

Álgebra Lineal I

Otoño 2024

Dr. Andreas Wachtel

A few nice symbols that I will use throughout:



A definition is for a mathematician like an anchor for a ship or a root for a tree.

Each provides something you can hold on to in some proofs.

Similar like leafs  hold on to a tree, mathematical results are based on definitions.

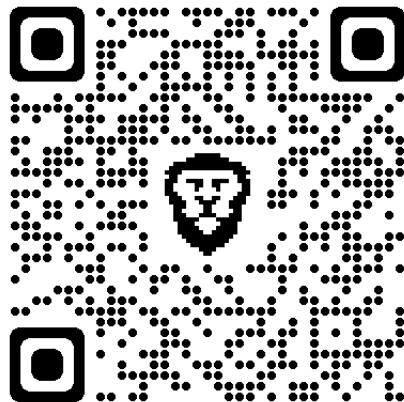
Other scientists usually base their argumentos on definitions and mathematical results.



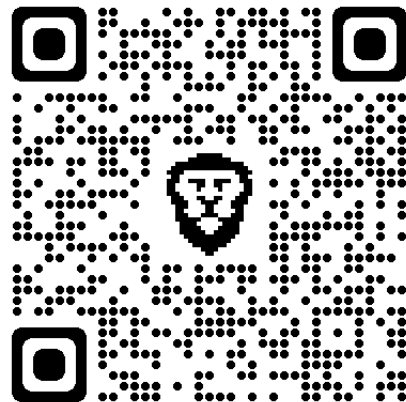
This symbol indicates a “positive” result or consequence. Something worth remembering.

Software que usaré para resolver, calcular
y visualizar algunas cosas en vivo:

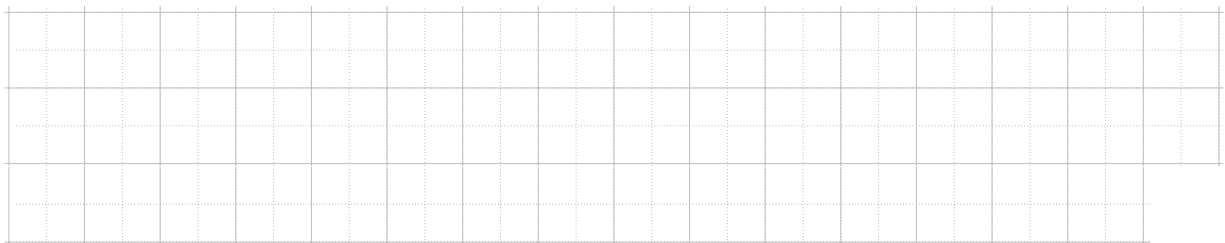
Geogebra



Google Colab



Para escribir o dibujar algo:



Índice general

1. Sistemas de Ecuaciones Lineales	1
1.1. Formas de visualizar un sistema de ecuaciones	1
1.1.1. Forma 1: Por fila / ecuación	1
1.1.2. Forma 2: Por columnas	2
1.1.3. Visualización en sistemas de 3 incógnitas	4
1.1.4. Casos regulares (cuando existen soluciones)	7
1.1.5. Casos singulares (cuando no existen soluciones)	7
1.2. Representación matricial	12
1.2.1. La forma matricial de un sistema lineal	13
1.3. Sistemas de m ecuaciones y n incógnitas	18
1.3.1. Sistemas equivalentes	19
1.3.2. Operaciones elementales	19
1.4. Eliminación Gaussiana y el método de Gauss-Jordan	25
1.4.1. Los pasos del método de Gauss-Jordan.	29
1.4.2. Resumen	31
1.5. Combinaciones lineales y sistemas homogéneos	35
1.6. Aplicaciones de sistemas de ecuaciones lineales	38
2. Álgebra matricial	47
2.1. Operaciones con matrices	48
2.1.1. Sumas de matrices	48
2.1.2. Producto entre una matriz y un escalar	50
2.1.3. Productos de matrices	51
2.2. Propiedades de matrices transpuestas	62
2.3. Matrices especiales	64
2.4. Motivación del siguiente tema	66
3. Espacios vectoriales	67
3.1. Ejemplos	68
3.1.1. El espacio vectorial R^n	68
3.1.2. El espacio vectorial de matrices reales	70
3.1.3. El espacio vectorial de los polinomios	73
3.2. Propiedades adicionales en espacios vectoriales	76
3.3. Subespacios vectoriales	78
3.3.1. Ejemplo 1	79

3.3.2. Ejemplo 2	80
3.4. Sub-espacios asociados a una matriz	80
4. Combinaciones lineales	83
4.1. Conjuntos generados	89
4.2. Dependencia e independencia lineal	96
4.3. Bases y dimensión	101
4.3.1. Bases y dimensión de sub-espacios	104
4.3.2. Resultados generales e importantes	112
4.4. Resumen y un mind map	117
5. Matrices inversas	119
5.1. Existencia de la inversa (I)	125
5.2. Propiedades de la inversa	126
5.3. Matrices elementales y sus inversas	128
5.3.1. Eliminación Gaussiana con matrices elementales	130
5.4. Una formula rápida para la inversa de 2×2	133
6. Determinantes	135
6.1. El determinante es una función multilineal.	137
6.2. Propiedades de determinantes	140
6.2.1. La primera propiedad	140
6.2.2. Determinantes y operaciones elementales	142
6.2.3. Determinantes de productos	144
6.2.4. Determinantes de transpuestas	145
6.3. La función determinante	147
6.3.1. La fórmula explícita para tamaño 3×3	149
6.3.2. La fórmula explícita para tamaño $n \times n$	150
6.4. Resultados generales	155
6.5. La regla de Cramer	158
6.6. La matriz de cofactores (de adjuntos)	160
7. Valores y vectores propios y Diagonalización de Matrices	161
7.1. Propiedades y resultados importantes	171
7.1.1. Eigenvalores y matrices inversas	173
7.1.2. Eigenvalores de matrices triangulares en bloques	174
7.1.3. Sistemas dinámicos discretos y cadenas de Markov	175
7.2. Bases de eigenvectores	177
7.3. Semejanza y Diagonalización	183
8. Transformaciones Lineales	193
8.1. Propiedades	199

8.2. Rango (Imagen) y núcleo	200
8.2.1. Transformaciones lineales son muy restrictivas	202
8.2.2. Transformaciones lineales preservan propiedades	203
8.3. Transformaciones inyectivas, sobres y biyectivas.	203
8.4. Composición e inversas	210
8.4.1. Inversas	212
8.4.2. Ap. Transporte seguro de datos binarios	214
8.4.3. Ap. Mover brazos en robótica	214
9. Un espacio vectorial con producto punto	215
9.1. Ángulos y ortogonalidad	216
9.2. Transformaciones rígidas	220
9.3. Matrices positivas definidas	222
Bibliografía	225

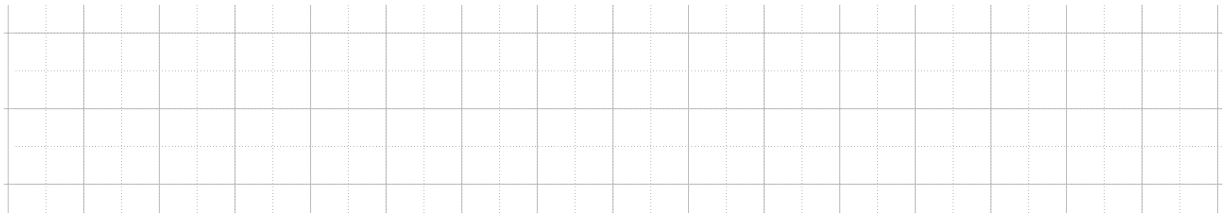
1. Sistemas de Ecuaciones Lineales

El problema central de Álgebra Lineal es **resolver** un *sistema lineal* (también “*sistema de ecuaciones lineales*”) y **dar criterios** bajo los cuales el sistema tiene ninguna solución, una única solución o varias soluciones.

Una ecuación es lineal significa que las incógnitas solamente estarán multiplicadas por números y nunca con incógnitas, es decir, nunca veremos xy, x^2, y^2, xy^2 , etc.

1.1. Formas de visualizar un sistema de ecuaciones

Nuestro primer sistema lineal de ecuaciones lineales será:



1.1.1. Forma 1: Por fila / ecuación

Esta primera forma separa el sistema por filas y nos concentraremos en una ecuación (fila) a la vez. Ambas ecuaciones producen una línea (recta) en el plano $x - y$.

La recta de la primera ecuación _____ pasa por los puntos:

La recta de la segunda ecuación _____ pasa por los puntos:



1. Sistemas de Ecuaciones Lineales

Al punto _____ se le conoce como la solución del sistema de ecuaciones y será denotado mediante un conjunto de solución:



La visualización por filas nos mostró dos líneas que se intersecan en un solo punto el cual será la solución. Cuando, las rectas son paralelas (y no iguales) no existe la intersección, y por tanto no habrá un punto que será solución.

1.1.2. Forma 2: Por columnas

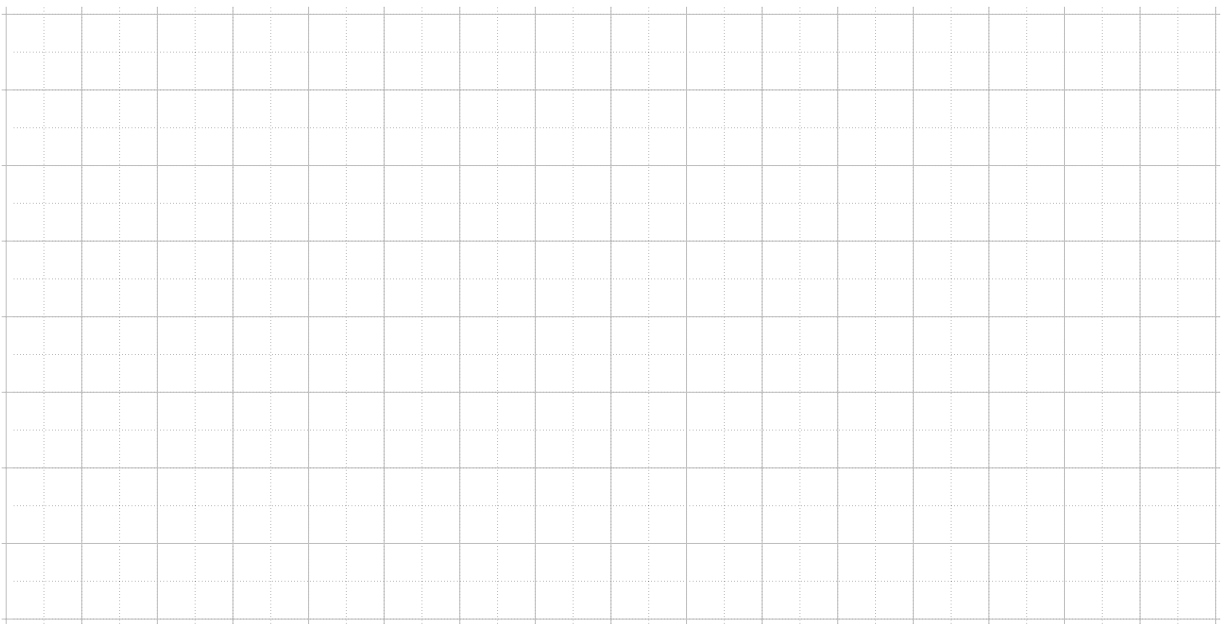
En esta visualización usaremos **dos operaciones que serán el corazón** del Álgebra Lineal.

1. La suma de dos vectores

2. La multiplicación de un vector por escalar

Para efectuar suma de vectores (de la misma longitud) la suma se aplica entrada por entrada. Similarmente, en la multiplicación “escalar (número) * vector” cada entrada del vector se multiplica por el escalar, *e.g.* $2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 * 1 \\ 2 * 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$.

Ejemplos:





Definición. Cuando se combinan estas dos operaciones el resultado se conoce como una **Combinación Lineal** de los vectores involucrados.



De hecho, cada sistema lineal representa una combinación lineal de columnas igual a una columna $\vec{b}$ (el lado derecho).

Ejemplo: Si separamos el sistema _____ en sus columnas, entonces obtenemos una combinación lineal de columnas igualada a una columna $\vec{b}$. Es decir:



El **problema ahorra** es encontrar valores para las incógnitas de tal manera que la combinación lineal de los vectores $\vec{v}, \vec{w}$ sea igual al vector $\vec{b}$.

Ejemplo: Dibujamos para encontrar la solución geométrica:



La visualización por columnas combina los vectores columna del lado izquierdo para producir al vector $\vec{b}$ del lado derecho.

Ejercicio: Proporcione la solución geométrica del siguiente sistema de ecuaciones a través de las visualizaciones por filas y columnas.

$$2x - y = 1$$

$$x + y = 5.$$

1.1.3. Visualización en sistemas de 3 incógnitas

A continuación usaremos el siguiente sistema de ecuaciones lineales:

$$x + 2y + 3z = 6$$

$$2x + 5y + 2z = 4$$

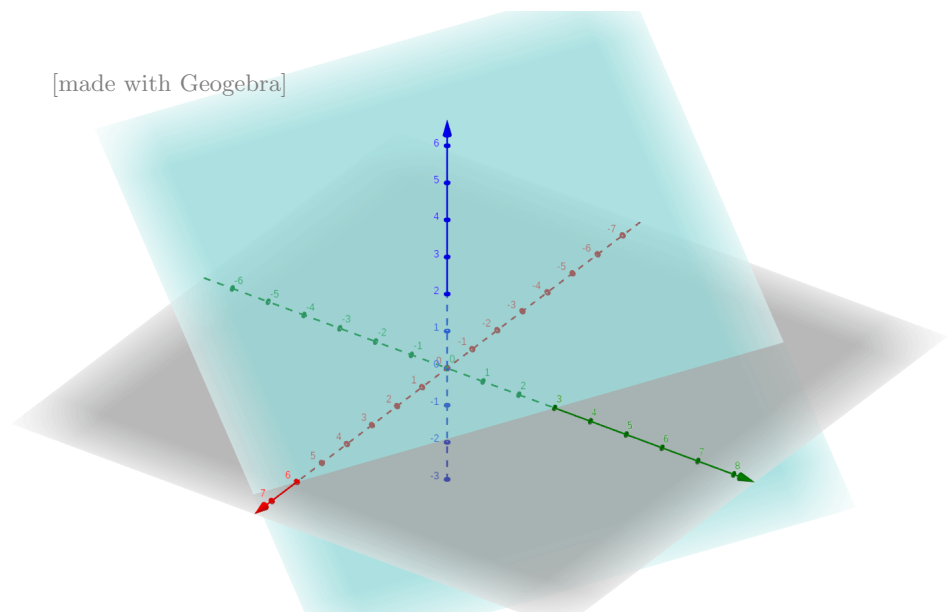
$$6x - 3y + z = 2$$

Estudiaremos este sistema por filas y columnas:

Empezamos por filas:

Cada fila es una ecuación que produce un plano en el espacio de 3 dimensiones ($x - y - z$).

La primera fila es la ecuación



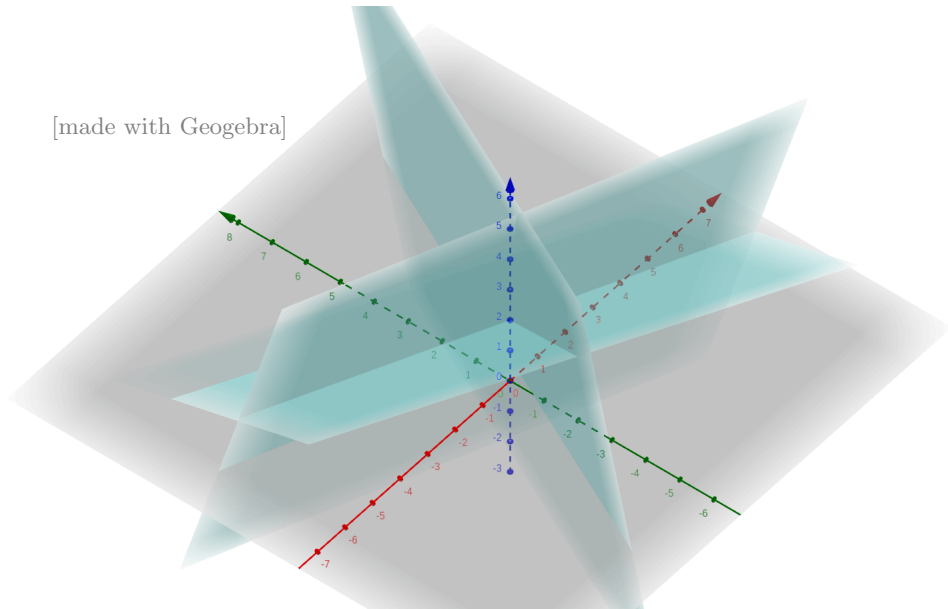
La segunda fila es la ecuación $2x + 5y + 2z = 4$ describe al plano que pasa por los puntos



e interseca al primer plano en una línea L . El resultado usual de intersecar dos ecuaciones con 3 incógnitas es una línea L , que será el conjunto solución, siempre y cuando las 2 ecuaciones representan 2 planos en $\mathbb{R}^3$ que no son paralelos. (“en $\mathbb{R}^3$ ” es corto para “en dimensión 3”).



