) \oplus z = x \oplus (y \oplus z)$.
2. *Conmutatividad:* $x \oplus y = y \oplus x$.
3. *Inverso aditivo:* $x \oplus (p - x) = 0$.

$\otimes$:

1. *Asociatividad:* $(x \otimes y) \otimes z = x \otimes (y \otimes z)$.
2. *Conmutatividad:* $x \otimes y = y \otimes x$.

Propiedad distributiva:

- $(x \oplus y) \otimes z = (x \otimes z) \oplus (y \otimes z)$.

6.2. Concepto de espacio vectorial

Definición: Un conjunto no vacío $\mathbb{E}$ (con sus vectores) es un espacio vectorial (e.v.) sobre un cuerpo $\mathbb{K}$ (con sus escalares) si tiene definidas dos operaciones binarias y cerradas:

- Suma $\oplus$: $\mathbb{E} \times \mathbb{E} \longrightarrow \mathbb{E}$.
- Ponderación $\odot$: $\mathbb{E} \times \mathbb{K} \longrightarrow \mathbb{E}$.

que cumplen los siguientes axiomas:
Axiomas de la Suma:

1. *Conmutatividad.* $(\forall x, y \in \mathbb{E})\, x + y = y + x$
2. *Asociatividad.* $(\forall x, y, z \in \mathbb{E})\, x + (y + z) = (x + y) + z$.
3. Existencia de *elemento neutro aditivo* 0_E. $0_E :=$ neutro aditivo de $\mathbb{E}$.
4. A cada $x \in \mathbb{E}$ le corresponde su *inverso aditivo.* $(\forall x \in \mathbb{E})(\exists \overline{x} \in \mathbb{E})\, x \oplus (-x) = 0_E$.

Axiomas de la Ponderación:

1. Existencia de un *neutro multiplicativo* 1_K: $(\forall x \in \mathbb{E})\, x \odot 1_K = x$.
2. *Asociatividad.* $(\forall \alpha, \beta \in \mathbb{K})(\forall x \in \mathbb{E})\, (\alpha\beta) \odot x = \alpha \odot (\beta \odot x) = \beta \odot (\alpha \odot x)$.
3. *Distributiva respecto* a la suma escalar.
4. *Distributiva respecto* a la suma vectorial.

Ejemplos: Los siguientes ejemplos se asumen inmediatamente como espacios vectoriales:

- $\mathbb{E} = \mathbb{R}^n$.
- $\mathbb{E} = \mathbb{M}_{m \times n}(\mathbb{R})$. Matrices de $m \times n$.
- $\mathbb{E} = \mathcal{F}(\mathbb{R})$. Todas las funciones reales $f : \mathbb{R} \longrightarrow \mathbb{R}$.
- $\mathbb{E} = C^1(\mathbb{R})$.
- $\mathbb{E} = \mathcal{P}(\mathbb{R})$. Conjunto de todos los polinomios con coeficientes reales.
- $\mathbb{E} = \mathcal{P}_n(\mathbb{R})$. Conjunto de todos los polinomios con grado menor o igual a 3. **Ojo:** Un e.v. con todos los polinomios de grado exactamente igual a 3 no es un cuerpo.

Se puede sustituir $\mathbb{R}$ por cualquier otro cuerpo $\mathbb{K}$ y obtener resultados.

Teorema: Sea $\mathbb{E}$ un e.v. sobre un cuerpo $\mathbb{K}$, entonces:

- El elemento neutro aditivo 0_E es único.
- Para cada elemento $x \in \mathbb{E}$, su inverso aditivo $(-x)$ es único.
- $0_K \odot x = 0_E$.
- $\alpha \odot 0_E = 0_E$.

- $(-1) \odot x = (-x)$.
- $\alpha \odot x = 0 \Longrightarrow \alpha = 0 \vee x = 0_E$.
- Existe un único vector x que es solución para $u \oplus x = v$.

6.3. Combinaciones lineales

Definición: Sea $\mathbb{E}$ un e.v. sobre $\mathbb{K}$ y $S = \{v_1, \dots, v_n\}$ un conjuunto donde todo $v_i \in \mathbb{E}$ entonces:

1. Una *combinación lineal* (c.l.) es una expresión de la forma:
$$\alpha_1 v_1 + \ldots + \alpha_n v_n \qquad \alpha_i \in \mathbb{K},\, 1 \leq i \leq n$$
2. Se define el *conjunto generado* por S como el conjunto de todas las c.l. de los vectores de S.
$$\langle S \rangle = \{x \,:\, x \,\text{es c.l. de}\, S\}$$
3. El conjunto S es *linealmente independiente* (l.i.) si la única forma de escribir 0_E como c.l. de vectores de S es la trivial. Es decir,
$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0_E \Longleftrightarrow (\forall i)\, \alpha_i = 0$$
4. El conjunto S es *linealmente dependiente* (l.d.) si existen α_i no todos nulos de modo que:
$$\alpha_1 v_1 + \ldots + \alpha_n v_n = 0_E$$

En general:

- $A \subseteq \mathbb{E}$ es l.i. si cualquier subconjunto finito A_i es l.i.
- $A \subseteq \mathbb{E}$ es l.d. si algún subconjunto finito A_i es l.d.

7. Subespacios vectoriales

7.1. Definición

Sea $\mathbb{E}$ un e.v. sobre $\mathbb{K}$, S es un *subespacio vectorial* (s.e.v.) si:

- $S \subseteq \mathbb{E}$ y $S \neq \varnothing$.
- S verifica los axiomas para las operaciones heredadas de $\mathbb{E}$.

Notación: $S \leqslant \mathbb{E}$.

Teorema: Sea S un subconjunto no vacío de $\mathbb{E}$, e.v. sobre $\mathbb{K}$. S es s.e.v. si y sólo si se cumple que:

1. S es cerrado bajo la suma: $x, y \in S \Longrightarrow x + y \in S$.
2. S es cerrado bajo la ponderación por escalar: $x \in S$, $\alpha \in \mathbb{K} \Longrightarrow \alpha \odot x \in S$.

Teorema: S es un s.e.v. de $\mathbb{E}$ si existe algún conjunto A tal que $S = \langle A \rangle$.

7.1.1. Para demostrar s.e.v.

Con los dos teoremas anteriores se tienen dos métodos para demostrar que S es un s.e.v. de $\mathbb{E}$:

Método 1: Para demostrar que S es un s.e.v. de $\mathbb{E}$, basta demostrar que:[1]

1. $S \subseteq \mathbb{E}$ y $S \neq \varnothing$. En particular, demostrar que $0_E \in S$.
2. S es cerrado bajo suma.
3. S es cerrado bajo ponderación.

Método 2: Expresar como conjunto generado. Evidentemente, algo no del estilo $\langle \varnothing \rangle$, pero por ejemplo $\langle 0_E \rangle$ sí es un s.e.v.

7.2. Operaciones con s.e.v.

No olvidar:

1. $x \in A \cap B \Longleftrightarrow x \in A \wedge x \in B$.
2. $x \in A \cup B \Longleftrightarrow x \in A \vee x \in B$.
3. $A \cap B = \{0_E\} \Longleftrightarrow \alpha_1 a_1 + ... + \alpha_n a_n = \beta_1 b_1 + \ldots + \beta_n b_n$ tiene solución trivial.

Teorema: Si S_1 y S_2 son s.e.v. de $\mathbb{E}$, entonces $S_1 \cap S_2$ también es s.e.v. de $\mathbb{E}$.

Importante: $S_1 \cup S_2 \subseteq \mathbb{E}$, pero no necesariamente es s.e.v., pues puede no ser cerrado bajo la suma o la multiplicación.

Teorema: Sean S_1 y S_2 s.e.v. sobre $\mathbb{E}$, entonces $S_1 \cup S_2$ es s.e.v. de $\mathbb{E}$ si y sólo si $S_1 \subseteq S_2$ ó $S_2 \subseteq S_1$.

Demostración: [2] Recordar que para demostrar una equivalencia (si y sólo si) se debe demostrar la implicancia para ambos lados, es decir:

$$(p \Leftrightarrow q) \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$$

$(\Longleftarrow)$

Lo haremos para $S_1 \subseteq S_2$. La otra demostración es análoga.

Observaciones:

1. $A \subseteq A \cup B$ y $B \subseteq A \cup B$.
2. $S_1 \subseteq S_2 \Longleftrightarrow (x \in S_1 \Longrightarrow x \in S_2)$

[1] Y en estos casos, siempre revisar la definición que se nos da de $\mathbb{E}$

[2] Se dejó propuesta en clases. Aquí está desarrollada, como ejercicio.

1) Cerrado bajo la suma.

$x \in S_1, y \in S_2$. Se sigue que $x \in S_2$. Luego, es evidente que $x + y \in S_1 \cup S_2$, pues $x + y \in S_2 \subseteq S_1$.

2) Cerrado bajo la ponderación.

Como S_1 es s.e.v., entonces si $x \in S_1$ se tiene que$\alpha x \in S_1 \subseteq S_1 \cup S_2$. De forma análoga ocurre para algún $y \in S_2$.

Por lo tanto, siendo $S_1 \subseteq S_2$ ó $S_2 \subseteq S_1$, se tiene que $S_1 \subseteq S_2$ ó $S_2 \subseteq S_1$ es s.e.v. de $\mathbb{E}$.

$(\Longrightarrow)$

En el sentido de la lógica, tenemos que:

$$\underbrace{p \Longrightarrow q \vee r}_{\text{por demostrar}} \quad \equiv \quad \underbrace{(p \wedge \overline{q}) \Longrightarrow r}_{\text{expr. equivalente}}$$

Entonces, supongamos que $S_1 \cup S_2$ es s.e.v. y que $S_1 \not\subseteq S_2$. Tenemos que demostrar que $S_2 \subseteq S_1$.

Sea $s_1 \in S_1 \Rightarrow s_1 \in S_1 \cup S_2$ y $s_2 \in S_2 \Rightarrow s_2 \in S_1 \cup S_2$, como es cerrado bajo la suma, entonces $s_1 + s_2 \in S_1 \cup S_2$.

Pero, $s_1 + s_2 \in S_1 \cup S_2 \Longrightarrow s_1 + s_2 \in S_1 \vee s_1 + s_2 \in S_2$. Pero, como S_1 es s.e.v. por hipótesis, también es cerrado bajo la suma. Luego, si restamos $s_1 \in S_1$, debe cumplirse que:

$$s_1 + s_2 - s_1 = s_2 \in S_1$$

Hemos demostrado que $s_2 \in S_2 \Longrightarrow s_2 \in S_1 \equiv S_2 \subseteq S_1$. Por lo tanto, se cumple también esto.

Como se han verificado ambas implicancias, queda entonces demostrado.■

Definición: Sean S_1 y S_2 dos s.e.v. de $\mathbb{E}$, se define:

$$S_1 + S_2 = \{x \in \mathbb{E} \, : \, x = s_1 + s_2 \,\text{con}\, s_1 \in S_1 \,\text{y}\, s_2 \in S_2\}$$

Teorema: Si S_1 y S_2 son s.e.v. de $\mathbb{E}$, entonces $S_1 + S_2$ es s.e.v. de $\mathbb{E}$.

Teorema: Si $S_1 = \langle A \rangle$ y $S_2 = \langle B \rangle$, entonces $S_1 + S_2 = \langle A \cup B \rangle$.

Definición: Sean S_1 y S_2 subespacios vectoriales de $\mathbb{E}$, entonces $S_1 + S_2$ es suma directa si:

i) $S_1 + S_2 = \mathbb{E}$.

ii) $S_1 \cap S_2 = \{0_E\}$.

Se anota entonces $S_1 \oplus S_2 = \mathbb{E}$. Notar que: $\langle v_1, \dots, v_n \rangle = \langle v_1, \dots, v_{i-1} \rangle \oplus \langle v_i, \dots, v_n \rangle$.

Nota: *Razonamiento para la intersección de s.e.v.* Para probar que $U \cap V = \{0_E\}$, tomamos dos bases:

$B_U = \{u_1, \dots, u_m\}$ y $B_V = \{v_1, \dots, v_n\}$. Luego, por definición de intersección:

$U \cap V = \{x \,:\, x = \alpha_1 u_1 + \ldots + \alpha_m u_m \,\wedge\, x = \beta_1 v_1 + \ldots + \beta_n v_n\}$.

Por lo que:

$$\alpha_1 u_1 + \ldots + \alpha_m u_m = \beta_1 v_1 + \ldots + \beta_n v_n$$

Es decir,

$$\sum_{i=1}^{n} \alpha_i u_i - \sum_{k=1}^{m} \beta_k v_k = 0_E$$

O, en otras palabras, demostrar que la única forma de generar 0_E es la trivial: $\alpha_i = 0$ y $\beta_i = 0$, porque de esta forma, el único elemento que satisfacería esta condición sería exactamente 0_E. Es decir:

$$\underbrace{\sum_{i=1}^{n} \alpha_i u_i}_{\text{vector de } U} = 0_E, \text{ pues } \alpha_i = 0$$

$$\underbrace{\sum_{i=1}^{m} \beta_i v_i}_{\text{vector de } V} = 0_E, \text{ pues } \beta_i = 0$$

8. Bases y dimensión

8.1. Bases

Definición: Sea $\mathbb{E}$ un e.v. sobre $\mathbb{K}$ y S un s.e.v. de $\mathbb{E}$. Un conjunto B es una base de S si:

1. $B \subseteq \mathbb{E}$.
2. B es l.i.
3. $S = \langle B \rangle$.

Notar que $\{0_E\}$ es el único conjunto que es s.e.v. de todo e.v. y a la vez no tiene base, pues $\{0_E\} = \langle 0_E \rangle$, que no es un conjunto l.i.

8.2. Dimensión

Teorema: Sea $S \leqslant \mathbb{E}$ sobre $\mathbb{K}$ cuya base tiene n vectores. Todo conjunto de vectores l.i. de S tiene a lo más n vectores.

Teorema: Sea $S \leqslant \mathbb{E}$ sobre $\mathbb{K}$ cuya base tiene n vectores. Toda base de S tiene exactamente n vectores.

Definición: Sea S un s.e.v. de $\mathbb{E}$, entonces se define la dimensión de S como el número de vectores que tiene cualquier base de S. Se anota $\dim S$, y se tiene que:

- Si S tiene n elementos[3], entonces:
$$\dim S = n$$
- Si $S = \{0_E\}$, entonces:
$$\dim S = 0$$
- Si S tiene infinitos elementos, entonces:
$$\dim S = \infty$$

Teorema: Sea $S \leqslant \mathbb{E}$ sobre $\mathbb{K}$ con $\dim S = n$. Si $A \subseteq S$ tiene n vectores, entonces:

1. A es base de S si y sólo si A es l.i.
2. A es base de S si y sólo si $S = \langle A \rangle$.

Notar que una implica a la otra.

8.3. Completación de bases

Teorema: Sea $S \leqslant \mathbb{E}$ sobre $\mathbb{K}$. Sea A un conjunto l.i. de vectores en S, entonces A está contenido en alguna base de S.

Teorema: Sea $S_1 \leqslant \mathbb{E}$ y $S_2 \leqslant \mathbb{E}$, ambos con dimensión finita. Entonces:

$$\dim(S_1 + S_2) = \dim(S_1) + \dim(S_2) - \dim(S_1 \cap S_2)$$

Corolario: Si $\mathbb{E} = S_1 \oplus S_2$, entonces $\dim \mathbb{E} = \dim S_1 + \dim S_2$.

9. Sistema de coordenadas y cambio de base

9.1. Sistemas de coordenadas

Observación: En una base el orden de los vectores sí importa. Se habla de *bases ordenadas* de $\mathbb{E}$.

Teorema: Sea $B = \{b_1, \ldots, b_2\}$ una base de $\mathbb{E}$. Entonces existe una única c.l. de $x \in \mathbb{E}$ con esa base.

Definición: Sea $B = \{b_1, \ldots, b_n\}$ una base de $\mathbb{E}$. Para $x \in \mathbb{E}$ se define su *vector coordenada* como

$$[x]_B = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n \iff x = \alpha_1 b_1 + \ldots + \alpha_n b_n$$

De preferencia se puede entender que $b_1, \ldots, b_n$ están en términos de la base canónica.

$(*)$ Ejemplo de base canónica: $\mathbb{R}^3 = \left\langle \underbrace{e_1, \ldots, e_3}_{\text{base canónica}} \right\rangle$. ¡El conjunto es la base, y no el generado!

Teoremas: Sea $B = \{b_1, \ldots, b_n\}$ una base de $\mathbb{E}$ y x, y vectores de $\mathbb{E}$. Entonces:

1. El vector $[x]_B$ es único. $[x]_B$ es un biyección.
2. $[x + y]_B = [x_1]_B + [x_2]_B$.
3. $[\alpha x]_B = \alpha[x]_B$.
4. $\{x_1, \ldots, x_n\}$ es l.i. $\Longleftrightarrow \{[x_1]_B, \ldots, [x_n]_B\} \in \mathbb{K}^n$.

9.2. Matrices de cambio de bases

Definición: Sea $S \leqslant \mathbb{E}$ sobre $\mathbb{K}$. $[Id]_{B_1}^{B_2}$ es la única matriz de cambio de base que cambia las coordenadas en base B_1 a coordenadas en base B_2 para un vector $t \in S$ y B_1y B_2 bases de S. Esto implica que:

$$(\forall x \in S) \quad \Big[x\Big]_{B_2} = \Big[Id\Big]_{B_1}^{B_2} \Big[x\Big]_{B_1}$$

¿Cómo determinarla? (Razonamiento): Verlo de la forma lo más gráfica posible. Pensar primero en las bases:

- $B_1 = \{x_1, \ldots, x_n\}$
- $B_2 = \{y_1, \ldots, y_n\}$

Considerar que B_1 y B_2 son conjuntos l.i. Es decir, cada elemento de cada conjunto se puede expresar como combinación lineal del otro. En particular, cada x_i se puede expresar en términos de $\{y_1, \ldots, y_n\}$.