[made with Geogebra]



La visualización por columnas:

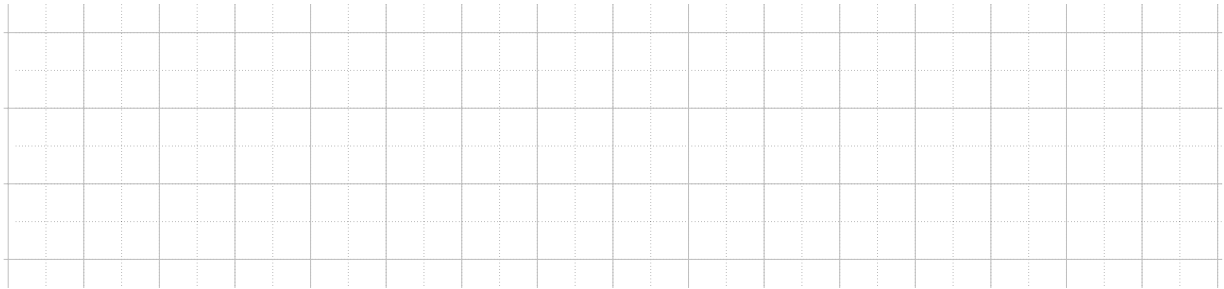
El sistema

$$x + 2y + 3z = 6$$

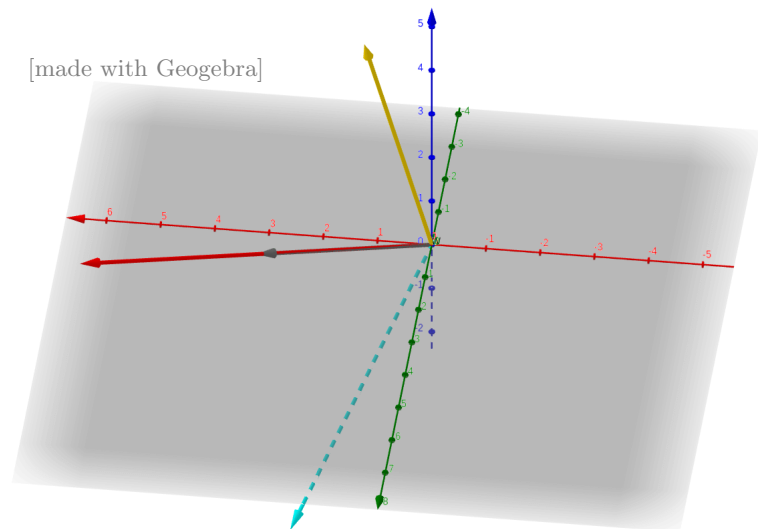
$$2x + 5y + 2z = 4$$

$$6x - 3y + z = 2$$

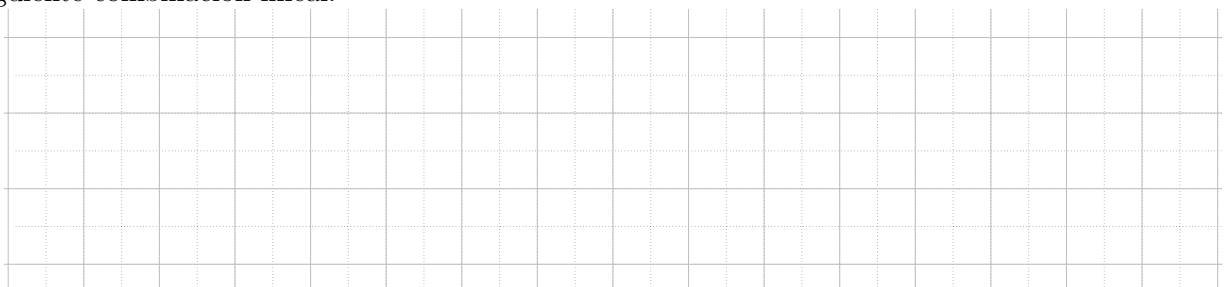
equivale a:



Las incógnitas x, y, z son los números (escalares) con los que multiplicaremos los tres vectores columnas con el fin de producir (llegar) a $\vec{b} = \begin{bmatrix} 6 \\ 4 \\ 2 \end{bmatrix}$.



De esto modo los coeficientes que se necesitan son $x = 0, y = 0, z = 2$. Es decir, tomamos la siguiente combinación lineal:



1.1.4. Casos regulares (cuando existen soluciones)

- La visualización por filas nos da la intersección de tres planos.
- En la visualización por columnas buscamos los números para formar una combinación lineal de las columnas y llegar al lado derecho $\vec{b}$.
- Soluciones de ecuaciones:

Puntos de intersección

son

los coeficientes en la combinación lineal de las columnas.

- El conjunto solución puede ser una recta L , por ejemplo, en el caso que tenemos 2 ecuaciones con 3 incógnitas. En este caso, se intersecan dos planos y el conjunto solución es una recta L .

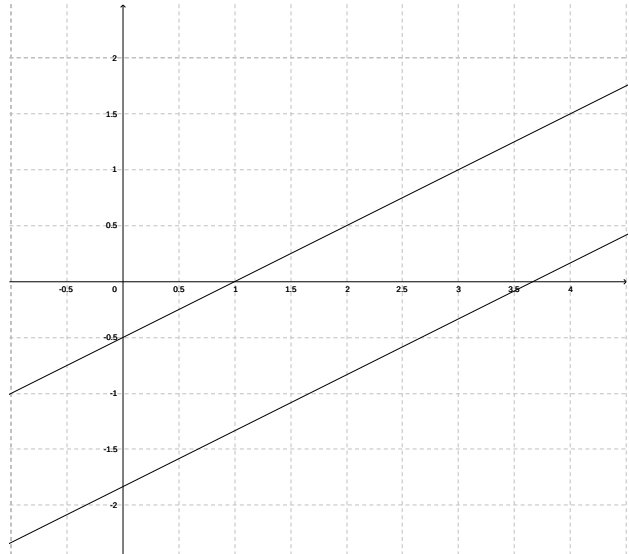
1.1.5. Casos singulares (cuando no existen soluciones)

En 2 dimensiones: Supongamos que tenemos el siguiente sistema de ecuaciones:

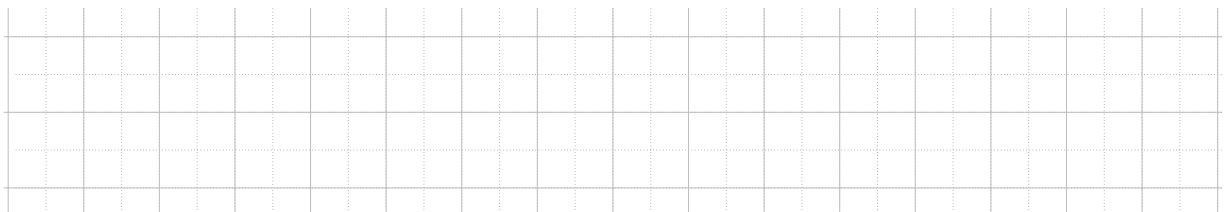
$$x - 2y = 1$$

$$3x - 6y = 11$$

Cuando se visualiza por filas observamos que las rectas son paralelas & no se intersecan.



Cuando observamos la representación por columnas tenemos



Dibujemos los vectores para ver mejor:

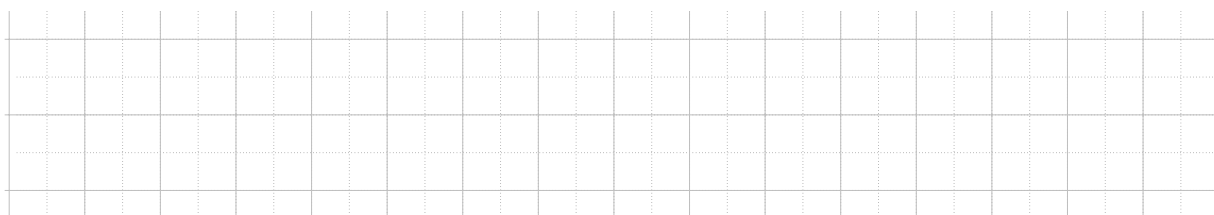


En 3 dimensiones: Si visualizamos el siguiente sistema veremos que dos planos son paralelos, por lo cual no se pueden intersectar y no habrá solución al sistema:

$$\begin{array}{rcl} 2x + y + z = 3 \\ 4x + 2y + 2z = 10 \\ 4x + 6z = -2 \end{array} \iff \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix} x + \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} y + \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} z = \begin{bmatrix} 3 \\ 10 \\ -2 \end{bmatrix}$$

Observe que en el sistema anterior las primeras 2 ecuaciones se contradicen, puesto que los coeficientes de la segunda ecuación son el doble de las coeficientes de la primera ecuación, pero en el lado derecho $2 * 3 \neq 10$.

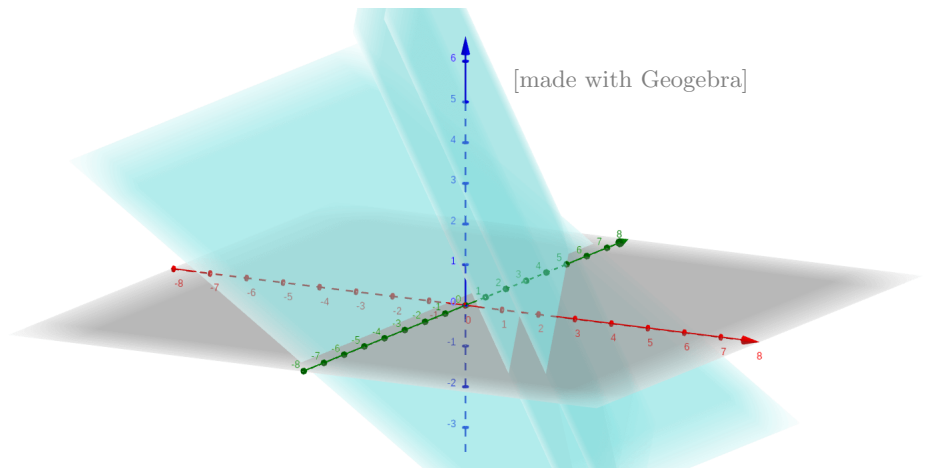
Para obtener algo más aterrizado hacemos los cálculos. Primero: (2* “Ecuación 1”)



Ahora, calculando 2* “Ecuación 1” – “Ecuación 2” llegamos que $0 = -4$:



Si lo visualizamos por filas, entonces 2 de 3 planos son paralelos & no iguales (no se intersectan). Por lo tanto, no existe la solución (un punto en el cual se intersectan los tres planos).



🍃 | Cuando la dimensión aumenta, hay más condiciones para la existencia de soluciones:

- No existe solución cuando por lo menos 2 de 3 planos son paralelos & no iguales. Abajo veremos un ejemplo que muestra que no existe una solución para 3 planos no paralelos.
 $\therefore$ Necesitamos la teoría que veremos más adelante.
- Por cierto: en $\mathbb{R}^2$ (dos dimensiones) hay solo tres posibilidades:
 - a) Solución única (un punto en $\mathbb{R}^2$) cuando tienes 2 líneas no paralelas.
 - b) Muchas soluciones (una recta en $\mathbb{R}^2$) cuando tienes 2 líneas iguales.
 - c) El negativo: no existe solución cuando tienes 2 líneas paralelas & distintas.

Ejemplo: El siguiente sistema muestra que a veces se necesitan considerar todas las filas para ver que no existe la solución:

$$\begin{aligned} x + y + z &= 2 \\ 2x \quad \quad + 3z &= 5 \\ 3x + y + 4z &= 6 \end{aligned}$$

Aquí del lado izquierdo (de la igualdad) la tercera fila es suma de las primeras dos. Entonces, si sumamos las primeras 2 ecuaciones y restamos la tercera da 0 del lado izquierdo.

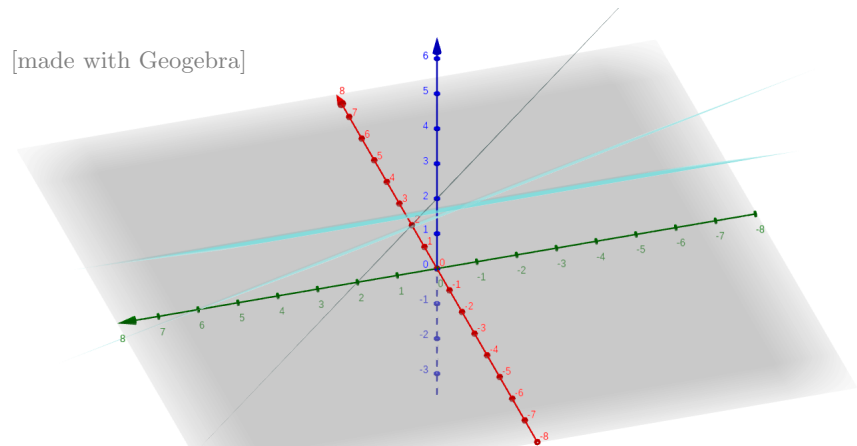
Es decir, calculando: “Ecuación 1” + “Ecuación 2” – “Ecuación 3” nos da:



Lo anterior contradice $0 \neq 1$, por lo cual se dice que el sistema es **inconsistente**.

1. Sistemas de Ecuaciones Lineales

Aquí, un dibujo (*plot*) es difícil porque solo de dos perspectivas se ve que la intersección de cada pareja de 2 planos es una línea que no se interseca con las otras. Nuestro mejor intento:



Desde un punto de vista de columnas no podemos encontrar una combinación lineal de las columnas (forma de combinar las columnas) y determinar los números x, y, z para obtener el vector $\vec{b}$.

Esto es:



Ejemplo: (Con una infinidad de soluciones). Del ejemplo anterior tomamos el sistema y cambiamos el lado derecho (cambiamos 6 por 7). Es decir, tenemos el siguiente nuevo sistema:

$$x + y + z = 2$$

$$2x + 3z = 5$$

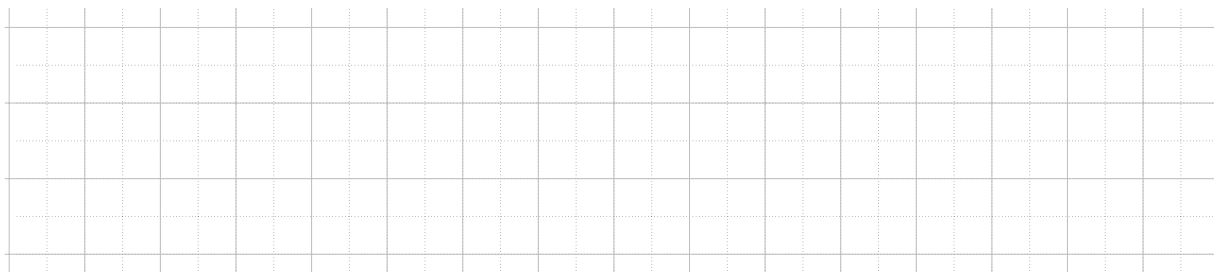
$$3x + y + 4z = 7$$

Del nuevo, calculando $\text{Fila}_1 + \text{Fila}_2 - \text{Fila}_3$ tenemos:



Interpretación: Cada punto que satisface las primeras dos ecuaciones satisface la tercera ecuación. Por lo tanto, la intersección de los primeros dos planos (cada punto de una línea) satisface la tercera ecuación y el conjunto solución es una línea (una infinidad de puntos).

La visualización por columna es:



Usando Geogebra podemos ver que el vector $\vec{b}$ (lado derecho) ahora pertenece al plano generado por las primeras dos columnas y los vectores $\vec{u}, \vec{v}, \vec{w}, \vec{b}$:

`Plane( (0,0,0), (1,2,3), (1,0,1) )`

`u = (1,2,3)`

`v = (1,0,1)`

`w = (1,3,4)`

`b = (2,5,7)`

1.2. Representación matricial

Por lo visto, tenemos dos formas de escribir el sistema, una por filas y otra por columnas. La existencia de soluciones depende de los coeficientes (números multiplicados con las incógnitas) y del lado derecho, pero no de la forma de escribir el sistema. Necesitamos el concepto de matrices para representar las dos formas de escribir el sistema y para estudiar existencia de soluciones.

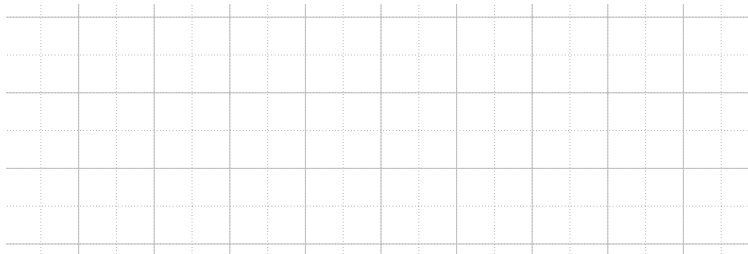


Definición. Una **matriz (real)** de tamaño



En corto: Muchos autores escriben $A \in \mathcal{M}(m \times n)$ o $A \in \mathbb{R}^{m \times n}$.

Ejemplo: Una matriz $A \in \mathbb{R}^{3 \times 3}$, es decir, una matriz (real) A de tamaño 3×3 puede ser



En PYTHON: Una forma de introducir matrices es definir un arreglo:

```
import numpy
A = numpy.array( [ [1,2,3], [2,5,2], [6,-3,1] ] )
```

Acceso a entradas:

- Si quiero una entrada de un arreglo (matriz), tengo que poner la matriz seguida de unas paréntesis `[ ]` con el número de fila y columna. Notar que PYTHON empieza contar filas y columnas con cero, por ejemplo, `A[2,1]` indica la entrada que esta en la tercera fila (de las filas $\{0, 1, 2\}$) y en la segunda columna de $\{0, 1, 2\}$.
- Si quiero toda la primera fila ocupo `A[0,:]` y para la fila i ocupo `A[i,:]` y para la columna j ocupo `A[:,j]`.