Siempre pensar que x_i e y_k son vectores, pero no de la forma $\mathbb{K}^n$, si no que vectores propios de su conjunto. Es decir, pueden ser matrices, polinomios, etc.

Considerar $[Id]_{B_1}^{B_2}$ por columnas. Es decir, $[Id]_{B_1}^{B_2} = [c_1|\ldots|c_2]$. Se sabe además que los vectores coordenadas son especificamente de la forma $\mathbb{K}^n$.

¿Cómo se obtiene la columna 1? Por operación de matrices se sabe que es $[Id]_{B_1}^{B_2} \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}}_{(*)}$. Pero, ¿qué es $(*)$? Es un vector coordenada de B_1. En particular, es tomar $[x_1|\ldots|x_n] \begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} = x_1$. La pregunta que falta es, ¿a qué se puede igualar esto? La idea detrás de la matriz de cambio de base es convertir un vector expresado en ciertas coordenadas a otra base. Por lo tanto, se debe buscar $\alpha_1, ..., \alpha_n$ de modo que $x_1 = \alpha_1 y_1 + \ldots + \alpha_n y_n$. De esta forma, se tiene que:

$$[\alpha_1 y_1 + \ldots + \alpha_n y_n]_{B_2} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = [Id]_{B_1}^{B_2} \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}}_{[x_1]_{B_1}} = c_1$$

$$\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = c_1$$

¿Cómo se obtienen las demás columnas? ¡De la misma forma! Por lo tanto, el proceso de determinación de la matriz de cambio de bases se puede resumir en lo siguiente:

Algoritmo: Considerar una base canónica B_e, tomar la siguiente matriz

$$\Big[[y_1]_{B_e}|\dots|[y_n]_{B_e}|[x_1]_{B_e}|...|[x_n]_{B_e}\Big]$$

y pivotearla. Es evidente, dado que ambos conjuntos son l.i., que los pivotes aparecerán en y_1. Incluso, como es base, y es un conjunto l.i., se pivoteará hasta llegar a $\mathbb{I}$. Luego, se puede resolver cada sistema "$[B_2]x_i$" individualmente. Sea u_i cada solución, se resume que:

$$FER\Bigg(\Big[[y_1]_{B_e}|\dots|[y_n]_{B_e}|[x_1]_{B_e}|...|[x_n]_{B_e}\Big]\Bigg) = [\mathbb{I}_n|\underbrace{u_1|\dots|u_n}_{[Id]_{B_1}^{B_2}}]$$

$$[Id]_{B_1}^{B_2} = [u_1|\dots|u_n]$$

El método general es resolver cada sistema, pero para operaciones no tan evidentes, este método puede simplificar mucho la tarea. Sin embargo, este método es válido, pues se sustenta en el hecho de que el cambio de coordenadas es una operación lineal.

Observación: La matriz de cambio de base $[Id]_{B_1}^{B_2}$ es:

1. Cuadrada, pues son n filas para n vectores de la base y n columnas para tener congruencia en el producto de matriz por vector.
2. Con columnas y filas l.i., pues la multiplicación por la matriz debe ser una operación biyectiva (a cada vector coordenada en B_1 le corresponde un único vector coordenada en B_2 y viceversa).
3. Por lo tanto, invertible y $\left([Id]_{B_1}^{B_2}\right)^{-1} = [Id]_{B_2}^{B_1}$. Algo esperable, pues la operación inversa a un cambio de base es regresar a la base original.

10. Transformaciones lineales

10.1. Conceptos básicos

Definición: Sean U y V e.v. sobre un mismo cuerpo $\mathbb{K}$. Una función $T : U \longrightarrow V$ es una transformación lineal (t.l.) si y sólo si:

1. $T(u_1 + u_2) = T(u_1) + T(u_2)$ con $u_1, u_2 \in U$.
2. $T(\alpha u_1) = \alpha T(u_1)$.

En ambos casos, la operación se hace primero en U y en el segundo lado de la igualdad en V.

Propiedades: Considerando nuevamente U y V.

1. $T(0_U) = 0_W$.
2. Siendo $A \subseteq U$, entonces $T(\langle A \rangle) = \langle T(A) \rangle$.
3. S es s.e.v. de $U \implies T(S)$ es s.e.v. de V. (Como cada s.e.v. se puede expresar como conjunto generado, es consecuencia de 2)

Definición: Sea $T : U \longrightarrow V$ una t.l. Se define:

1. el *kernel o núcleo* como $\mathrm{Ker}(T) = \{u \in U : T(u) = 0_V\}$.
2. la *imagen* como $\mathrm{Im}(T) = \{v \in V : \exists u \in U : T(u) = v\}$.

Teorema: Sea $T : U \longrightarrow V$ una t.l., entonces:

1. $\mathrm{Ker}(T)$ es un s.e.v. de U.
2. $\mathrm{Im}(T)$ es un s.e.v. de V.

Teorema: *(del Núcleo-Imagen)* Sean U y V e.v. sobre $\mathbb{K}$ y de dimensión finita y $T : U \longrightarrow V$ una t.l. Entonces:

$$\dim U = \dim(\mathrm{Ker}\ T) + \dim(\mathrm{Im}\ T)$$

Este teorema es clave para la determinación de núcleo e imágenes de transformaciones lineales, pues indica fácilmente la dimensión y el número de vectores de un conjunto generado conociendo el otro.

Para el caso de matrices, simplemente se puede resumir en:

$$\text{número de columnas} = \underbrace{\text{número de columas no l.i.}}_{\text{cantidad de vectores del Ker}} + \underbrace{\text{número de columnas l.i}}_{\text{N° de vectores que generan Im(A)}}$$

10.2. Isomorfismos

Definición: Sea $T : U \longrightarrow V$ una t.l. Entonces T es:

1. un *monomorfismo* si T es 1-1 o inyectiva.
2. un *epiformismo* si T es sobre o epiyectiva.
3. un *isomorfismo* si T es 1-1 y sobre, es decir, una biyección.

Definición: Dos espacios U y V son isomorfos ($U \cong V$) si existe una transformación T que sea isomórfica.

Teoremas: Sea $T : U \longrightarrow V$ una t.l. Entonces:

1. T es monomorfismo si y sólo si $\dim(\mathrm{Ker}(T)) = \{0\}$.

2. T es epimorfismo si y sólo si $\dim(\mathrm{Im}(T)) = \dim(V)$.

3. T es monomorfismo si y sólo si transforma conjuntos l.i. de U en conjuntos l.i. de V.

4. T es epimorfismo si y sólo si transforma conjuntos generadores de U en conjuntos generadores de V.

5. T es isomorfismo si y solo si transforma bases de U en bases de V.

Teorema: $U \cong V \iff \dim U = \dim V$.

10.3. Matriz de una transformación lineal

Definición: Sean U y V espacios vectoriales sobre $\mathbb{K}$ con dimensiones n y m respectivamente y sus respectivas bases B_U y B_V. Sea además $T : U \longrightarrow V$ una t.l., entonces se define la matriz $A_{m\times n}$ de dicha transformación lineal como aquella que cumple que para $u \in U$:

$$\Big[T(u)\Big]_{B_V} = A_{m\times n}[u]_{B_U}$$

¿Cómo determinarla?

1. Decidir o considerar las bases pertinentes. Para el caso de la definición se puede tomar:

$$B_U = \{u_1, \dots, u_n\}$$

$$B_V = \{v_1, \dots, v_m\}$$

2. Considerar A por columnas:

$$A = [a_1| \dots |a_n]$$

3. Se quiere sacar cada columna de A, luego, lo prudente sería ir multiplicándola por cada "e_i" correspondiente, como se hacía anteriormente. Recordar aquí que:

$$A\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix} = A[u_1]_{B_U} = a_1$$

4. A priori se conoce la transformación lineal. Por lo tanto, se conoce $T(u_i)$. Conociendo la base, es también posible determinar $\Big[T(u_i)\Big]_{B_V}$.