En nuestro curso para $A \in \mathbb{R}^{m \times n}$ usaremos la siguiente notación parecida :

Ejemplo: Si $C \in \mathbb{R}^{3 \times 4}$ es la matriz $C = \begin{bmatrix} 1 & 0 & 6 & 2 \\ 7 & 8 & 3 & 1 \\ 9 & 6 & 4 & 3 \end{bmatrix}$, entonces escribimos para el segundo renglón y la cuarta columna:



En Álgebra es **importante** notar lo siguiente:

- El renglón $C(2, :) \in \mathbb{R}^{1 \times 4}$ es un **vector renglón** con 4 entradas y (a la vez) **una matriz** con 1 renglón y 4 columnas.
- La columna $C(:, 4) \in \mathbb{R}^{3 \times 1}$ es un **vector columna** con 3 entradas y (a la vez) **una matriz** con 3 renglones y 1 columna.

1.2.1. La forma matricial de un sistema lineal

Recordamos y usamos el producto punto de “Geometría Analítica”, por ejemplo, el producto punto entre los vectores $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ y $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ es

Ejemplo: Usando la idea anterior, representamos el sistema

$$x + 2y + 3z = 6$$

$$2x + 5y + 2z = 4$$

$$6x - 3y + z = 2$$

1. Sistemas de Ecuaciones Lineales

mediante un producto de una matriz por un vector igualado al vector $\vec{b}$ del lado derecho. El producto de matriz por vector viene del producto (punto) de un vector fila por el vector columna $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$.

Por varias razones (“flojera”, “ahorrar papel”, ...) las operaciones anteriores se escriben utilizando la matriz, es decir, este sistema en **forma matricial** se ve como:

Escribimos el sistema nuevamente por columnas:



En esta página se reconocen 3 formas de escribir el mismo sistema de 3 ecuaciones lineales:

- La forma matricial $A\vec{x} = \vec{b}$ es la más usada de escribir un sistema lineal.
- La primera forma arriba (deja reconocer la forma por filas) y nos dice como calcular el producto de una matriz A con el vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, es decir, el producto $A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que aparece en la forma matricial del sistema.
- La representación por columnas nos dice **una segunda forma** de calcular el producto $A \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, es decir, primero multiplicamos x con la primera columna de A , y con la segunda y z con la tercera y después sumamos estos vectores escalados.
- La primera forma también permite paralelizar el calculo de productos “matriz” - “vector” y así acelerar algoritmos que usan este producto.



Ejemplo*: El sistema escrito por columnas permite ver con cual vector $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ debes multiplicar cuando quieres extraer una columna de la matriz A . ¿Con cuál vector sacas la primera, segunda y tercera columna de A ?

Ejemplo: (Producto matriz-vector).

¿Cuánto vale el vector $\vec{b}$ tal que $A\vec{x} = \vec{b}$ si la matriz A y el vector $\vec{x}$ son dados por:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad \& \quad \vec{x} = \begin{bmatrix} 1 \\ 4 \\ 9 \end{bmatrix} \quad ?$$

La matriz A es especial porque el producto $A\vec{x}$ nos da la diferencia de las coordenadas de $\vec{x}$:

Además, el siguiente ejemplo muestra que un sistema con la matriz anterior es fácil de resolver.

Ejemplo: (Resolver un sistema con la matriz A del ejemplo anterior). Encuentre el vector $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que es solución de $A\vec{x} = \vec{b}$ cuando $\vec{b} = \begin{bmatrix} 1 \\ 5 \\ 8 \end{bmatrix}$.



La solución fue fácil de encontrar porque la matriz A es un tipo de matriz que llamaremos **triangular inferior** $A = \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$, es decir, las entradas de A arriba de su diagonal son ceros.

Matrices especiales:

Existen matrices con características especiales.

- “Multiplicar un número con uno no cambia el número”.

Para “preservar” matrices y vectores el “UNO” es la **matriz identidad**, denotada I_n donde n indica el número de renglones y columnas y los elementos de la diagonal son 1's y fuera de ella son cero. Por ejemplo:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad I_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

En PYTHON: `numpy.eye(n)` construye la matriz I_n .

El nombre “matriz identidad” es porque, el producto $I_n\vec{x}$ siempre da $\vec{x}$ y si buscamos una solución $\vec{x}$ del sistema $I_n\vec{x} = \vec{b}$ entonces esa solución será $\vec{b}$.

- Algo especial son matrices cuyas entradas son cero o el vector $\vec{0} \in \mathbb{R}^n$ (el origen), por ejemplo

$$\mathcal{O}_{3 \times 5} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathcal{O}_{2 \times 1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \vec{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

En PYTHON: `numpy.zeros( (3,5) )` o `numpy.zeros( (2,1) )` construyen esas matrices.

Si multiplican un vector con estas matrices, siempre les dará cero o un vector de ceros.

- La **matriz de rotación** de α grados en sentido contrario a las manecillas del reloj es:

$$R_{\alpha} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

Ejemplo: Rotación por $\frac{\pi}{2}$ ó 90° del triángulo ABC con $A = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $B = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $C = \begin{bmatrix} -3 \\ 2 \end{bmatrix}$.



- La **matriz de reflexión** en la línea $y = x \tan \frac{\alpha}{2}$:

$$S_{\alpha} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix} \in \mathbb{R}^{2 \times 2}$$

Ejercicio: Usa la matriz de Reflexión con $\alpha = 90^\circ$ para ver que pasa con el triángulo.



Hasta ahora hemos visto sistemas de 2×2 & 3×3 , es decir, sistemas de 2 ecuaciones e incógnitas & 3 ecuaciones e incógnitas. A continuación veremos sistemas con m ecuaciones y n incógnitas.

1.3.1. Sistemas equivalentes



Definición. Se dice que dos sistemas S & S' son **equivalentes** denotado $S \sim S'$ si tienen el mismo conjunto solución. Esto es, cada solución del sistema S es solución de S' y al revés.

Ejemplos:

a) Intercambio de ecuaciones no cambia soluciones:

$$S = \begin{cases} x_1 - x_2 + 2x_3 = 1 \\ -x_1 + 2x_2 - x_3 = -1 \\ x_1 - 2x_2 - x_3 = 3 \end{cases} \quad S' = \begin{cases} \text{[Grid]} \end{cases}$$

b) Multiplicar una ecuación por una constante **distinta de cero** no cambia. Las siguientes sistemas son equivalentes, ya que ambos tienen la única solución $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$:

$$S = \begin{cases} x_1 - x_2 = -1 \\ x_2 = 3 \end{cases} \quad S' = \begin{cases} \text{[Grid]} \end{cases}$$

c) Los siguientes sistemas son equivalentes, ya que la única solución es $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ tanto para S como para S' .

$$S = \begin{cases} x_1 - x_2 = 3 \\ 2x_1 - x_2 = 5 \end{cases} \quad S' = \begin{cases} \text{[Grid]} \end{cases}$$

Note que tanto S como S' tienen la misma primera ecuación y que

“segunda ecuación de S' ” = “segunda ec. de S ” – “2 veces la primera ec. de S ”.

En los tres ejemplos anteriores hemos obtenido sistemas equivalentes, usando operaciones que nos llevan del primer sistema al segundo sistema. A estas operaciones se les conoce como operaciones elementales & las definimos formalmente.

1.3.2. Operaciones elementales



Definición. Se llaman **Operaciones Elementales** de ecuaciones o filas a las siguientes tres operaciones:

- (I) Intercambiar (permutar) dos ecuaciones (filas).
- (II) Multiplicar una ecuación por un número **diferente de cero**.
- (III) Sumar un múltiplo de una ecuación a **otra** ecuación.

1. Sistemas de Ecuaciones Lineales

A continuación presentamos unos de los teoremas mas importantes dentro del Álgebra Lineal que nos permitirá llevar un sistema lineal S a un sistema equivalente a través de operaciones elementales.



Teorema 1.1. Si S es un sistema y se realiza una operación elemental para obtener un sistema S' , entonces ambos sistemas son equivalentes.

Demostración. Sea $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$ una tupla que es solución del sistema S . Se tienen que considerar tres casos. Solamente efectuamos uno de ellos, ya que es el que mejor ejemplifica la demostración.

Caso (II), la segunda operación elemental.

Consideremos un número $i \in \{1, \dots, m\}$ fijo. Además, consideremos la i -ésima ecuación del sistema S , es decir, consideremos

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n = b_i.$$

Si multiplicamos la ecuación anterior por un número $t \neq 0$ se obtiene la ecuación i del segundo sistema S' , esto es:

$$t a_{i1}x_1 + t a_{i2}x_2 + \dots + t a_{in}x_n = t b_i.$$

Como $\vec{u}$ es solución de S , entonces se tiene que $\vec{u}$ satisface:

$$t a_{i1}u_1 + t a_{i2}u_2 + \dots + t a_{in}u_n = t (a_{i1}u_1 + a_{i2}u_2 + \dots + a_{in}u_n) = t b_i.$$

Concluimos que $\vec{u}$ también es solución de S' .

De forma similar se demuestra el caso (III) y el caso (I) se considera fácil. □



Dada un sistema lineal S , el teorema anterior permite generar una sucesión de sistemas lineales donde el último sistema es muy fácil de resolver. Las operaciones elementales se usan en algoritmos “Eliminación Gaussiana”, “método de Gauss-Jordan” y en futuras materias.

Ejemplo*: Aplique el teorema 1.1 al siguiente sistema para obtener sistemas equivalentes y determine el conjunto solución.

Solución: Denotaremos por $R_1 :=$ renglón 1, $R_2 :=$ renglón 2 & $R_3 :=$ renglón 3. Sea S el siguiente sistema de ecuaciones:

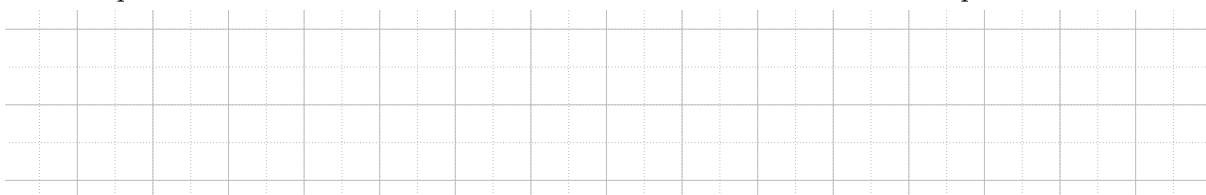
$$S = \begin{cases} 2x_1 + 2x_2 - 6x_3 = 6 \\ x_1 - 3x_2 + 2x_3 = 1 \\ x_1 + 5x_2 - 8x_3 = 5 \end{cases}$$

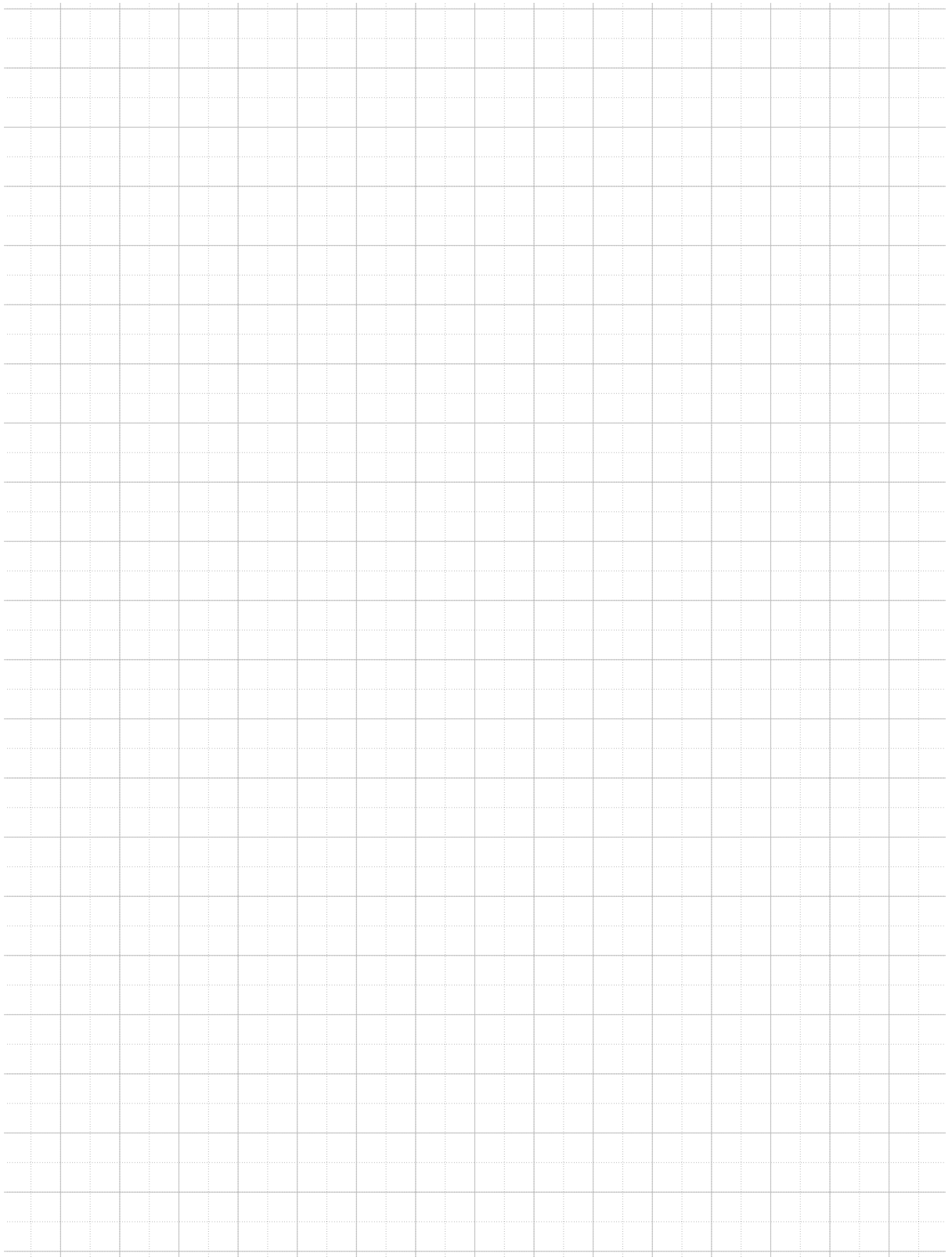
izquierda:

Sistemas equivalentes $<$

derecha:

$>$ lo que realmente cambia





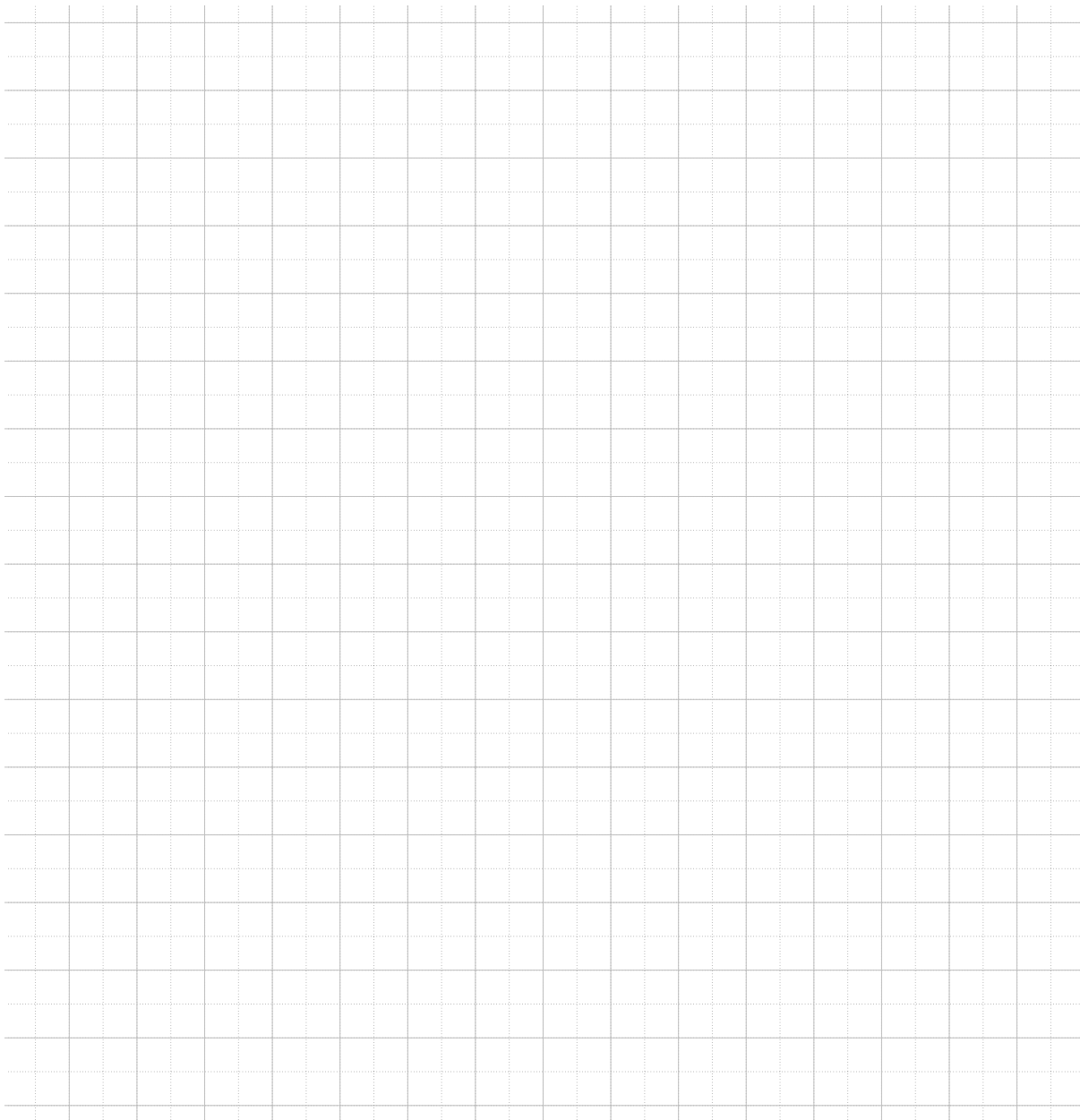
Del lado derecho vimos para cada sistema equivalente una forma corta [“matriz” | “lado derecho”] de representarlo. Esta forma $[A | \vec{b}]$ será más tarde la **matriz aumentada** del sistema $A\vec{x} = \vec{b}$.