5. ¿Qué falta por hacer? ¡Igualar!

$$\Big[T(u_i)\Big]_{B_V} = a_i \longleftarrow \text{columna } i\text{-ésima de } A$$

Concluimos bajo este razonamiento que:

$$\boxed{\text{A}=\left[\Big[\text{T}(u_1)\Big]_{B_V}\Big|\dots\Big|\Big[\text{T}(u_n)\Big]_{B_V}\right]}$$

Definición: La matriz A se llama formalmente matriz de la transformación T de la base B_U a la base B_V. Se anota:

$$\Big[T\Big]_{B_U}^{B_V}$$

Teoremas: Sean $T : U \longrightarrow V$, $B_U = \{u_1, \dots, u_n\}$ y $B_V = \{v_1, \dots, v_n\}$ las bases respectivas de los e.v. U y V y $[T]_{B_U}^{B_V}$ la matriz de la transformación T en las bases respectivas. Entonces, los siguientes teoremas muestran un comportamiento similar de las t.l. al comportamiento ya conocido de matrices:

1. $u \in \text{Ker}(T) \Longleftrightarrow [u]_{B_U}$ es solución del sistema $[T]_{B_U}^{B_V} x = 0$.
2. $\forall v \in \text{Im}(T)$ el sistema $[T]_{B_U}^{B_V} x = v$ es compatible.
3. T es un monomorfismo si y solo si $[T]_{B_U}^{B_V}$ tiene inversa por la izquierda (la matriz sería 1-1).
4. T es un epimorfismo si y solo si $[T]_{B_U}^{B_V}$ tiene inversa por la derecha (la matriz sería sobre).
5. T es un isomorfismo si y solo si $\left(\Big[T\Big]_{B_U}^{B_V}\right)^{-1}$ existe.

Corolario: $\Big[T^{-1}\Big]_{B_U}^{B_V} = \left(\Big[T\Big]_{B_U}^{B_V}\right)^{-1}$.

10.4. Álgebra de transformaciones lineales

Propiedades: Las propiedades de las matrices de transformación son las mismas que las de las matrices ya estudiadas:

1. $\Big[T_1 + T_2\Big]_{B_U}^{B_V} = \Big[T_1\Big]_{B_U}^{B_V} + \Big[T_2\Big]_{B_U}^{B_V}$.
2. $\Big[\alpha T_1\Big]_{B_U}^{B_V} = \alpha \Big[T_1\Big]_{B_U}^{B_V}$.
3. $\Big[T_3 \circ T_1\Big]_{B_U}^{B_W} = \Big[T_3\Big]_{B_V}^{B_W} \cdot \Big[T_1\Big]_{B_U}^{B_V}$.

Para todos los cálculos, el álgebra de matrices puede ser aplicada, por lo que se verifican las mismas propiedades para todas las transformaciones lineales.

Observación: Sean

- B_U^1 y B_U^2 bases de U.
- B_V^1 y B_V^2 bases de V.
- $T : U \longrightarrow V$ una t.l.
- $\Big[T\Big]_{B_U^1}^{B_V^1}$ y $\Big[T\Big]_{B_U^2}^{B_V^2}$ matrices de la transformación T con respecto a las bases indicadas.

Entonces:

$$\Big[T\Big]_{B_U^2}^{B_V^2} = \Big[Id\Big]_{B_V^1}^{B_V^2} \Big[T\Big]_{B_U^1}^{B_V^1} \Big[Id\Big]_{B_U^2}^{B_U^1}$$

10.5. Anexo: resolución de problemas tipo

Problema tipo 1: Sea $T : U \longrightarrow V$ una transformación lineal de modo que para $x \in U$ existe algún $y \in V$ tal que:

$$T(x) = y$$

(Es decir, se entrega la definición de la transformación lineal como se estime pertinente)

Determinar bases A y B para U y V respectivamente de modo que la matriz de la transformación sea

$$T = [c_1 | \dots | c_n] \quad c_i \in \mathbb{K}^m$$

(Se entrega la matriz completa, pero nosotros la consideramos por columnas)

Resolución: Este problema se resuelve considerando la siguiente secuencia:

1. Determinar base para el conjunto de partida.
2. Determinar base para el conjunto de llegada.

Considerar que la matriz es de $m \times n$. Entonces, buscamos bases que cumplan que:

- $A = \{a_1, \dots, a_n\}$
- $B = \{b_1, \dots, b_m\}$

Sabemos, por definición que:

- $c_1 = \Big[T(a_1)\Big]_B$
- $\vdots$
- $c_n = \Big[T(a_n)\Big]_B$

Si aparece una columna de la forma $c_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ quiere decir que $a_i \in \operatorname{Ker}(A)$.

En dichos casos, será necesario determinar los elementos del Ker de la matriz a partir de la definición entregada. En la posición i-ésima de la base deberá aparecer un elemento del núcleo de la matriz. Los demás elementos de la base se seleccionan de modo que el conjunto quede l.i. (lo más sencillo, por lo tanto, es elegir elemento de la base canónica o la base que simplifique lo más posible los cálculos).

La base del conjunto de partida debiera estar determinada. Ahora importa la base del conjunto de llegada. Como sabemos cada elemento a_i, podemos determinar sin complicación cada $T(a_i)$, pues sabemos la definición de la trasnformación que se nos entrega, además de la matriz. En caso contrario, el ejercicio no tendría solución.

Pensemos de nuevo en la base. Esta está determinada por

$$B = \{b_1, \ldots, b_m\}$$

Cada b_i es lo que se desea determinar. Sabemos que:

$$\Big[T(a_i)\Big]_B = c_i = \begin{pmatrix} \eta_{i1} \\ \vdots \\ \eta_{im} \end{pmatrix}$$

Por la definición, esto es:

$$T(a_i) = \eta_{i1}b_1 + \ldots + \eta_{im}b_m$$

Donde conocemos $T(a_i)$ y cada η_{ik}. Tomando una base canónica D, el problema se reduce a determinar una matriz $[b_1|\ldots|b_n]$ conociendo soluciones al sistema $\Big[\Big[b_1\Big]_D|\ldots|\Big[b_n\Big]_D\Big]c_i = \Big[T(a_i)\Big]_D$. El problema sería que pueden aparecer vectores que no son l.i., o que el conjunto quede con una dimensión menor que el conjunto de llegada. ¿Como se soluciona? Se completan bases, de modo que el conjunto quede l.i., y sólo los primeros términos representen a la imagen del conjunto.

11. Valores y vectores propios

11.1. Valores y vectores propios de una matriz cuadrada

Definición: Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Se define:

- El *valor propio* como aquel $\lambda \in \mathbb{R}$ para el que existe un $\overrightarrow{x} \in \mathbb{R}^n$ distinto de $\overrightarrow{0}$ tal que:

$$A\overrightarrow{x} = \lambda\overrightarrow{x}$$

- El *vector propio* es aquel $\overrightarrow{x}$ que satisface la condición del valor propio.

Observaciones:

1. Los vectores propios **no** pueden ser nulos.
2. $Ker(A)$ son los vectores propios para $\lambda = 0$, **si y solo si** es no trivial.
3. Un vector propio genera lo que se denomina una *dirección fija* en A, pues $Ax = \lambda x$, pero también αx es un vector propio para ese λ, pues

$$A(\alpha x) = \alpha Ax = \alpha\lambda x = \lambda(\alpha x)$$

Es decir, se tiene que una recta, o dirección fija, satisface la condición para ese valor propio.

Observación: La forma pertinente de calcular es primero el valor propio, y luego los vectores propios asociados a ese valor.

11.2. Cálculo de valores propios

Teorema: Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, entonces:

$$\lambda \text{ es valor propio de } A \Longleftrightarrow \det(A - \lambda\mathbb{I}_n) = 0$$

Definición: Para la matriz A $\det(A - \lambda\mathbb{I}_n)$ origina un polinomio que se denomina *polinomio característico* y se anota:

$$p_A(\lambda) = \det(A - \lambda\mathbb{I}_n)$$

Este polinomio es de grado n, y por Teorema Fundamental del Álgebra, se sabe que tiene n raíces complejas (no necesariamente distintas).

Definición: Sea λ_i raíz de $p_A(\lambda)$ (y por lo tanto, un valor propio de A). Entonces, se entiende por *multiplicidad algebraica* de λ_i al número de veces que λ_i es raíz de $p_A(\lambda)$. Es decir, el polinomio característico puede factorizarse como:

$$p_A(\lambda) = (\lambda - \lambda_i)^{a_i} q_A(\lambda) \quad a_i \in \mathbb{N}$$

¿Cómo calcular valores propios?: Sacar raíces de $p_A(\lambda)$ según se necesiten las reales y/o las complejas.

11.3. Cálculo de vectores propios

Método: Para $A_{n\times n}$.

1. Calcular valores propios, seleccionar los valores reales si es que se requieren solo estos. Considerar el conjunto $\{\lambda_1, \ldots, \lambda_n\}$.
2. Se busca resolver:
$$\begin{array}{rcl}
Ax_1 &=& \lambda_1 x_1 \quad \Longleftrightarrow \quad \underbrace{(A - \lambda_1\mathbb{I}_n)}_{\text{¡Ker!}} x_1 = 0 \\
\vdots & & \qquad\qquad\qquad \vdots \\
Ax_n &=& \lambda_n x_n \quad \Longleftrightarrow \quad (A - \lambda_n\mathbb{I}_n)x_n = 0
\end{array}$$
3. Con lo anterior, se ve que simplemente hay que calcular individualmente el Ker de cada matriz $A - \lambda x_1$ por pivoteo o el método que se estime conveniente (i.e. Regla de Cramer).
4. Verificar que **ninguna** solución al Ker sea trivial, pues en ese caso λ_i no corresponde a un valor propio ($\det(A - \lambda_i\mathbb{I}_n) \neq 0$).