1. Sistemas de Ecuaciones Lineales

Observe que S' tiene 2 ecuaciones y 3 incógnitas, es decir, el conjunto solución del sistema S es igual a la intersección de estos dos planos (una recta en el espacio). En otras palabras se dice que tenemos una $(1 = 3 - 2)$ **variable libre**:

$$S' = \begin{cases} x_1 + x_2 - 3x_3 = 3 \\ x_2 - \frac{5}{4}x_3 = \frac{1}{2} \end{cases} \iff S' = \left\{ \begin{array}{c} \text{ } \\ \text{ } \\ \text{ } \end{array} \right.$$

Si consideramos a la variable libre x_3 igual a un parámetro $t \in \mathbb{R}$, entonces sustituciones hacia atrás permiten despejar las otras variables en función de t :



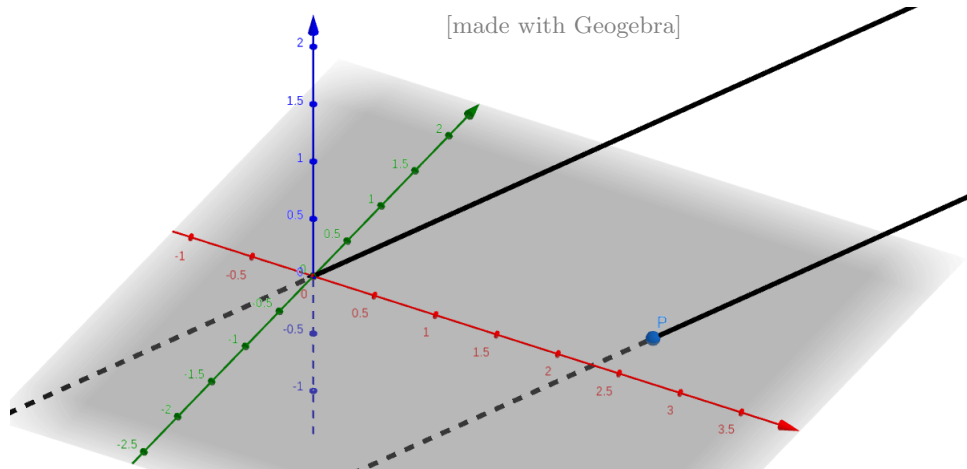
Nuestro conjunto solución del problema anterior es el siguiente:

$$\mathcal{L}_s = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5/2 \\ 1/2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 7/4 \\ 5/4 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\} = \{ \vec{p} + t\vec{v} : t \in \mathbb{R} \}.$$

Ejercicio: Verifique que $A\vec{p} = \vec{b}$ y $A\vec{v} = \vec{0}$ donde $A \in \mathbb{R}^{3 \times 3}$ es la matriz original del ejemplo.



Concepto general: El vector $\vec{p}$ es una **solución particular** de $A\vec{x} = \vec{b}$ y el vector $\vec{v}$ es una solución del sistema homogéneo $A\vec{x} = \vec{0}$.



Aplicación de programación lineal: El siguiente problema de minimización equivale a minimizar una función en una variable:

$$\begin{aligned} &\text{minimizar} && x_1 - x_2 - x_3 \\ &\text{sujeto a} && 2x_1 + 2x_2 - 6x_3 = 6 \\ &&& x_1 - 3x_2 + 2x_3 = 1 \\ &&& x_1 + 5x_2 - 8x_3 = 5 \\ &&& x_1, x_2 \geq 0 \quad \& \quad 0 \leq x_3 \leq 4 \end{aligned}$$

Las soluciones del sistema lineal $A\vec{x} = \vec{b}$ pertenecen al conjunto

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathcal{L}_s = \left\{ \begin{bmatrix} 5/2 \\ 1/2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 7/4 \\ 5/4 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\} = \{ \vec{p} + t\vec{v} : t \in \mathbb{R} \}.$$

Además, las desigualdades $0 \leq x_3 \leq 4$ limitan la variable t al intervalo $[0, 4]$. Junto, la estructura de la solución y la restricción $t \in [0, 4]$ nos permiten escribir el problema de la siguiente forma:

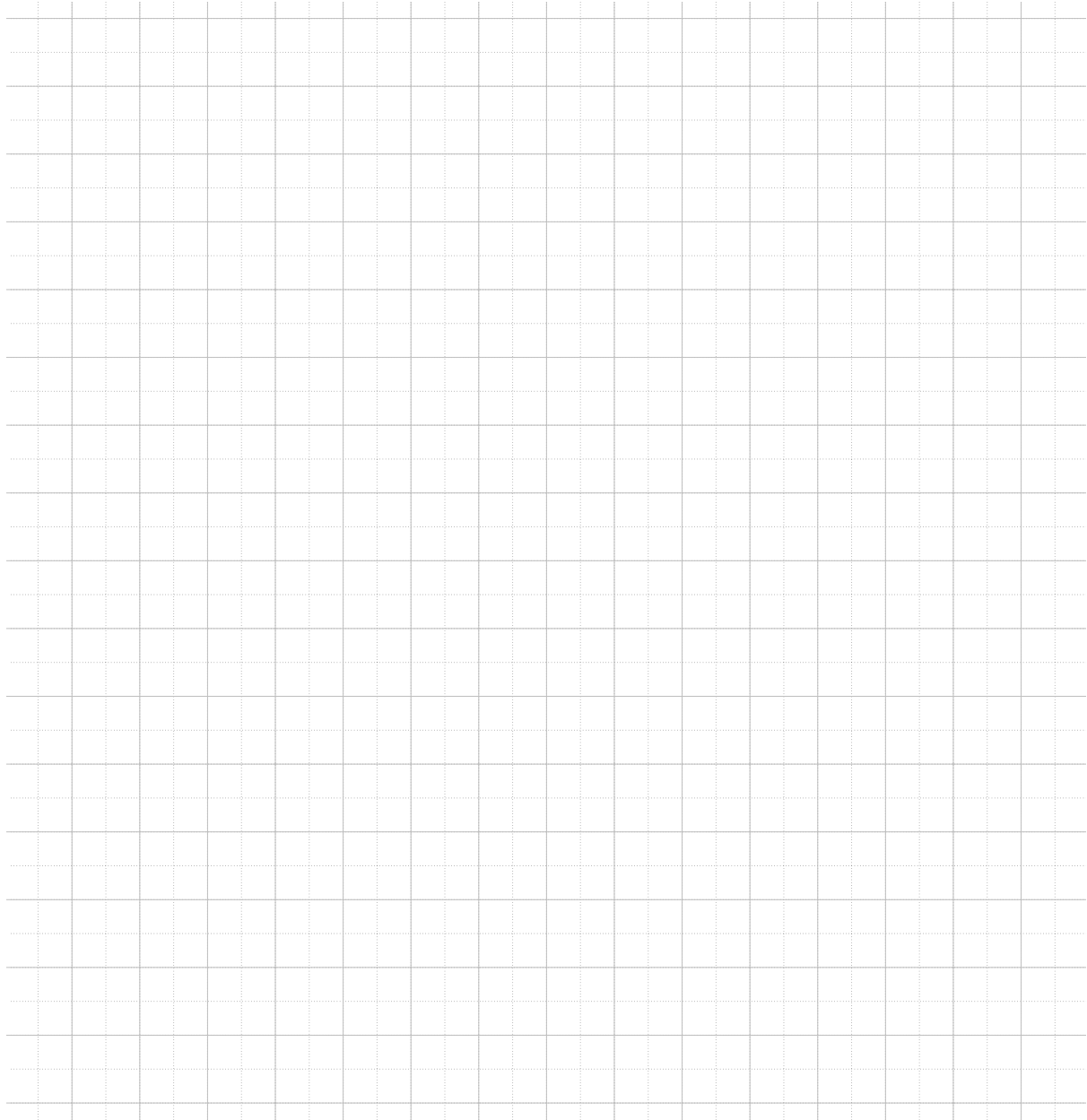
$$\min_{t \in [0, 4]} \left(\frac{5}{2} + t\frac{7}{4} \right) - \left(\frac{1}{2} + t\frac{5}{4} \right) - t.$$



Definición. Para cada sistema lineal $A\vec{x} = \vec{b}$ la matriz $[A|\vec{b}]$ se llama **matriz aumentada**.

Ejemplo: (Sistema inconsistente). Sea S el siguiente sistema lineal, sobre el cual efectuaremos operaciones elementales (página 19). El objetivo es llevar el sistema a uno con matriz triangular:

$$S = \begin{cases} 3x_1 - 4x_2 + 5x_3 = 0 \\ 2x_1 + x_2 - 2x_3 = -2 \\ x_1 - 5x_2 + 7x_3 = 3 \end{cases}$$



Para efectuar la 3ª operación elemental es conveniente que en una fila (alguna ecuación) el primer coeficiente distinto de cero es 1. Así se pueden evitar algunas fracciones (como apenas se hizo).



Definición. Un sistema lineal es **consistente** si y solo si (syss) tiene al menos una solución. Si el conjunto solución es vacío, entonces diremos que el sistema es **inconsistente**.

1.4. Eliminación Gaussiana y el método de Gauss-Jordan

Hemos visto como transformar el sistema lineal original a un sistema equivalente con operaciones elementales. Vimos que esas operaciones se pueden efectuar en algún orden flexible (depende de la persona). Una persona (también un algoritmo) debe tener un objetivo “en mente”, *e.g.* un sistema estándar (fácil de resolver) al que quiere llegar.

Esa forma estándar se conoce como *forma escalonada (reducida)* y la definimos en seguida.



Definición 1.1. Una matriz está en **forma escalonada (FE)** syss se cumplen las siguientes tres condiciones:

1. Si la matriz tiene renglones que consisten únicamente de ceros, esos renglones están en su parte inferior.
2. Si un renglón tiene coeficientes distintos de cero, entonces el primer (de izquierda a derecha) de esos coeficientes distintos de cero es igual a uno. Ese uno es llamado un **uno principal**.
3. Cada uno principal aparece a la derecha y por debajo de los unos principales de los renglones precedentes (por arriba).

Una matriz está en **forma escalonada reducida (FER)** si esta en FE y además satisface:

4. Cada uno principal es el único elemento diferente de cero dentro de la columna en la cual se encuentra.

Ejemplos:





La tercera condición de la definición 1.1 implica que si $A \in \mathbb{R}^{m \times n}$ está en forma escalonada, entonces su número de unos principales es menor o igual al mínimo entre m y n .



Definición. Sea $A \in \mathbb{R}^{m \times n}$ una matriz. Se dice que $B \in \mathbb{R}^{m \times n}$ es una **forma escalonada de A** si B está en FE y B se obtiene aplicando operaciones elementales empezando con A .

Se dice que $B \in \mathbb{R}^{m \times n}$ es la **forma escalonada reducida de A** si B está en FER y B se obtiene aplicando operaciones elementales empezando con A .

Con ejemplos se puede ver que una matriz puede tener varias FE. Por otro lado, se necesita demostrar que cualquier matriz tiene una y solo una FER, ver por ejemplo el Teorema 2.28 en el libro de Mota & Rumbos 2020, [MR20].

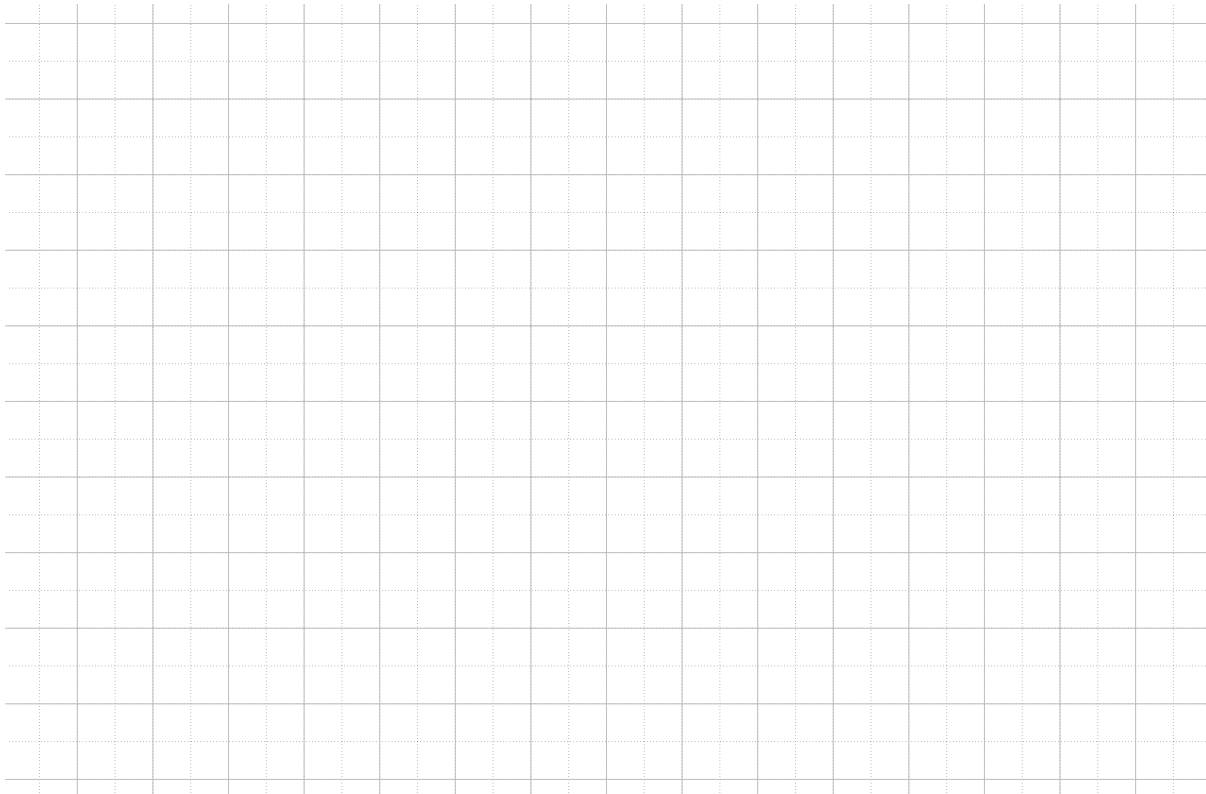


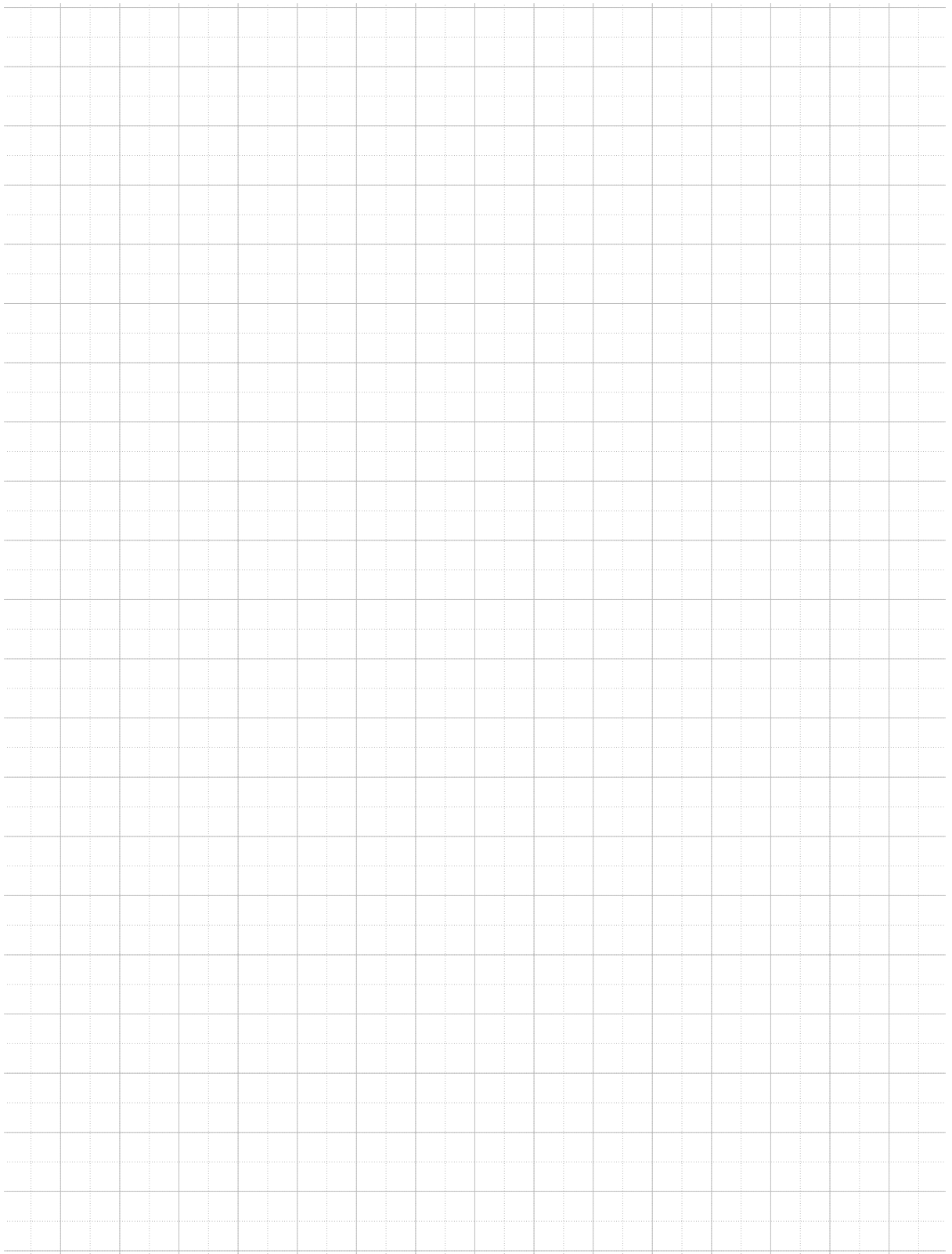
Definición. Eliminación Gaussiana toma una matriz aumentada $[A | \vec{b}]$ de un sistema lineal y efectúa operaciones elementales (pág. 19) hasta llegar a **una** forma escalonada (definición 1.1).

El **método de Gauss-Jordan** continua este proceso (efectuando más operaciones elementales) para llegar a **su** forma escalonada reducida.

Ejemplo: (FE y FER con infinidad de soluciones). Usando operaciones elementales convierte la matriz aumentada del siguiente sistema (a) en FE y (b) en FER. De ambas formas determine su conjunto solución.

$$S = \begin{cases} 2x_1 + 2x_2 - 6x_3 = 6 \\ x_1 - 3x_2 + 2x_3 = 1 \\ x_1 + 5x_2 - 8x_3 = 5 \end{cases}$$





Definición. Piensa en una FE de una matriz. Si x_i es multiplicada con un UNO principal, entonces le decimos **variable básica** (ó dependiente). En otro caso, le decimos **variable libre**.

La FER anterior indica que $\{x_1, x_2\}$ son variables básicas y x_3 es variable libre.

1. Sistemas de Ecuaciones Lineales

Ejemplo: Encuentre la solución general del sistema lineal cuya matriz aumentada se ha reducido a:

$$\left[\begin{array}{ccccc|c} 0 & 0 & 2 & -8 & -1 & 3 \\ 1 & 6 & 2 & -5 & -2 & -4 \\ 0 & 0 & 0 & 0 & 1 & 7 \end{array} \right]$$

Observe que solo falta un 1 principal y un cambio de renglones para estar en FE.





1.4.1. Los pasos del método de Gauss-Jordan.

Hago un intento de describir lo que son los pasos. Unos pasos parecidos se encuentran en las paginas 24-25 del libro de Mota & Rumbos (2020).

-
- *Paso 0:* Definir la matriz aumentada del sistema lineal y el renglón actual ($r = 1$).
 - *Paso 1:* Localizar la primera columna (de izquierda a derecha) con entradas diferentes de cero en renglones $i \geq r$ (es decir, en renglón r o por debajo de él) y elegir alguna entrada distinta de cero en esa columna. Supongamos que dicha entrada está en el i -ésimo renglón.
Si no existe una entrada diferente de cero, entonces el algoritmo termina aquí.
 - *Paso 2:* Cambiar renglón i con el renglón actual r . Después, la columna actual (fijada en paso 1) tendrá una entrada diferente de 0 en el renglón r .
 - *Paso 3:* Generar un 1 principal en el renglón actual r , es decir, usar la 2^a operación elemental para dividir ese renglón entre su primer elemento diferente de cero (debido al paso 2).
 - *Paso 4:* (Solo si $r < m$, es decir, r no es el último renglón.) Aplicar la 3^a operación elemental (pág. 19) para eliminar entradas diferentes de cero en la columna actual (fijado en paso 1) y por debajo del 1 principal del renglón actual r (generado en paso 3).
 - *Paso 5:* Si $r < m$, aumentar “ $r \mapsto r + 1$ ” para usar el siguiente renglón como el actual y repetir los pasos 1 a 4. La siguiente “columna actual” estará a la derecha de esta.
Cuando $r = m$ hemos llegado a una FE de la matriz inicial.
 - *Paso 6:* Usar la 3^a operación elemental (pág. 19) para eliminar entradas diferente de ceros en los renglones arriba de los unos principales.
Con este paso, llegaremos a la FER de la matriz inicial.
-



1. Sistemas de Ecuaciones Lineales

Ejemplo: (del Libro Mota & Rumbos) Considere el siguiente sistema y determine cuales condiciones deben cumplir a, b, c para que el sistema sea consistente:

$$3x - y = a$$

$$x - y = b$$

$$2x + y = c$$



1.4.2. Resumen

Después de aplicar sucesivamente las **tres** operaciones elementales (página 19) al sistema:

$$(S) \left\{ \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{array} \right.$$

se obtiene un sistema S' que esta en forma escalonada (FE) con $r \leq \min\{m, n\}$ unos principales. Para simplificar la discusión de los hechos relevantes supongamos que los unos principales se ubican en las primeras r columnas de S' :

$$(S') \left\{ \begin{array}{l} a'_{11}x_1 + a'_{12}x_2 + \dots + a'_{1r}x_r + \dots + a'_{1n}x_n = b'_1 \\ a'_{22}x_2 + \dots + a'_{2r}x_r + \dots + a'_{2n}x_n = b'_2 \\ \vdots \\ a'_{rr}x_r + \dots + a'_{rn}x_n = b'_r \\ 0 = b'_{r+1} \\ \vdots \\ 0 = b'_m \end{array} \right.$$

donde $a'_{11} = a'_{22} = \dots = a'_{rr} = 1$ (por estar en FE).

Todas las (posibles) formas escalonadas de una matriz comparten las siguientes características:

1. Cada FE es el resultado de aplicar eliminación Gaussiana (las **tres** operaciones elementales, pág. 19) y tiene el mismo conjunto solución que el sistema original (por teorema 1.1).
2. El número $r \leq \min\{m, n\}$ de unos principales es igual al de la FER del sistema, ya que la FER se obtiene eliminando coeficientes arriba de los unos principales por medio de la tercera operaciones elemental (pág. 19), es decir, haciendo el último paso del método de Gauss-Jordan (pág. 29).

Conceptos prevalentes:



Definición. Para una matriz $A \in \mathbb{R}^{m \times n}$, es decir, con m renglones y n columnas, el número r de unos principales en una FE(R) de A se llama **rango de A** y siempre satisface $r \leq \min\{m, n\}$.



En este curso: Las incógnitas que se multiplican con (las columnas que contienen) los unos principales de la FER serán **variables básicas** (variables dependientes) y las $(n - r)$ variables restantes serán **variables libres**.



Caracterización de sistemas y sus conjuntos de solución:

	$r = m$	$r < m$ $b'_{r+1} = \dots = b'_m = 0$	$r < m$ $\exists b'_i \neq 0$ p.a. $i > r$
$r = n$	solución única	solución única	$\mathcal{L}_s = \emptyset$
$r < n$	varias soluciones y $(n - r)$ variables libres	varias soluciones y $(n - r)$ variables libres	$\mathcal{L}_s = \emptyset$

Ejemplos vistos: Con solución única (en páginas 1–6), sin solución (en páginas 7–10 y 24) y con muchas soluciones ($r < n$ en páginas 11, 20–23, 26–27).

Ejemplos vistos con algunos detalles:

Página	m	n	r	variables libres	$\mathcal{L}_s$
20–21	3	3			
24					
25 (la FER)	4	8			
26–27					
28					
30					



Consecuencias de operaciones elementales para **sistemas homogéneos**:

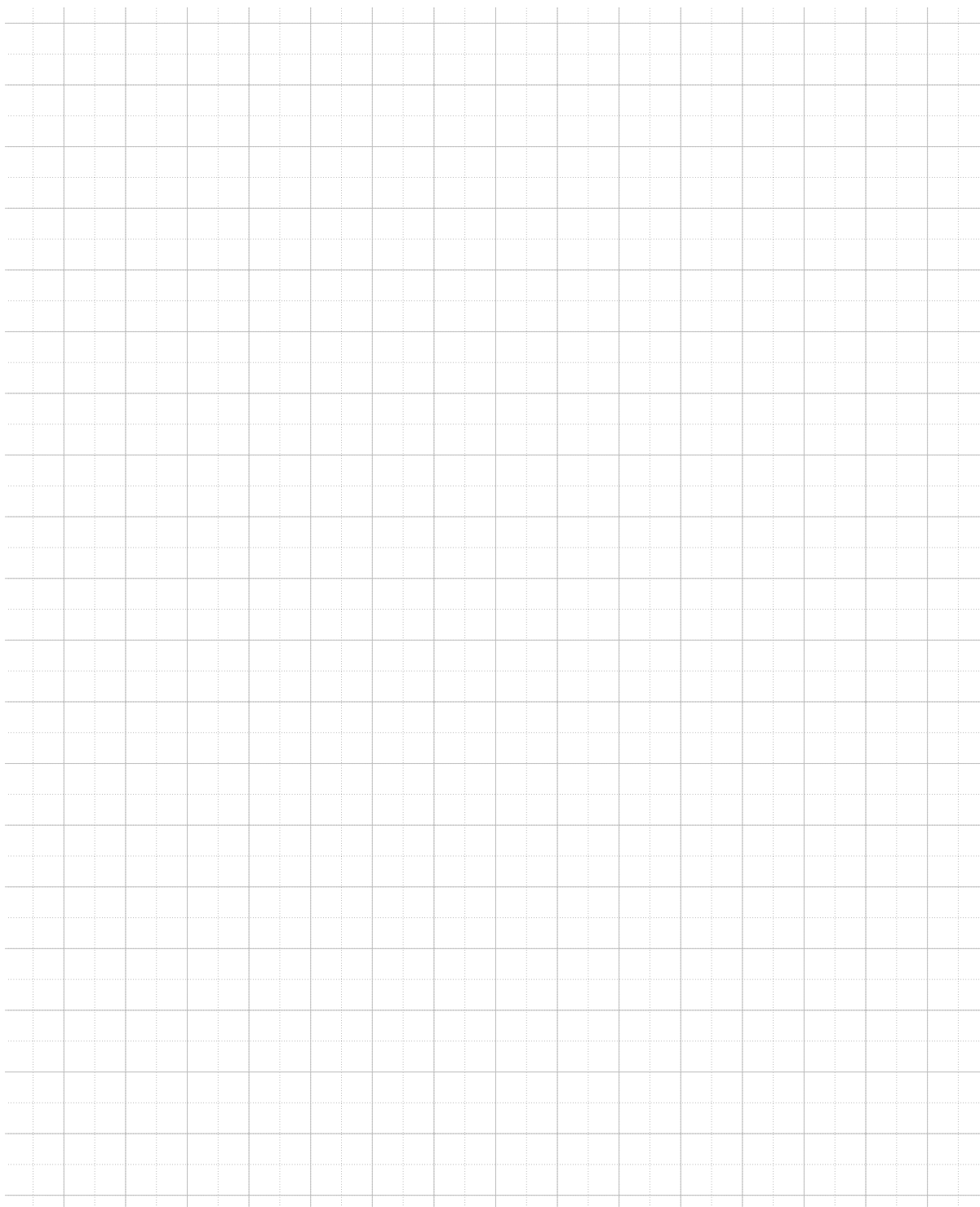
- Si un sistema es homogéneo, entonces cualquier sistema equivalente será homogéneo.
- Un sistema homogéneo siempre tiene por lo menos una solución, la **solución trivial** que es el vector de ceros $\vec{0}$.
- Si el sistema homogéneo satisface $r < n$, es decir, si en la FER hay menos ecuaciones que incógnitas, entonces el sistema también tiene soluciones no-triviales.

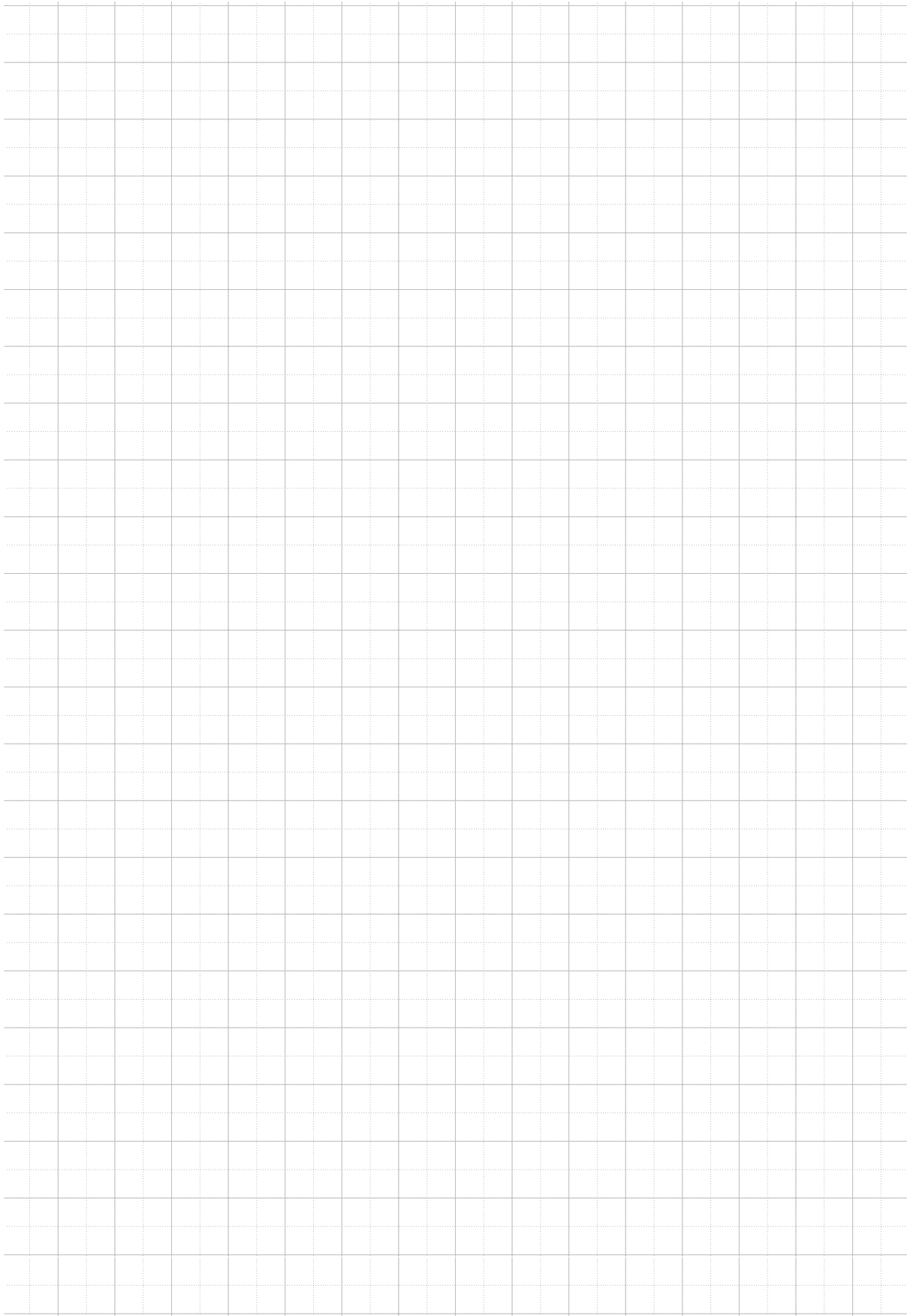
Ejercicio: Resolviendo el siguiente sistema te darás cuenta que los tres puntos son ciertos:

$$\begin{aligned}
 w + x + y + z &= 0 \\
 w + x &= 0 \\
 x + 2y + z &= 0
 \end{aligned}
 \quad \text{La solución es: } \mathcal{L}_s = \left\{ t \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} : t \in \mathbb{R} \right\} \subset \mathbb{R}^4$$

Ejemplo: Encontrar los valores de b que hacen que el siguiente sistema (dado en forma de su matriz aumentada) tenga solución, infinitas soluciones & ninguna solución:

$$\left[\begin{array}{ccc|c} 1 & b & -2 & 2 \\ -1 & (b-2) & 2 & -2 \\ 2 & 2 & (b-4) & 3 \end{array} \right]$$





 **Lema 1.2.** Sea S_H un sistema homogéneo con n incógnitas. El conjunto de soluciones de S_H satisface la **propiedad de cerradura** “bajo combinaciones lineales”. Es decir, si conocemos dos vectores (n -tuplas) $\vec{u} \in \mathbb{R}^n$ y $\vec{v} \in \mathbb{R}^n$ que son soluciones del mismo sistema homogéneo S_H , entonces las combinaciones lineales $\alpha\vec{u} + \beta\vec{v}$ para $\alpha, \beta \in \mathbb{R}$ también serán soluciones de S_H .