Definición: Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, entonces:

1. Para cada λ_i el $\text{Ker}(A - \lambda_i \mathbb{I}_n)$ es un s.e.v. (pues es un conjunto generado) y se conoce como *subespacio propio* de λ_i.

2. Se define la *multiplicidad geométrica* (g_i) de λ_i como:

$$g_i = \dim\Big(\text{Ker}(A - \lambda \mathbb{I}_n)\Big)$$

Teorema: Sea λ_i valor propio de $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ con multiplicidad algebraica a_i y multiplicidad geométrica g_i, entonces:

$$1 \leq g_i \leq a_i$$

Propiedades:

1. Si existe $x \neq 0$ tal que $Ax = \lambda x$ con $\lambda \in \mathbb{R}$, entonces $A^k x = \lambda^k x$.

2. Si A es invertible y tiene a $\lambda \neq 0$ como valor propio, entonces $\dfrac{1}{\lambda}$ es valor propio de A^{-1}. Ambos tienen asociados los mismos vectores propios.

3. $\lambda = 0$ es valor propio de A si y solo si $Ker(A) \neq 0$. Es decir, si A es no invertible.

4. Las matrices A y B son *similares* si $A = C^{-1}BC$. Las matrices similares tienen el mismo polinomio característico: $p_A(\lambda) = p_B(\lambda)$.

5. Si $\lambda_1, \ldots, \lambda_n$ son los valores propios de una matriz A, entonces $p_\lambda(0) = \prod \lambda_i$. Es decir, el determinante es igual al producto de los valores propios de A. (?)

Definición: Se define la *traza* de una matriz cuadrada A como la suma de los elementos de la diagonal principal. Es decir,

$$\text{Si } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \text{ entonces } tr(A) = \sum_{i=1}^{n} a_{ii}$$

Teorema: Se tiene que $tr(A) = \displaystyle\sum_{i=1}^{n} \lambda_i$ donde $\{\lambda_1, \ldots, \lambda_n\}$ es el conjunto de los valores propios de A (sin ser necesariamente distintos).

11.4. Diagonalización de una matriz cuadrada

Definición: Una matriz $A \in \mathcal{M}_{n\times n}$ se dice *diagonalizable* si las bases de los subespacios propios correspondientes a los distintos valores propios de A forman una base de $\mathbb{R}^n$.

Teorema: Sea $\lambda_1 \neq \lambda_2$ valores propios de una matriz A. Si x_1 es vector propio correspondiente a λ_1 y x_2 es vector propio correspondiente a λ_2, entonces el conjunto $\{x_1, x_2\}$ es l.i.

Corolario: $\text{Ker}(A - \lambda_1 \mathbb{I}) \cap \text{Ker}(A - \lambda_2 \mathbb{I}) = \{\overrightarrow{0}\}$.

Corolario: Sea $\{U_1, \dots, U_m\}$ el conjunto de los subespacios propios de A. Entonces, A es diagonalizable si $\dim(\mathbb{R}^n) = \displaystyle\sum_{i=1}^{m} \dim(U_i) = n$.

Definición: Sea $A \in \mathcal{M}_{n\times n}$. Diremos que A es diagonalizable si y solo si A es similar a una matriz diagonal. Es decir, si existen dos matrices $P_{n\times n}$ invertible y $D_{n\times n}$ en $\mathcal{M}_{n\times n}$ tales que:

$$A = PDP^{-1}$$

Donde $P = [x_1|\dots|x_n]$ y x_i es un vector propio de A y

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Teorema: $A \in \mathcal{M}_{n\times n}$ es diagonalizable si y solo si:

1. existe una base de $\mathbb{R}^n$ compuesta por vectores propios de A.
2. $a_i = g_i$ para todo λ_i valor propio de A.

11.4.1. Aplicaciones de matrices diagonalizables

1. Sea A diagonalizable, entonces $A = PDP^{-1}$. Luego,

$$A^k = \underbrace{(PDP^{-1})\dots(PDP^{-1})}_{k \text{ veces}} = PD^kP^{-1}$$

2. Sea $A_{n\times n}$ y sea $q(x)$ un polinomio con coeficientes reales. Si $q(x) = a_0 + a_1x + \dots + a_kx^k$, se define $q(A)$ como:

$$q(A) = a_0\mathbb{I} + a_1A + \dots + a_kA^k$$

Teorema: Si v es un vector propio de A correspondiente al valor propio λ, entonces:

$$q(A)v = q(\lambda)v$$

Teorema: Si A es diagonalizable, entonces:

$$q(A) = Pq(D)P^{-1}$$

3. **Teorema** (*de Cayley-Hamilton*): Sea $A_{n\times n}$ diagonalizable y $p_A(\lambda)$ si polinomio característico, entonces $p_A(A) = \emptyset$.

12. Ortogonalidad

12.1. Espacios vectoriales con producto interno

Definición: Sea V un espacio vectorial cualquiera. Una función $\varphi : V \times V \longrightarrow \mathbb{R}$, será un *produto interno* (p.i.) de V si $(\forall x, y, z \in V)(\forall \alpha \in \mathbb{R})$ se cumple que:

1. Simétrica: $\varphi(x, y) = \varphi(y, x)$.
2. Lineal en su primer argumento:
   - $\varphi(x + y, z) = \varphi(x, z) + \varphi(y, z)$.
   - $\varphi(\alpha x, y) = \alpha \varphi(x, y)$.
3. Definida positiva: $\varphi(x, x) \geq 0$ y $\varphi(x, x) = 0 \Longleftrightarrow x = 0$.

Definición: El par de objetos V (e.v.) y φ (p.i.) se llama *espacio con producto interno* o *espacio euclídeo.*

Propiedad: El p.i. también es lineal en su segundo argumento.

$$\varphi(x, y + z) = \varphi(x, y) + \varphi(x, z)$$

Notación: El producto interno también se nota como $\langle \cdot, \cdot \rangle$. Es decir, $\langle x, y \rangle = \varphi(x, y)$.

12.1.1. Matriz de un producto interno

Definición: Sea $Bi = \{v_1, \ldots, v_n\}$ base de V y φ un p.i. de V, y x e y dos vectores de V. Entonces la matriz $[\varphi]_B$ es la *matriz del producto interno* φ y esta representada por:

$$[\varphi]_B = \begin{bmatrix} \varphi(v_1, v_1) & \ldots & \varphi(v_1, v_n) \\ \vdots & \ddots & \vdots \\ \varphi(v_n, v_1) & \ldots & \varphi(v_n, v_n) \end{bmatrix}$$

Y si $\varphi(x, y) = \alpha$, entonces $[x]_B^T [\varphi]_B [y]_B = \alpha$.

Propiedades:

1. $[\varphi]_B$ es simétrica.
2. $\varphi(x, x)$ es una forma cuadrática positiva definida.
3. **Teorema:** Sea V un e.v. sobre $\mathbb{R}$ con $\dim(V) = n$ y B una base de V. Entonces una función $\varphi : V \times V \longrightarrow \mathbb{R}$ es un producto interno si y solo si existe una matriz $M_{n \times n}$ simétrica y positiva definida tal que:
$$(\forall x, y \in V) \quad \varphi(x, y) = [x]_B^T M [y]_B$$

12.1.2. Norma

Definición: Sea (V, φ) un espacio euclídeo. Se define la longitud o norma de un vector $x \in V$ como:

$$\|x\| = \sqrt{\varphi(x,x)}$$

Propiedades: Son las mismas indicadas para el producto punto en $\mathbb{R}^n$.

1. $\|\alpha x\| = |\alpha| \, \|x\|$.
2. $\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\varphi(x,y)$.
3. $\|x\| \geq 0$ y $\|x\| = 0 \Longleftrightarrow x = 0$.
4. *Desigualdad de Cauchy-Schwartz*:
$$|\varphi(x,y)| \leq \|x\| \, \|y\|$$
5. *Desigualdad triangular*:
$$\|x+y\| \, leq \, \|x\| + \|y\|$$

Definición: Se define el ángulo entre dos vectores x e y en el espacio euclídeo (V, φ) como:

$$\angle(x,y) = \arccos\left(\frac{\varphi(x,y)}{\|x\| \, \|y\|}\right)$$

12.1.3. Concepto de ortogonalidad

Definición: Sea (V, φ) un espacio euclídeo y x e $y \in V$. Entonces:

1. x e y son *ortogonales* si $\varphi(x,y) = 0$ y se anota $x \perp y$.
2. x es *ortogonal* al conjunto C (¡no vacío!) si $(\forall c \in C) \quad x \perp c$. Se anota $x \perp C$.
3. Se define el *complemento ortogonal* $C^\perp$ al conjunto no vacío $C \subseteq V$ como $C^\perp = \{x \in V \, : \, x \perp C\}$.

Notar que $\{0\}^\perp = V$ y a su vez $V^\perp = \{0\}$.