Para cada fila $i \in \{1, \dots, m\}$ hemos demostrado: $a_{i1} w_1 + a_{i2} w_2 + \dots + a_{in} w_n = 0$, lo cual equivale a $A\vec{w} = \vec{0}$. Por lo tanto, la combinación lineal $\alpha\vec{u} + \beta\vec{v}$ es solución de S_H para todo $\alpha, \beta \in \mathbb{R}$. $\square$



Nota: Los argumentos de esta demostración también sirven para demostrar $A(\alpha\vec{u}) = \alpha(A\vec{u})$, “podemos sacar factores escalares $\alpha \in \mathbb{R}$ ” y $A(\alpha\vec{u} + \beta\vec{v}) = \alpha A\vec{u} + \beta A\vec{v}$, “una ley distributiva”.



Definición. Dados n escalares $x_1, \dots, x_n \in \mathbb{R}$ y los siguientes n -vectores de m -entradas cada uno

$$\vec{v}_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{m1} \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{m2} \end{pmatrix}, \dots, \quad \vec{v}_n = \begin{pmatrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{mn} \end{pmatrix}$$

definimos la **combinación lineal** de ellos como $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_n\vec{v}_n$, es decir

$$\begin{aligned} x_1 \begin{pmatrix} v_{11} \\ v_{21} \\ \vdots \\ v_{m1} \end{pmatrix} + x_2 \begin{pmatrix} v_{12} \\ v_{22} \\ \vdots \\ v_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} v_{1n} \\ v_{2n} \\ \vdots \\ v_{mn} \end{pmatrix} &= \begin{pmatrix} x_1 v_{11} + x_2 v_{12} + \dots + x_n v_{1n} \\ x_1 v_{21} + x_2 v_{22} + \dots + x_n v_{2n} \\ \vdots \\ x_1 v_{m1} + x_2 v_{m2} + \dots + x_n v_{mn} \end{pmatrix} \\ &= \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \end{aligned}$$

En términos generales una combinación lineal de vectores es la suma de los vectores una vez multiplicados por los escalares.

Ejemplo: Escribe el lado izquierdo del siguiente sistema homogéneo (a) como combinación lineal de vectores, (b) como producto matrix-vector. Después, resuelva el sistema:

$$x_1 + 2x_2 + 3x_3 - x_4 = 0$$

$$-x_1 - x_2 - x_3 - x_4 = 0$$



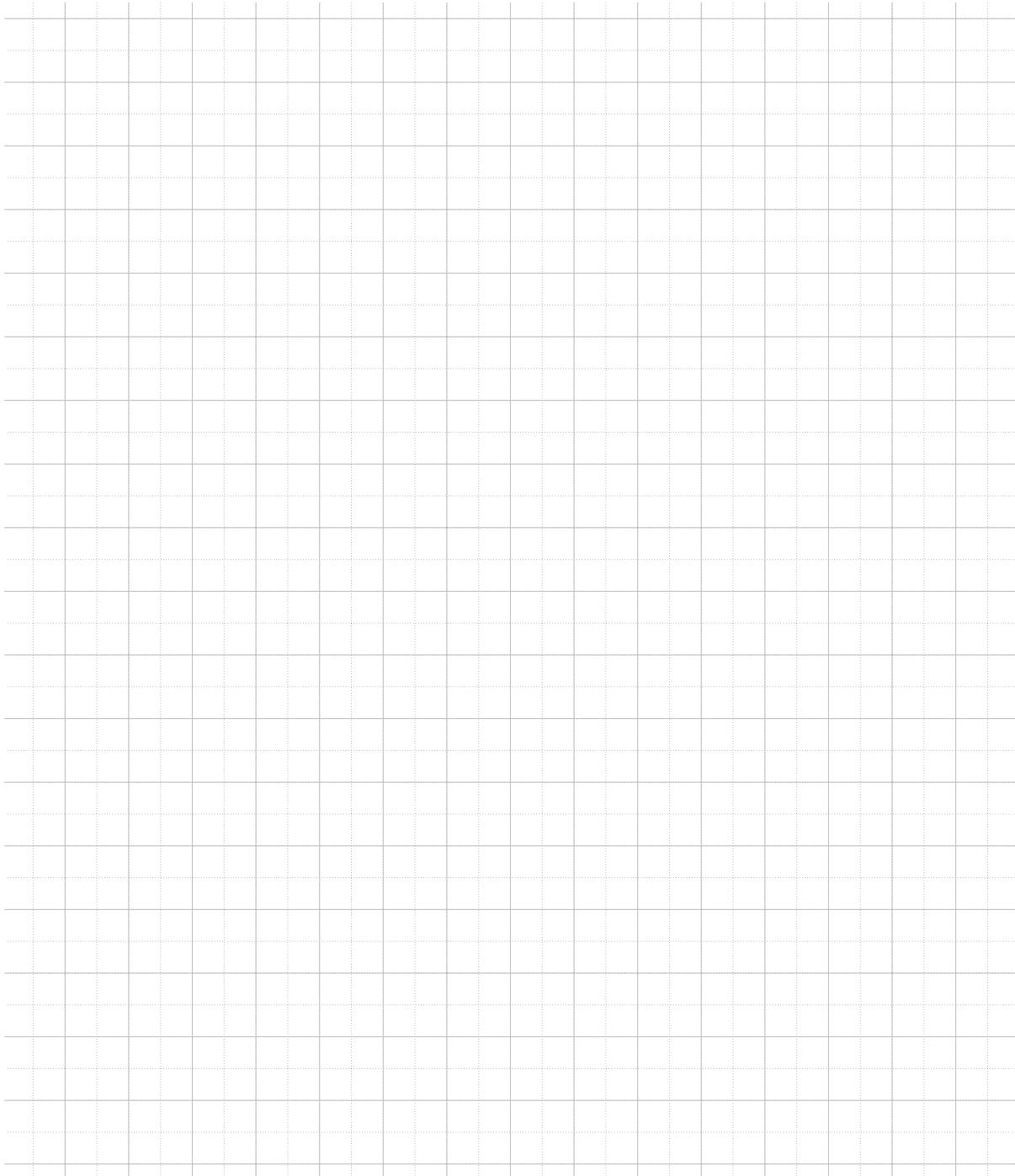
Ese conjunto solución ejemplifica lema 1.2.

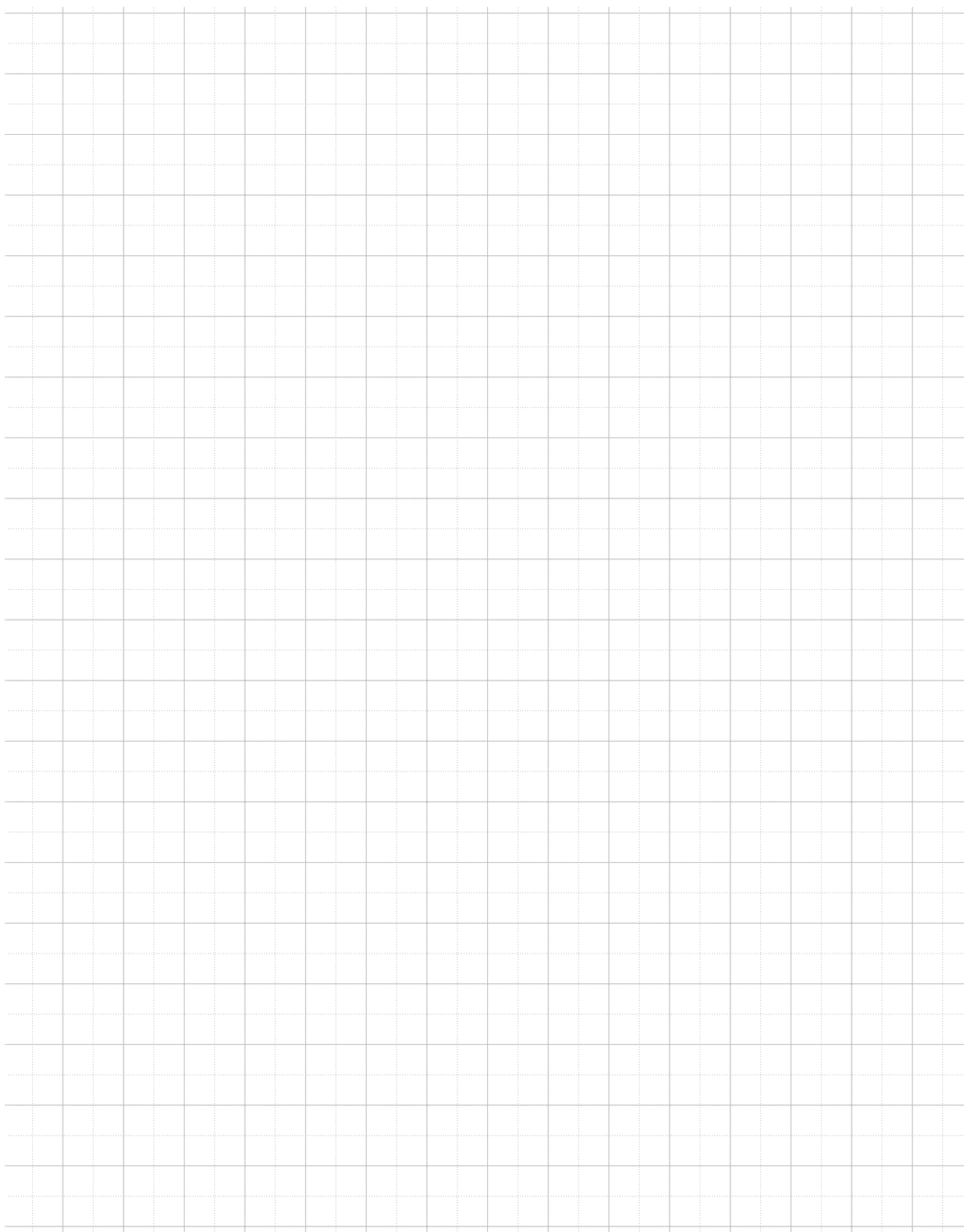


1.6. Aplicaciones de sistemas de ecuaciones lineales

En esta sección presentaremos problemas que pueden ser modelados mediante sistemas de ecuaciones lineales. A continuación presento un ejercicio de las tareas.

Ejercicio: La Combustión del amoníaco (NH_3) en oxígeno (O_2) produce nitrógeno (N_2) y agua (H_2O). Encuentre una ecuación química balanceada de esta reacción. Es decir, encuentre cuantas moléculas de amoníaco y oxígeno son necesarias para producir nitrógeno y agua recordando que los átomos no se pierden al balancear la ecuación.





En la realidad química queremos moléculas enteras y su número debe ser no-negativo. Por eso, el conjunto solución práctico sería el siguiente subconjunto de $\mathcal{L}_s$:

$$\left\{ \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = k \begin{bmatrix} 4 \\ 3 \\ 2 \\ 6 \end{bmatrix} : k \in \mathbb{N} \right\} \subset \mathcal{L}_s$$

1. Sistemas de Ecuaciones Lineales

Ejercicio: Un espía sabe que en cierto aeropuerto secreto hay estacionados 60 aviones, entre cazas y bombarderos. El espía desea determinar cuántos de los 60 aviones son cazas y cuántos son bombarderos. Hay un tipo de cohete que es transportado por ambas clases de aviones; el caza porta seis de estos cohetes y el bombardero sólo dos. El agente sabe que con 240 cohetes quedan pertrechados por completo todos los aviones que se encuentran en el aeropuerto. Además, se enteró de que en esa base el número de aviones caza es el doble del número de bombarderos. Calcule el número de aviones caza y bombarderos que hay en el aeropuerto o bien muestre que la información del agente debe ser incorrecta, ya que es inconsistente.

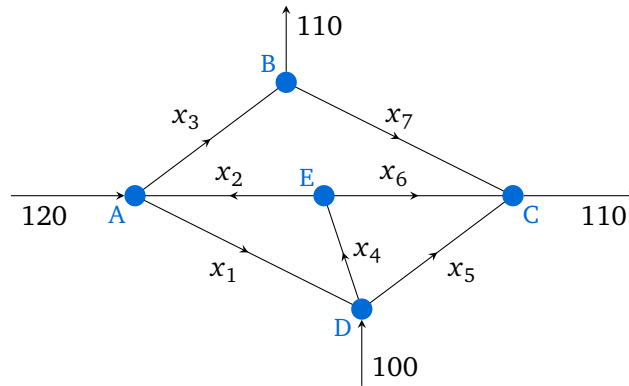
Solución: Lo primero es introducir las variables.

A large grid of graph paper, consisting of 20 columns and 30 rows of small squares, intended for the student to write their solution.

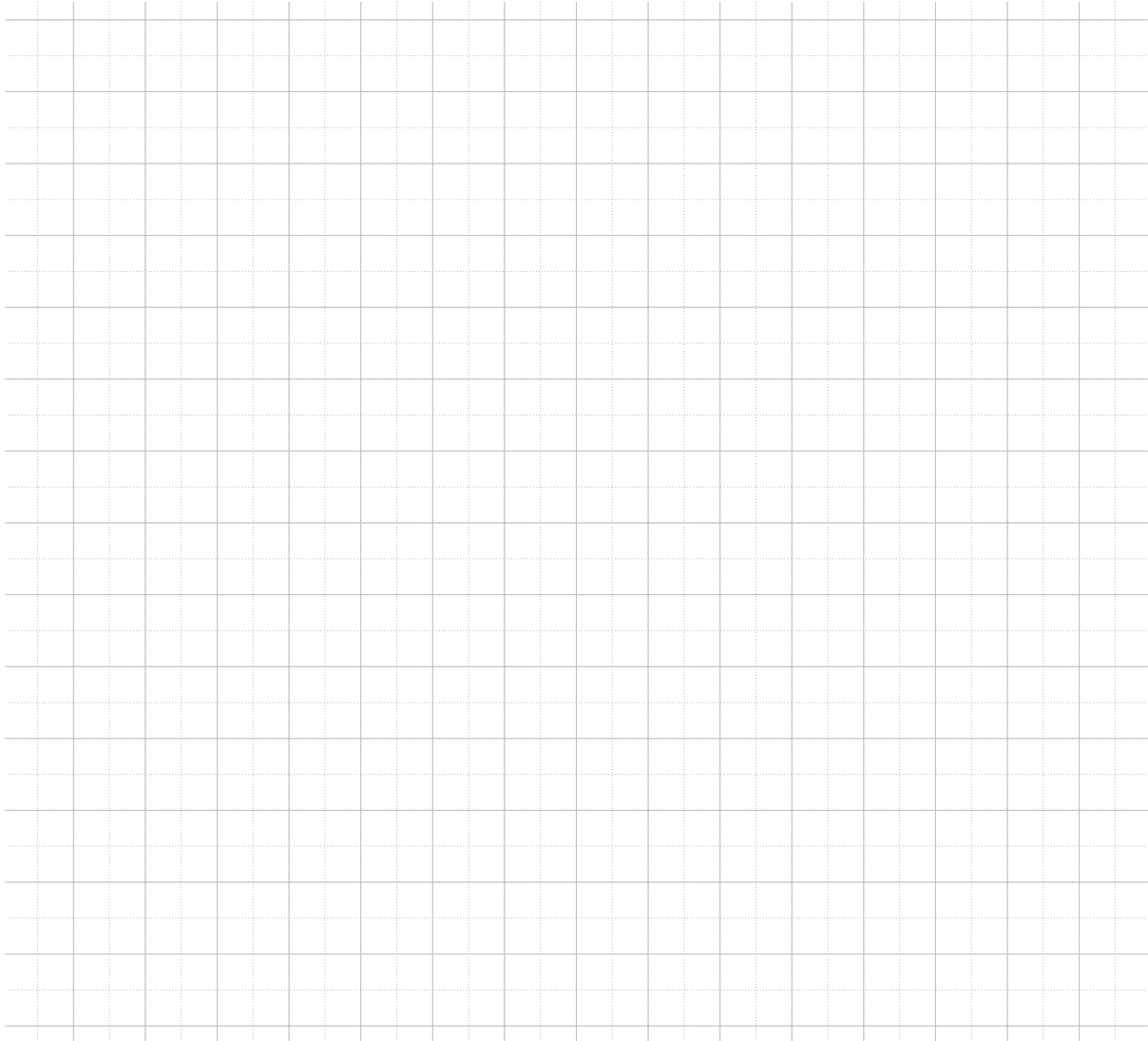
∴ La información del espía es incorrecta, puesto que el sistema es inconsistente.

Ejemplo: Tomado del libro de Álgebra lineal de Rafael Bru, Joan Josep Climent (página 46).

Se desea organizar el tráfico por calles de una ciudad que se representa mediante el diagrama de la siguiente figura. Llegan 120 y 100 coches por hora a los puntos A y D, mientras que de los puntos B y C salen 110 coches por hora. Hallar los posibles flujos de tráfico entre calles.



Solución: Sean $x_1, x_2, \dots, x_7 =$ “# de coches que pasan por las calles $1, \dots, 7$ ”, respectivamente.