Observación importante: Un espacio euclídeo genera una geometría propia en dicho espacio, ya que se pueden definir la distancia y el ángulo entre vectores, por esta razón se pueden demostrar propiedades como el teorema de Pitágoras o similares.

Propiedades: Sea (V, φ) un espacio euclídeo, entonces:

1. $C^\perp \leqslant V$.
2. Para $x \in V$ se tiene que $x \perp \langle s_1, \ldots, s_m \rangle \Longleftrightarrow x \perp s_1, \ldots, x \perp s_m$.
3. Si B es una base de $S \leqslant V$, entonces $B^\perp = S^\perp$.
4. Si $C_1 \subseteq C_2$, entonces $C_2^\perp \subseteq C_1^\perp$.

12.1.4. Bases ortogonales y proceso de Gram-Schmidt

Definición: Sea (V, φ) un espacio euclídeo de dimensión n. Entonces, el conjunto $\{u_1, \ldots, u_n\}$ es una *base ortogonal* si $\varphi(u_i, u_j) = 0$ para todo $i \neq j$. Será una *base ortonormal* si además $\varphi(u_i, u_i) = 1$ para todo $i = 1, \ldots, n$.

Teorema (*Proceso de Gram-Schmidt*): La demostración es análoga al procedimiento para tener una base ortogonal. Sea (V, φ) un espacio euclídeo, entonces V tiene una base ortonormal.

Demostración y Procedimiento: Todo e.v. tiene al menos una base. Sea esta base llamada $B = \{b_1, \ldots, b_n\}$, y esta siempre debe ser conocida.

Parte 1: Determinación de la base ortogonal.

Es evidente que la base a obtener debe tener n elementos. Supongamos que estos son $\{v_1, \ldots, v_n\}$.

1. Partimos tomando $v_1 = b_1$.
2. $v_2 = b_2 - \alpha_{21} v_1$. El escalar α_{21} se debe elegir de modo que $v_1 \perp v_2$.
$$\varphi(v_1, v_2) = \varphi(v_1, b_2 - \alpha_{21} v_1) = 0$$
$$\varphi(v_1, b_2) - \alpha_{21}\varphi(v_1, v_1) = 0$$
$$\alpha_{21} = \frac{\varphi(v_1, b_2)}{\varphi(v_1, v_1)} \longrightarrow v_2 = b_2 - \frac{\varphi(v_1, b_2)}{\varphi(v_1, v_1)} v_1$$
3. $v_3 = b_3 - \alpha_{31} v_1 - \alpha_{32} v_2$. Nuevamente, α_{31} y α_{32} se eligen de modo que $v_3 \perp v_1$ y $v_3 \perp v_2$. Por este procedimiento análogo, tenemos que:
$$\alpha_{31} = \frac{\varphi(v_1, b_3)}{\varphi(v_1, v_1)} \quad \text{y} \quad a_{32} = \frac{\varphi(v_2, b_3)}{\varphi(v_2, v_2)}$$
4. Se sigue análogamente que:
$$v_k = b_k - \sum_{j=1}^{k-1} \frac{\varphi(v_j, b_k)}{\varphi(v_j, v_j)} v_j$$

¿Cómo garantizamos que el conjunto es l.i.? Porque al ser ortogonales, por teorema, son l.i. No hace falta, por lo tanto, demostrar que sean l.i., si no más bien señalar esta condición.

De esta forma, obtenemos cada v_k.

Parte 2: Determinación de la base ortonormal.

Supongamos que la base a buscar es $\widetilde{V} = \{\widetilde{v}_1, \ldots, \widetilde{v}_n\}$. Tenemos que:

$$\widetilde{v}_k = \frac{v_k}{\|v_k\|}$$

Pues de esta forma $\|\widetilde{v}_k\| = \left\| \dfrac{v_k}{\|v_k\|} \right\| = \dfrac{\|v_k\|}{\|v_k\|} = 1$.

12.1.5. Proyecciones ortogonales

Definición: Sea (V, φ) un espacio euclídeo de dimensión finita y $W \leqslant V$. Entonces para $x \in V$ se define la *proyección ortogonal* como aquel vector w que satisface que:

- $w \in W$.
- $b - w \perp W$.

Teorema: *La proyección ortogonal es única.* Sea (V, φ) un espacio euclídeo de dimensión finita y $W \leqslant V$. Para cada $b \in V$ existe un único vector $w \in V$ tal que $b - w \perp W$.

Obtención de la proyección ortogonal:

Razonamiento:

Tomemos la base $B = \{w_1, \ldots, w_k\}$ de W. Tomemos el b fijo que buscamos. Tenemos que:

$$w = \alpha_1 w_1 + \ldots + \alpha_k w_k$$

Entonces:

$$\begin{aligned} b - w \perp W \quad &\iff \quad \varphi(b - w, w_i) = 0 \quad \text{para cada } w_i \\ &\iff \quad \varphi(b - \alpha_1 w_1 - \ldots - \alpha_k w_k, w_i) = 0 \\ &\iff \quad \varphi(b, w_i) = \alpha_1 \varphi(w_1, w_i) + \ldots + \alpha_k \varphi(w_k, w_i) \end{aligned}$$

Es decir, para cada elemento de la base B nos queda:

$$\begin{gathered} \varphi(b, w_1) = \alpha_1 \varphi(w_1, w_1) + \ldots \alpha_k \varphi(w_k, w_1) \\ \vdots \\ \varphi(b, w_k) = \alpha_1 \varphi(w_1, w_k) + \ldots \alpha_k \varphi(w_k, w_k) \end{gathered}$$

Estas ecuaciones se conocen como *ecuaciones normales de proyección*. Esto se resume matricialmente en el sistema:

$$\underbrace{\begin{bmatrix} \varphi(w_1, w_1) & \ldots & \varphi(w_1, w_k) \\ \vdots & \ddots & \vdots \\ \varphi(w_k, w_1) & \ldots & \varphi(w_k, w_k) \end{bmatrix}}_{[\varphi_W]_B} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} \varphi(b, w_1) \\ \vdots \\ \varphi(b, w_k) \end{pmatrix}$$

Ambas partes del sistema son calculables. Este sistema **siempre tiene** solución única, ya que:

- $[\varphi_W]_B$, la matriz que representa al producto interno de W, es positiva definida, como todo producto interno.
- Como es positiva definida, es necesariamente invertible, ya que $\det([\varphi_W]_B) \neq 0$.

Una vez obtenidos $\alpha_1, \ldots, \alpha_k$ se evalua la combinación lineal para obtener w. **Ojo:** se pide w, no los α_i.

Forma mecánica:

Para un vector dado $b \in W$ se busca la proyección ortogonal w. Sea $B = \{w_1, \ldots, w_k\}$, se resuelve el sistema:

$$[\varphi]_B [w]_B = \begin{pmatrix} \varphi(b, w_1) \\ \vdots \\ \varphi(b, w_k) \end{pmatrix}$$

Luego, $[w]_B$ se convierte a lo que corresponda según el s.e.v. y su base dada.

Teorema: Sea $W \leqslant V$, donde V es un espacio euclídeo con producto interno φ. Entonces:

$$V = W \oplus W^{\perp}$$

Corolario: Para $b \in V$ existen únicos vectores $w \in W$ y $w^* \in W^{\perp}$ tales que $b = w + w^*$.

Corolario: Si (V, φ) es un e.v. de dimensión finita y $W \leqslant V$, entonces:

$$(W^{\perp})^{\perp} = W$$

12.1.6. El operador proyección P_W

Definición: Sea (V, φ) un espacio euclídeo y $W \leqslant V$, entonces se define el *operador proyección* $P_W : V \longrightarrow V$ por:

$$P_W(b) = w$$

Donde w es la proyección ortogonal de b sobre W.

Teoremas:

- P_W es una transformación lineal.
- $\text{Im}(P_W) = W$ y $\text{Ker}(P_W) = W^{\perp} = \text{Im}\,(P_{W^{\perp}})$.
- $P_W \circ P_W = P_W$.
- $P_W + P_{W^{\perp}} = Id$.
- P_W tiene valores propios $\lambda = 0$ con $W^{\perp}$ como subespacio asociado y $\lambda = 1$ con W como subespacio asociado ($P_W(b) = b$).

12.1.7. El operador reflexión R_W

Definición: El operador reflexión R_w busca entregar el vector simétrico a B en torno a W.

$$R_W(b) = b + 2 \cdot (w - b) = 2w - b$$
$$R_W(b) = 2P_W(b) - b$$

12.2. Producto punto en $\mathbb{R}^n$

Se considera el espacio euclídeo (V, φ) donde $V = \mathbb{R}^n$ y φ es el producto punto.

Teorema: $\mathrm{Im}(A)^{\perp} = \mathrm{Ker}(A^T)$.

Corolario: Como consecuencia de lo anterior, se tiene que:

1. $\mathrm{Ker}(A)^{\perp} = \mathrm{Im}(A^T)$.
2. $\mathrm{Im}(A) \oplus \mathrm{Ker}(A^T) = \mathbb{R}^m$ y $\mathrm{Im}(A^T) + \mathrm{Ker}(A) = \mathbb{R}^n$.
3. Sea $W \leqslant \mathbb{R}^n$, con una base B y una matriz A cuyas columnas son cada elemento de la base respectivo, entonces $W^{\perp} = \mathrm{Ker}(A^T)$.

Definición: Sea A la matriz cuyas columnas son base del espacio euclídeo $(\mathbb{R}^n, \cdot)$, entonces la matriz del producto punto está determinada por $A^T A$. Esto es fácil verlo, pues para:

$$A = [c_1 | \dots | c_n] \longrightarrow A^T = \begin{bmatrix} c_1^T \\ \vdots \\ c_n^T \end{bmatrix} \longrightarrow A^T A = \begin{bmatrix} c_1 \cdot c_1 & \dots & c_n \cdot c_1 \\ \vdots & \ddots & \vdots \\ c_1 \cdot c_n & \dots & c_n \cdot c_n \end{bmatrix}$$

12.2.1. Factorización QR y matrices ortogonales

Definición: Se entiende por factorización QR aquella factorización matricial para matrices $A \in \mathcal{M}_{m\times n}$ cuyas columnas son l.i.

$$A = QR$$

Donde $Q = [\widehat{v}_1 | \dots | \widehat{v}_n]$ y $R = \overbrace{\begin{bmatrix} \|v_1\| & \widehat{v}_1 \cdot u_2 & \dots & \widehat{v}_1 \cdot u_n \\ 0 & \|v_2\| & \dots & \widehat{v}_2 \cdot u_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \|v_n\| \end{bmatrix}}^{\text{aumentar } u_i \longrightarrow}$.

Razonamiento de la factorización:

Sea $A_{m\times n}$ con columnas l.i.

$$A = [u_1 | \dots | u_n]$$

Entonces $\{u_1, \ldots, u_n\}$ es una base $\mathbb{R}^m$. Se le puede aplicar el proceso de Gram-Schmidt para obtener la base ortogonal $\{v_1, \ldots, v_n\}$.

Sabemos que: $v_k = u_k - \sum_{i=1}^{k-1} \frac{v_i \cdot u_k}{v_i \cdot v_i} v_i \Longrightarrow u_k = v_k + \sum_{i=1}^{k-1} \frac{v_i \cdot u_k}{v_i \cdot v_i} v_i$. Esto se puede escribir matricialmente como:

$$u_k = [v_1|\ldots|v_n] \begin{pmatrix} \frac{v_1 \cdot u_k}{v_1 \cdot v_1} \\ \vdots \\ \frac{v_{k-1} \cdot u_k}{v_{k-1} \cdot v_{k-1}} \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \longrightarrow \text{posición } k\text{-ésima}$$

$$A = \underbrace{[v_1|\ldots|v_n]}_{(1)} \underbrace{\begin{bmatrix} 1 & \alpha_{21} & \alpha_{31} & \alpha_{41} & \ldots & \ldots & \alpha_{n1} \\ 0 & 1 & \alpha_{32} & \alpha_{42} & & & \vdots \\ \vdots & & 1 & \alpha_{43} & & & \vdots \\ \vdots & & & 1 & & & \vdots \\ \vdots & & & & \ddots & & \vdots \\ \vdots & & & & & \ddots & \alpha_{n(n-1)} \\ 0 & \ldots & \ldots & \ldots & \ldots & \ldots & 1 \end{bmatrix}_{n \times n}}_{(2)}$$

Donde $\alpha_{ki} = \frac{v_i \cdot u_k}{\|v_i\|^2}$. Considerando que $\widehat{v}_i = \frac{v_i}{\|v_i\|}$, podemos hacer uso de operaciones matriciales elementales. Tenemos que:

$$A = (1) \underbrace{E_i(\frac{1}{\alpha}) E_i(\alpha)}_{\mathbb{I}_n} (2)$$

Donde E_i es la matriz de la operación elemental de escalamiento. Como $E_i = E_i^T$, a la matriz (1) se le aplica la operación por columnas, y a A se le aplica la operación por filas. Luego, podemos hacer esto para cada columna de (1) y obtener los vectores normales. Es decir,

$$A = [\widehat{v}_1|\ldots|\widehat{v}_n] \begin{bmatrix} \|v_1\| & \alpha_{21}\|v_1\| & \ldots & \alpha_{n1}\|v_1\| \\ 0 & \|v_2\| & & \vdots \\ \vdots & & \ddots & \alpha_{n(n-1)}\|v_{n-1}\| \\ 0 & 0 & \ldots & \|v_n\| \end{bmatrix}$$

Pero $\alpha_{ki}\|v_i\| = \dfrac{v_i \cdot u_k}{\|v_i\|} = \widehat{v}_i \cdot u_k$. Entonces:

$$A = \underbrace{[\widehat{v}_1|\dots|\widehat{v}_n]}_{Q} \underbrace{\begin{bmatrix} \|v_1\| & \widehat{v}_1 \cdot u_2 & \dots & \widehat{v}_1 \cdot u_n \\ 0 & \|v_2\| & & \vdots \\ \vdots & & \ddots & \widehat{v}_{n-1} \cdot u_n \\ 0 & 0 & \dots & \|v_n\| \end{bmatrix}}_{R}$$

- R es de $n \times n$, triangular superior e invertible.

Teorema: Sea $A \in \mathcal{M}_{m\times n}$ y $A = QR$. Entonces, $Q^TQ = \mathbb{I}$.

Es fácil de notar, pues:

$$\begin{bmatrix} v_1^T \\ \vdots \\ v_k^T \end{bmatrix} \begin{bmatrix} v_1 & \dots & v_k \end{bmatrix} = \begin{bmatrix} v_1^T v_1 & \dots & v_1^T v_k \\ \vdots & \ddots & \vdots \\ v_k^T v_1 & \dots & v_k^T v_k \end{bmatrix}$$

Como $\{v_1, \dots, v_k\}$ es una base ortonormal, tenemos que $v_1^T v_1 = \|v_1\|^2 = 1$ y que $v_i^T v_j = v_i \cdot v_j = 0$ para $i \neq j$. Entonces,

$$Q^TQ = \begin{bmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{bmatrix} = \mathbb{I}$$

Definición: Una matriz A es *ortogonal* si:

- es cuadrada ($A \in \mathcal{M}_{n\times n}$).
- $A^TA = \mathbb{I}_n$.

Entonces, si $A_{n\times n}$ es invertible $|A| \neq 0$, tiene factorización QR, y por lo tanto Q es ortogonal. **Ojo:** No toda matriz Q de QR es ortogonal, porque no es necesario que Q sea cuadrada.

Teorema: Sea $Q \in \mathcal{M}_{n\times n}$, entonces las siguientes proposiciones son equivalentes entre sí:

1. Q es ortogonal.
2. $Q^TQ = QQ^T = \mathbb{I}_n \Longrightarrow Q^{-1} = Q^T$.
3. Las columnas de Q forman una base ortonormal de R^n.
4. $\|Qx\| = \|x\|$. Ojo, es la norma, y no el vector.
5. $Qx \cdot Qy = x \cdot y$.
6. Si $\{x_1, \dots, x_n\}$ es una base ortonormal, entonces $\{Qx_1, \dots, Qx_n\}$ también es base ortonormal.

Observaciones importantes: Las matrices ortogonales adquiere especial importancia por las siguientes razones:

1. Las transformaciones lineales representadas por matrices ortogonales respetan los ángulos y distancias entre los vectores de los espacios de partida y de llegada.
2. Si las columnas de A forman una base ortogonal de $\text{Im}(A)$, entonces $A^T A$ será una matriz diagonal. Si forman una base ortonormal, entonces $A^T A = \mathbb{I}$.

Observación: $A^T A$ es definida positiva $\iff A$ tiene columnas l.i $\implies A = QR$. Por lo tanto,

$$A^T A = R^T Q^T Q R = R^T R$$

Es decir, tiene una factorización de Cholesky con raíces. Además $A^T A$ representa la matriz del producto punto en la base formada por las columnas de A y el e.v. $\text{Im}(A)$.

12.2.2. Matrices de proyección

El cálculo de proyecciones se vuelve matricial en el espacio euclídeo canónico $\mathbb{R}^n$. El proceso para determinar la matriz es el siguiente:

Razonamiento: Sea $W = \left\langle \underbrace{w_1, \ldots, w_k}_{\text{base}} \right\rangle \leqslant \mathbb{R}^n$. Se tiene que:

$$\text{Para } w \in W,\ w = x_1 w_1 + \ldots + x_k w_k = Cx$$

Donde $C = [w_1 | \ldots | w_k]$ y $x = (x_1, \ldots, x_k)^T$. Para $b \in \mathbb{R}^n$ se debe cumplir que:

$$b - w = b - Cx \perp W$$

Debemos buscar el x, pues este nos entrega la información de la combinación lineal para determinar w. Por lo tanto, se debe cumplir que:

$$w_i \cdot (b - Cx) = w_i^T (b - Cx) = 0$$

Las matrices se usan para aglomerar información simultáneamente, como sistemas de ecuaciones, por ejemplo. Por lo tanto, esto se puede resumir en:

$$C^T (b - Cx) = 0, \text{ donde } C^T = \begin{bmatrix} w_1^T \\ \vdots \\ w_k^T \end{bmatrix}$$

Lo que, reescrito, significa que $C^T b = C^T C x$. Donde $C^T b$ es un vector calculable y $C^T C$ es una matriz también calculable con la información entregada. Por lo tanto, es fácil resolver x por los métodos ya conocidos.