Resolvemos el sistema. La primera operación elemental es $R'_4 := R_4 + R_1$:

$$\begin{aligned}
 \left[\begin{array}{ccccccc|c} 1 & -1 & 1 & 0 & 0 & 0 & 0 & 120 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ -1 & 0 & 0 & 1 & 1 & 0 & 0 & 100 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \end{array} \right] & \sim & \left[\begin{array}{ccccccc|c} 1 & -1 & 1 & 0 & 0 & 0 & 0 & 120 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & -1 & 1 & 1 & 1 & 0 & 0 & 220 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \end{array} \right] \\
 & \sim & \left[\begin{array}{ccccccc|c} 1 & 0 & 1 & -1 & 0 & 1 & 0 & 120 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 220 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \end{array} \right] \\
 & \sim & \left[\begin{array}{ccccccc|c} 1 & 0 & 0 & -1 & 0 & 1 & 1 & 10 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \end{array} \right] \\
 & \sim & \left[\begin{array}{ccccccc|c} 1 & 0 & 0 & -1 & 0 & 1 & 1 & 10 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \end{array} \right] \\
 & \sim & \left[\begin{array}{ccccccc|c} 1 & 0 & 0 & -1 & 0 & 1 & 1 & 10 \\ 0 & 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 110 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 110 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]
 \end{aligned}$$

De esta FER podemos leer el conjunto solución:

$$L_S = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} = \begin{pmatrix} 10 + x_4 - x_6 - x_7 \\ x_4 - x_6 \\ 110 + x_7 \\ x_4 \\ 110 - x_6 - x_7 \\ x_6 \\ x_7 \end{pmatrix} : x_i \in \mathbb{N} \right\}$$

El siguiente ejemplo es una aplicación del álgebra lineal en “Cadenas de Markov”.

Ejemplo: Suponga que únicamente dos empresas rivales R, S fabrican cierto producto. Cada año, la empresa R conserva $1/4$ de sus clientes mientras que $3/4$ cambian a S . En el mismo lapso, S conserva $2/3$ de sus clientes mientras $1/3$ cambia a R .

Al comenzar por primera vez la fabricación del producto, R tiene $x_r = 3/5 = 60\%$ del mercado (el mercado es la cantidad total de los clientes), mientras S tiene los otros $x_s = 2/5 = 40\%$.

Escribimos las cantidades iniciales y la distribución de los clientes en un año como vectores:

$$\vec{x}_0 = \begin{bmatrix} x_r \\ x_s \end{bmatrix} = \begin{bmatrix} 3/5 \\ 2/5 \end{bmatrix}$$



Notar que la primera entrada representa la fracción del mercado que “esta con / se queda con / se cambia a” la empresa R y la segunda entrada dice lo mismo sobre los clientes de S .

Entonces podemos definir la matriz $A \in \mathbb{R}^{2 \times 2}$ de transición por

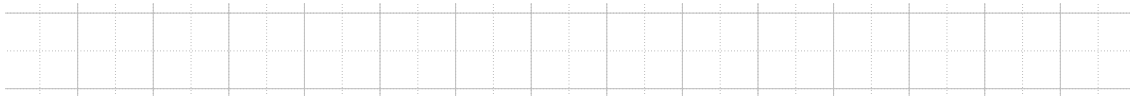


y un año después, la distribución del mercado es:



Para tener un número de clientes (y no un porcentaje) tomamos que el mercado inicial tiene κ -personas, digamos $\kappa = 1200$, y supongamos que este número no se modifica con el paso del tiempo.

En este caso R tiene (después de un año):



y S tiene (después de un año):



1. Sistemas de Ecuaciones Lineales

De manera análoga, al final de dos años, la distribución del mercado estará dado por

$$\vec{x}_2 = A\vec{x}_1 = A(A\vec{x}_0).$$

A las empresas les interesa mantener su número de clientes.

🍃 | Si esto ocurre, es decir, cuando $A\vec{x}_0 = \vec{x}_0$, la distribución $\vec{x}_0$ del mercado se llama **estable**.

Un pregunta natural:

¿Podemos determinar una distribución del mercado que sea la misma año con año?

La respuesta: Sea $\vec{x}_0 = \begin{bmatrix} a \\ b \end{bmatrix}$.

Como R y S controlan todo el mercado, debemos tener:

Además, queremos que la distribución no se modifique después de un año:

Las tres ecuaciones se reducen a dos que podemos resolver:



Ejemplo (un circuito eléctrico):

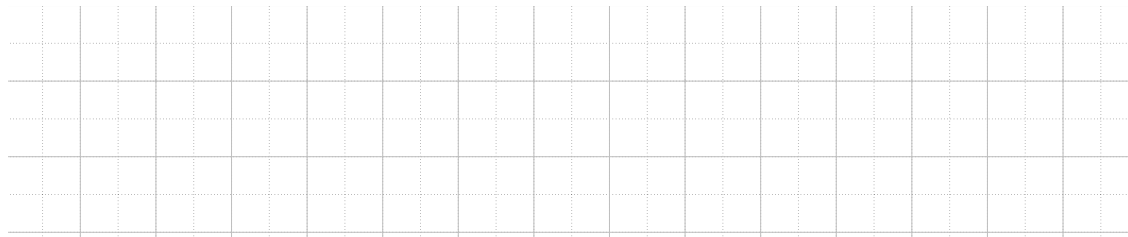
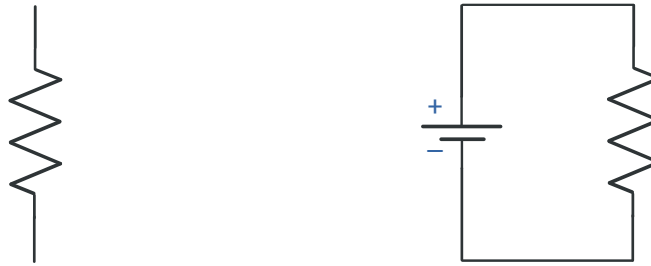
A continuación usamos la siguiente simbología:



Para este ejemplo una *red eléctrica* consiste de estas partes y dentro de ella cada conexión cerrada será un *circuito cerrado*.

La ley de Ohm:

El flujo de una corriente I que pasa por una resistencia (ej. una lámpara) requiere un potencial (un voltaje V) y es proporcional al mismo $V \sim I$. El factor de proporcionalidad depende de la resistencia R y se cumple la *ley de Ohm*: $V = IR$ donde el voltaje V se mide en *volts* $[V]$, el flujo de corriente I en *amperes* $[A]$ y la resistencia R en *ohms* $[\Omega]$.



Lado izquierdo: “flujo natural”

Lado derecho (circuito cerrado): batería proporciona un voltaje igual a la “caída de voltaje RI ”.

Una ley que generaliza la ley de Ohm:**Ley de Kirchhoff del voltaje:**

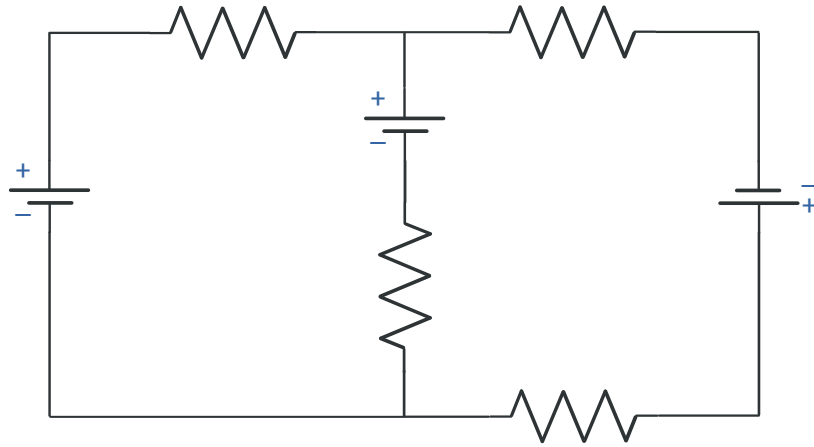
La suma algebraica de las caídas de voltaje RI en una dirección a lo largo de un circuito cerrado es igual a la suma algebraica de las fuentes de voltaje existentes en la misma dirección alrededor del circuito cerrado.

Para aplicar esa ley necesitamos dos convenciones:

- Dentro de cada circuito cerrado C_k habrá un flujo de corriente I_k al cual le asignamos una dirección adivinada. Más tarde, tras haber calculado el flujo I_k , el signo de I_k nos dirá la dirección real del flujo. Precisamente, si $I_k > 0$, la dirección real es la dirección adivinada, en otro caso, la dirección real es opuesta a la dirección adivinada.

1. Sistemas de Ecuaciones Lineales

- Una batería $\begin{array}{c} | \\ | \end{array}$ nos proporciona voltaje. Está considerada positiva (en los cálculos) si la dirección adivinada de la corriente sugiere que el flujo sale del lado positivo (más largo) de la batería; en el caso contrario, el voltaje será negativa.



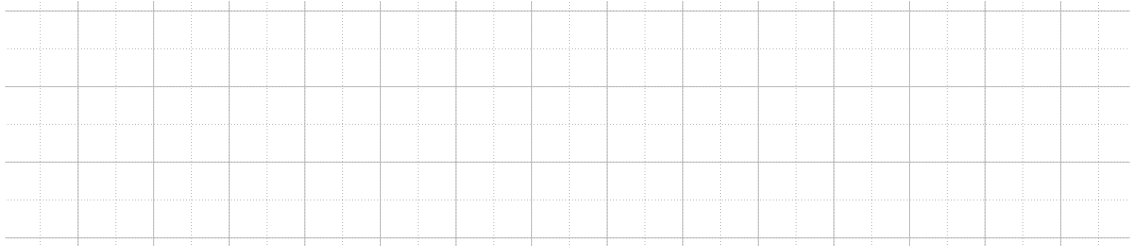
2. Álgebra matricial

Historia: El término de “matriz” fue introducido por primera vez en 1850 por Sylvester, un año después Cayley iniciaba su investigación en ellas.

Hemos definido matrices (en sección 1.2) y visto algunas operaciones entre matrices y vectores. En este capítulo estudiaremos más a fondo las matrices y operaciones entre ellos y otros elementos (vectores, escalares). Las operaciones definidas aquí son la base para entender estadística, optimización y códigos de algoritmos de “inteligencia artificial”.

Notación y formas de matrices:

- Una **matriz cuadrada** es una matriz del tamaño $n \times n$, es decir, ella tiene el mismo número de filas/renglones que de columnas. También, se les denomina matrices $n \times n$. La **diagonal principal** son los elementos $\{a_{11}, a_{22}, \dots, a_{nn}\}$. *Ejemplos:*



La diagonal principal de A es:

La diagonal principal de B es:

La diagonal principal de C es:

- Una **matriz columna** (también **vector columna**) es una matriz de tamaño $n \times 1$.
- Una **matriz fila** / **renglón** (también **vector renglón**) es una matriz de tamaño $1 \times n$.
- Las matrices anteriores son casos especiales de matrices del tamaño $m \times n$:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = [a_{ij}]_{m \times n}$$

Algunas veces no será necesario escribir la matriz completa y bastará con escribir

$A = [a_{ij}]$ ó $A = [a_{ij}]_{m \times n}$ donde $m \times n$ indica el tamaño y los límites de i, j .

2.1. Operaciones con matrices

Primero definimos sumas que son la extensión natural de sumas de vectores. Después, de manera similar, productos con escalares y finalmente productos entre matrices. Habrá que demostrar algunas leyes de distribución y asociatividad. En general **productos NO** serán conmutativos.

2.1.1. Sumas de matrices



Definición. Si dos matrices A y B tienen el mismo tamaño $m \times n$, su **suma** $A + B$ tendrá el mismo tamaño y se efectúa entrada por entrada. Es decir, sean A y B como a continuación se escribe

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ & & \vdots & \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

entonces $A + B$ es:

$A + B =$

Ejemplos:

Hecho I: La suma (de matrices) es conmutativa.

Demostración. Nuestro objetivo es demostrar que $A + B = B + A$.

Sean A y B matrices del mismo tamaño, es decir

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}, \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ & & \vdots & \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

Calculamos:

$$A + B =$$

□

Hecho II: La suma (de matrices) es asociativa.

Demostración. Nuestro objetivo es demostrar que $(A + B) + C = A + (B + C)$.

Usamos notación corta. Sean A, B y C matrices del mismo tamaño, es decir

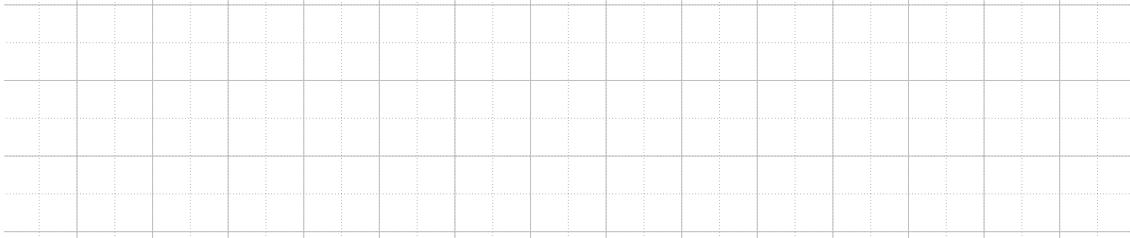
$$A = [a_{ij}]_{m \times n}, \quad B = [b_{ij}]_{m \times n}, \quad C = [c_{ij}]_{m \times n}$$

□

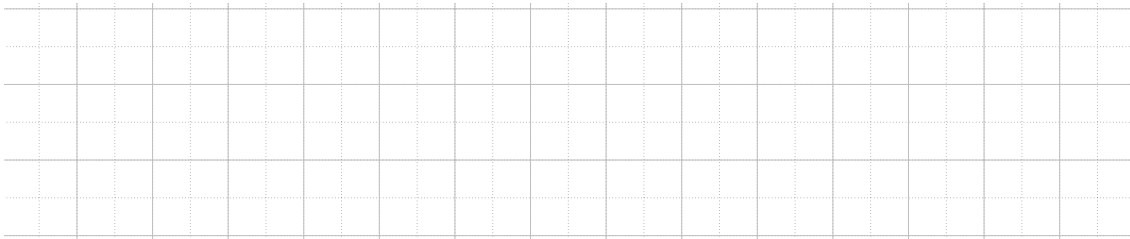
2.1.3. Productos de matrices

Ya hemos conocido los siguientes casos especiales:

- El producto punto que multiplica una matriz renglón $1 \times n$ con una matriz columna $n \times 1$ y da como resultado un número real (un escalar).



- El producto de matriz por vector columna (matriz columna) en sistemas lineales da como resultado un vector con tantas renglones como la matriz.



 **Observación:** Al cambiar una matriz renglón (en un producto punto) por una matriz con m renglones, el resultado (anteriormente un escalar “matriz 1×1 ”) gana renglones ya que el nuevo resultado será una matriz columna de la forma $m \times 1$.

Aún no definido, podemos pensar en tomar un producto punto y cambiar al vector columna por una matriz que tiene varias columnas. ¿Te parece lógico que el resultado debe ganar columnas?

 **Regla de oro (anterior):** En ambos casos era necesario que el número de columnas de la primera matriz (izquierda) era igual al número de renglones (entradas) del vector columna (matriz derecha).

La siguiente regla de oro y definición respetan los casos especiales anteriores:

 **Regla de oro:** Sean A, B dos matrices. Entonces, el producto $A * B$ está definido **si y solamente si** el número de columnas de la matriz **izquierda** A es igual al número de renglones de la matriz **derecha** B . Otra forma de decir eso: La **dimensión interior** (de los lados de matrices que tocan la operación $*$) debe ser igual.

 **Definición 2.1.** Si $A \in \mathbb{R}^{m \times p}$ y $B \in \mathbb{R}^{p \times n}$ entonces el **producto** $A * B$ es una matriz con tantas renglones como tiene la **primera** matriz (m renglones) y tantas columnas como tiene la **segunda** matriz (n columnas), donde la columna j del producto es el producto de A con j -ésima columna de B , es decir

$$A * B = \begin{bmatrix} AB_{:1} & AB_{:2} & \dots & AB_{:n} \end{bmatrix}.$$

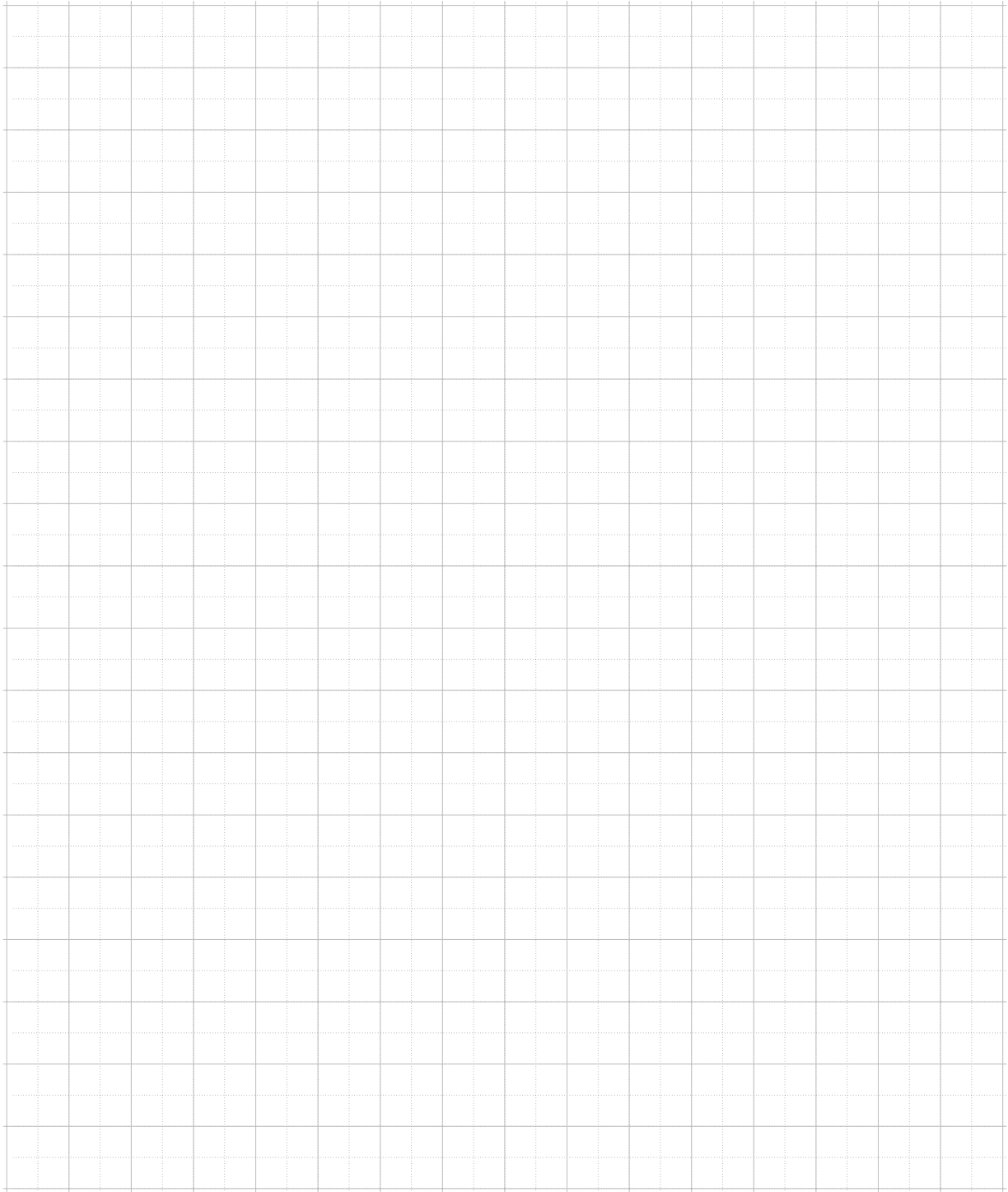
Otra forma de decirlo: El producto $A * B$ tiene las “dimensiones exteriores” (de los lados de las matrices que no tocan la operación $*$).