Observación: Este sistema tiene solución única, ya que $C^T C$ representa al producto punto interno de W y es definida positiva, por lo tanto invertible.

Es decir,

$$x = (C^T C)^{-1} C b$$

Como $w = P_w(b) = Cx$, entonces $P_w(b) = C(C^T C)^{-1} C b$. Hemos encontrado la matriz de la proyección.

Ojo: $C^T C$ es por sí invertible, pero no quiere decir que C^T y C lo sean. Recordar que es una implicancia y no una equivalencia este teorema.

Definición: Sea $C \in \mathcal{M}_{k\times n}$ una matriz con columnas l.i., entonces $C(C^T C)^{-1} C^T$ es una matriz de proyección, y representa a la operación de proyección sobre $\mathrm{Im}(C)$.

Propiedades:

1. Si C y D pertenecen a $\mathcal{M}_{k\times n}$, tienen columnas l.i. e $\mathrm{Im}(C) = \mathrm{Im}(D)$. Entonces,
$$C(C^T C)^{-1} C^T = D(D^T D)^{-1} D^T$$
2. Si $A = C(C^T C)^{-1} C^T$, entonces $\mathrm{Im}(A) = \mathrm{Im}(C)$ y $\mathrm{Ker}(A) = \mathrm{Ker}(C^T)$.
3. Si A es la matriz de proyección sobre W, entonces la matriz de proyección sobre $W^\perp$ es $\mathbb{I} - A$.
4. Si A es la matriz de proyección sobre W, entonces la matriz de reflexión sobre W es $2A - \mathbb{I}$ y sobre $W^\perp$ es $\mathbb{I} - 2A$.
5. P es matriz de proyección sobre W si y solo si $P^2 = P$ y $P^T = P$. Luego, $W = \mathrm{Im}(P)$.
6. Si W tiene una base B ortonormal, entonces C es ortogonal, y la matriz de proyección viene a ser CC^T.

Observación importante: Sea P la matriz de proyección sobre un s.e.v. $W \leqslant \mathbb{R}^n$ y Q la matriz de proyección sobre $W^\perp$. Como:

$$\dim(W) + \dim(W^\perp) = n$$

Entonces, se puede elegir calcular la matriz de proyección que corresponda al s.e.v. de menor dimensión, pues se simplifican drásticamente los cálculos efectuados sobre $C(C^T C)^{-1} C$. Luego de haber calculado, basta recordar que:

$$P + Q = \mathbb{I}$$

12.2.3. Mínimos cuadrados

Definición: Sea $A_{m\times n}$, entonces todo vector x_0 que es solución del sistema $Ax = P_W(b)$ es una solución por mínimos cuadrados del sistema $Ax = b$. Si $b \in \mathrm{Im}(A)$, entonces las solución por mínimos cuadrados coinciden con las soluciones del sistema. Es decir,

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|^2 = \|Ax_0 - b\|^2$$

Razonamientos: La distancia mínima de Ax a b cuando no tiene solución esta dada por $w = Ax - b$, que pertenece a $\mathrm{Im}(A)^\perp = \mathrm{Ker}(A^T)$. Luego, como $w \in \mathrm{Ker}(A^T)$, se tiene que:

$$A^T w = 0 \longrightarrow A^T(Ax - b) = 0$$
$$A^T Ax = A^T b$$

Calcular mínimos cuadrados:

1. Consideramos el sistema $Ax = b$. ¿Tiene soluciones?

   Sí: Solucionamos este sistema. Esa es la solución

   No: Continue.

2. Se multiplica por la izquierda A^T. Entonces se tiene $A^T Ax = A^T b$. ¿Tiene columnas l.i.?

   Sí: Entonces $A^T A$ es positiva definida, ya que representa al producto punto con la base dada por las columnas de A, y por lo tanto invertible. Luego, tiene solución única el sistema. Por lo tanto, $x = (A^T A)^{-1} A^T b$.

   No: Las soluciones son infinitas. Simplemente se resuelve el sistema. Todos los datos son conocidos.

Casos concretos / Problemas tipo:

1. <u>Buscar mínimos de formas cuadráticas</u>: Una expresión de la forma:

   $$\min_{x,y\in\mathbb{R}} (x + y + \alpha_0)^2 + (x - y + \alpha_1)^2 + (x + y + \alpha_2)^2$$

   O similares, se pueden calcular de esta forma. Esta expresión se corresponde con la norma al cuadrado de un vector u. En particular, con:

   $$u = \begin{pmatrix} x + y + \alpha_0 \\ x - y + \alpha_1 \\ x + y + \alpha_2 \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}$$

   $$u = \underbrace{\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}}_{A} \underbrace{\begin{pmatrix} x \\ y \end{pmatrix}}_{\vec{v}} + \underbrace{\begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix}}_{-\vec{b}}$$

Osea, buscamos:

$$\min_{x,y\in\mathbb{R}} \|u\|^2 = \min_{v\in\mathbb{R}^n} \|Av-b\|^2$$

Desde aquí, ya es posible resolver.

2. Distancia punto-hiperplano: Una forma es con proyecciones ortogonales, otra es con mínimos cuadráticos. Tenemos un hiperplano $\mathcal{H}: x_0+\langle v_1,\dots,v_k\rangle$ en $\mathbb{R}^n$ y un punto c en $\mathbb{R}^n$ determinado por sus coordenadas respectivas. Tenemos que cualquier punto del hiperplano es de la forma:

$$x\in\mathcal{H} \iff x = x_0+\alpha_1 v_1+\dots\alpha_k v_k \quad \text{con } \alpha_i\in\mathbb{R}$$

$$x = x_0 + [v_1|\dots|v_k]\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_k\end{pmatrix}$$

Por lo tanto, la distancia del punto al hiperplano es la mínima que se pueda dar, ajustando los α_i para que se cumpla esto. Es decir, buscamos también aquel punto del hiperplano con aquel que la distancia al punto sea mínima.

Sabemos que este punto es la proyección ortogonal (se puede ver gráficamente), y buscamos por lo tanto:

$$\min_{\alpha_i\in\mathbb{R}}\left\|x_0+[v_1|\dots|v_k]\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_k\end{pmatrix}-c\right\|^2 = \min_{\alpha_i\in\mathbb{R}}\left\|\underbrace{[v_1|\dots|v_k]}_{A}\underbrace{\begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_k\end{pmatrix}}_{x}\underbrace{-c+x_0}_{-b}\right\|^2$$

El problema se reduce a:

$$\min_{x\in\mathbb{R}^n}\|Ax-b\|^2$$

Esto es algo que ya sabemos hacer. Una vez obtenido el x, que tiene los coeficientes de nuestra combinación lineal, simplemente calculamos el valor de la norma, y tendremos la distancia.

Respecto a hiperplanos: Para un hiperplano de ecuación $\mathcal{H}: n\cdot x = 0$ se tiene la ecuación cartesiana aproximada $n_1x_1+\dots+n_mx_m = 0$. Luego, se tiene que $\mathcal{H}^\perp = \langle n\rangle$, ya que este representa al vector normal al plano.

12.3. Valores y vectores propios de matrices simétricas

Teorema: Sea $A\in\mathcal{M}_{n\times n}(\mathbb{R})$ y tal que $A^T = A$, entonces:

1. Todos los valores propios de A son reales.

2. Si $\lambda_1\neq\lambda_2$ son valores propios de A, entonces $\mathrm{Ker}(A-\lambda_1\mathbb{I})\perp\mathrm{Ker}(A-\lambda_2\mathbb{I})$.

Observación: Entonces, para cada valor propio, podemos tomar una base y, mediante el proceso de Gram-Schmidt, obtener una base normal para cada una. Por el teorema anterior, juntar todos los vectores de todas las bases obtenidas generaría en efecto una base ortonormal.

Teorema: Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y tal que $A^T = A$, entonces A es diagonalizable.

Entonces, para una matriz simétrica, se puede hacer la factorización $A = PDP^{-1}$. Considerando la base ortonormal para esta factorización, se tiene que:

$$P = [v_1|\ldots|v_n] \quad \text{con } \|v_1\|^2 = 1$$

Entonces, es una matriz ortogonal, y $P^{-1} = P^T$.

Definición: Se dirá que $A \in \mathcal{M}_{n\times n}$ y $A^T = A$ fue *diagonalizada ortogonalmente* si:

$$A = PDP^T$$

Donde, A es diagonalizable, y en este caso en particular, P es ortogonal.