2. Álgebra matricial

Ejemplo: (no conmutativo $2 \times 3 * 3 \times 2$, por dimensión del resultado)

Para A, B dadas abajo calcule $A * B$ y $B * A$:

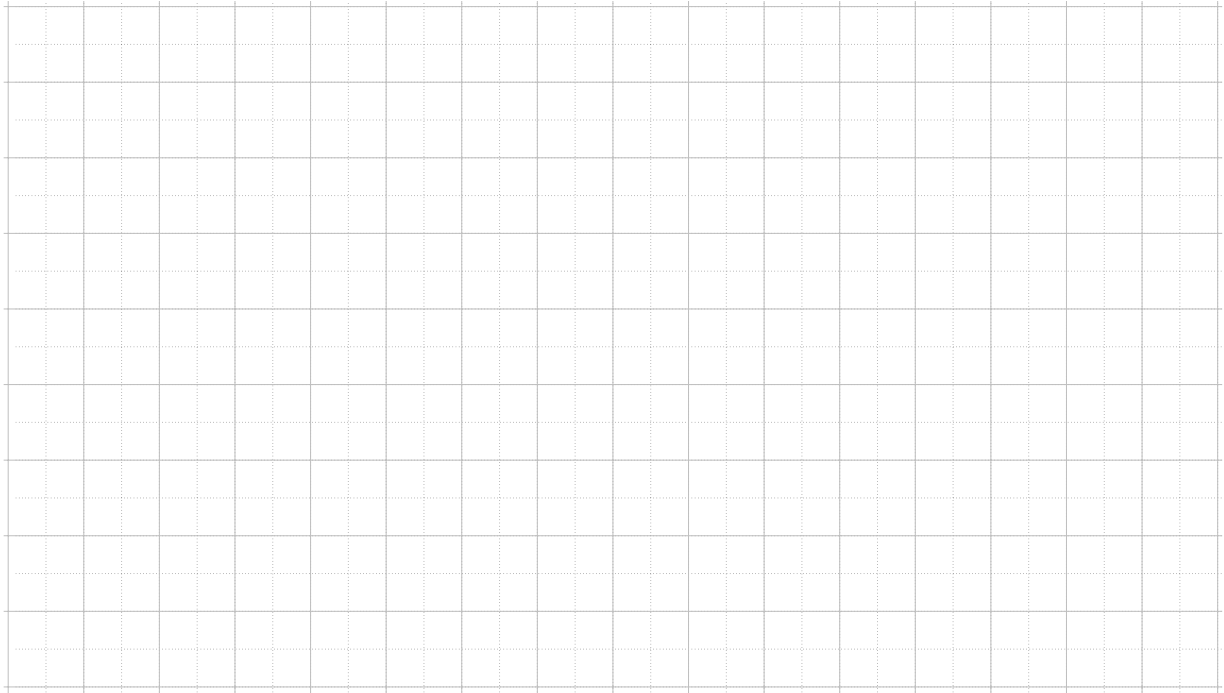
$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 3 & 4 \end{bmatrix}$$



Ejemplo: (no conmutativo $2 \times 3 * 3 \times 1$, por no ser definido)

Para A, B dadas abajo calcule $A * B$ y justifique que $B * A$ no está definido:

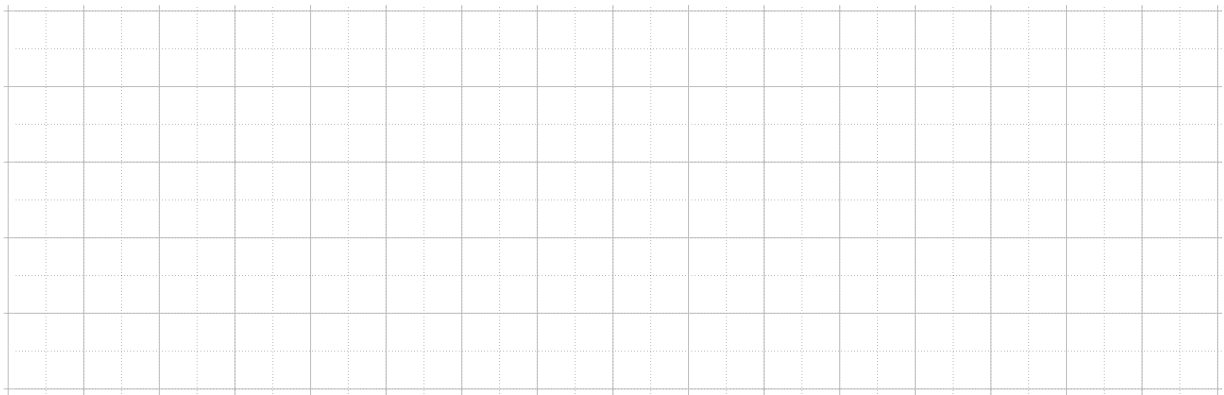
$$A = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$$



Ejemplo: (no conmutativo $2 \times 2 * 2 \times 2$, por resultado distinto)

Calcular los productos $A * B$ & $B * A$ con

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}.$$



 | En matrices puede suceder que $AB = \mathcal{O}$ siendo $A \neq \mathcal{O}$ & $B \neq \mathcal{O}$.

2. Álgebra matricial

Ejemplo: (Productos “exteriores” $3 \times 1 * 1 \times 3$)

Calcular los productos $A * B$ & $B * A$ y $C * D$ & $D * C$ con

$$A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \& \quad B = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \quad \& \quad C = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \quad \& \quad D = \begin{bmatrix} 0 & 3 & 1 \end{bmatrix} .$$



La transpuesta de una matriz.



Definición. Si $A = [a_{ij}]_{m \times n}$ es una matriz (de tamaño $m \times n$), entonces **la matriz transpuesta de A** , se denota por A^T y es de tamaño $n \times m$ donde la entrada en el j -ésimo renglón y en la i -ésima columna es

$$(A^T)_{ji} = a_{ij} .$$

En palabras: la matriz A^T (transpuesta de A) es la matriz cuyas renglones son las columnas de A .



Los productos de matrices (tal vez) más importantes en estadística (matriz de covarianza) y en algunas demostraciones matemáticas son del tipo $AA^T \in \mathbb{R}^{m \times m}$ y $A^T A \in \mathbb{R}^{n \times n}$.

Ejemplo: Escribir las transpuestas de las matrices A, B y calcular los productos AA^T y B^TB .



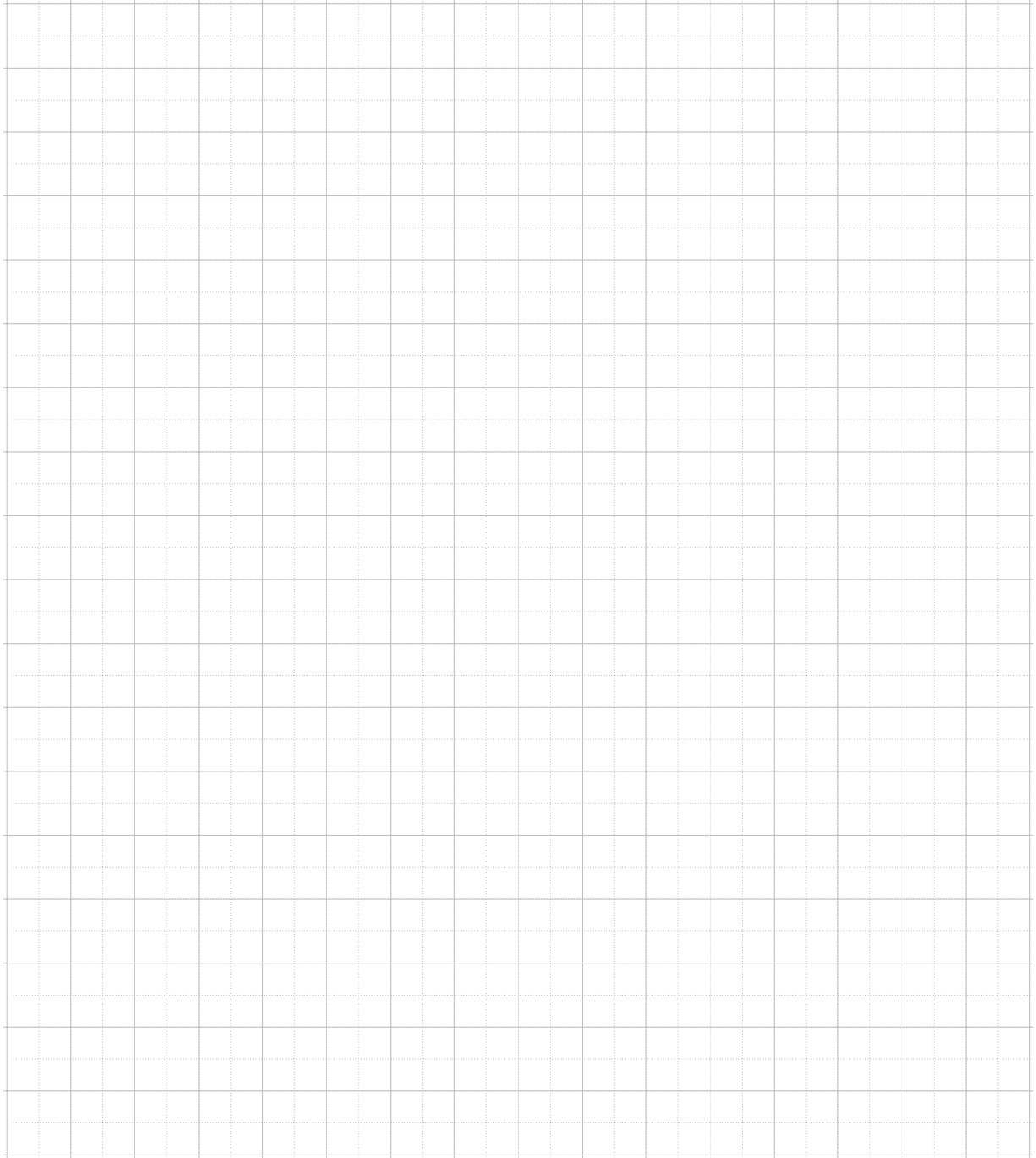
Definición. El vector $\vec{e}_j \in \mathbb{R}^{n \times 1}$ denota al j -ésimo **vector canónico**, es decir, la j -ésima entrada del vector $\vec{e}_j$ es 1 y las otras son 0. Para visualizar escribo su transpuesta:

$$\vec{e}_j^T = [0 \quad \dots \quad 0 \quad \overset{j}{1} \quad 0 \quad \dots \quad 0]$$

2. Álgebra matricial

Ejemplo: Vectores canónicos en $\mathbb{R}^3$ y la matriz identidad como suma de productos “exteriores” de ellos.

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$



En general tenemos la igualdad:

$$I_n = \sum_{j=1}^n \vec{e}_j \vec{e}_j^{\top}$$



Hecho I: La matriz identidad es el elemento neutral del producto.

Vimos que la multiplicación de matrices no es conmutativa.

Por eso, la identidad del lado izquierdo puede ser (por su dimensión) distinta del lado derecho.

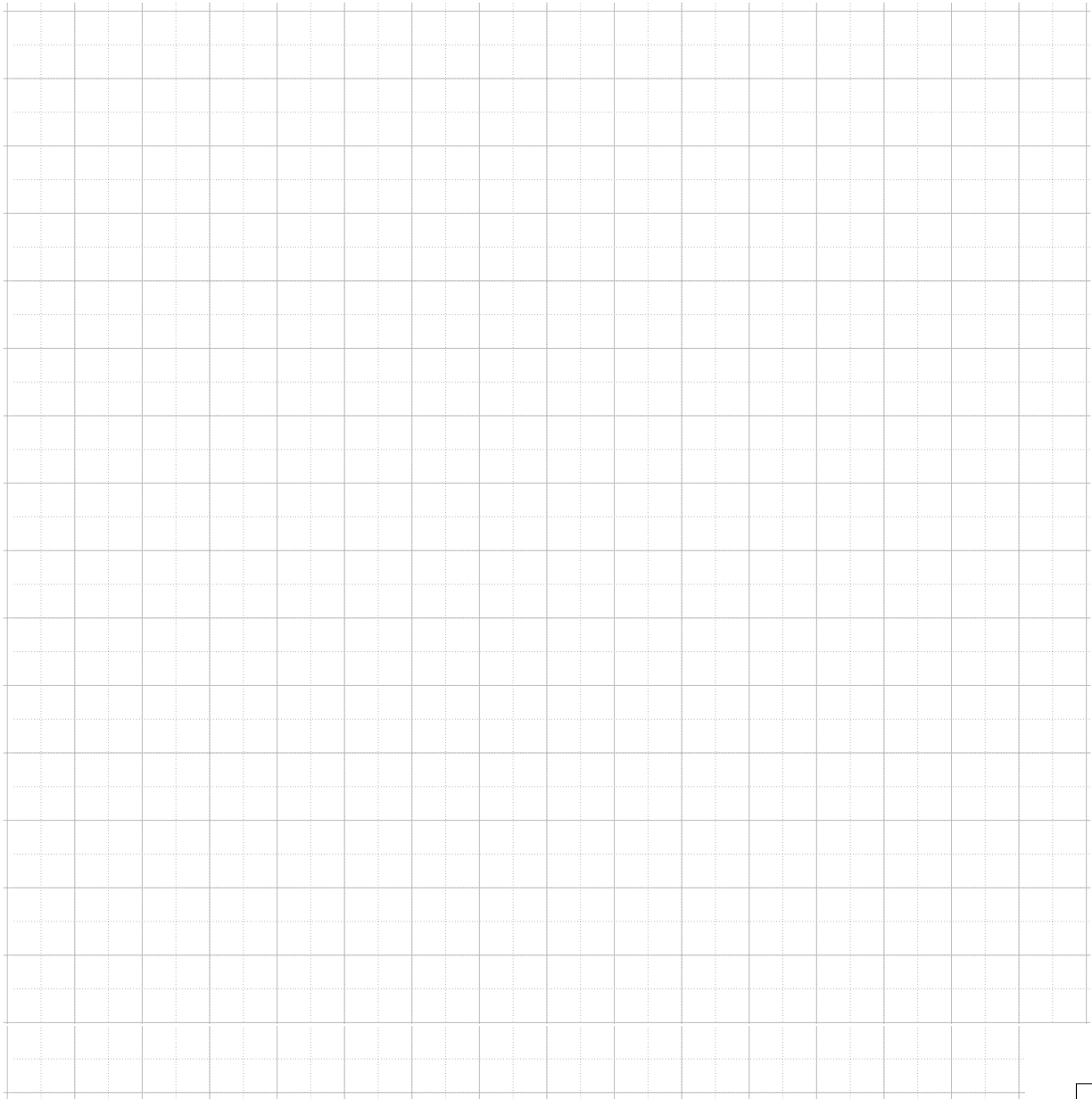


Lema 2.1. Sea $A \in \mathbb{R}^{m \times n}$ una matriz (de tamaño $m \times n$). Entonces, la matriz identidad (por la izquierda) es $I_m \in \mathbb{R}^{m \times m}$ y la matriz identidad (por la derecha) es $I_n \in \mathbb{R}^{n \times n}$ puesto que se cumplen las siguientes identidades:

$$I_m A = A = A I_n .$$

Demostración. Usar la definición 2.1 (del producto de matrices).

Haremos uso de la propiedad $I\vec{x} = \vec{x}$.



□

Hecho II: Leyes distributivas

Teorema 2.2. Sea A una matriz de tamaño $m \times n$, y sean B, C matrices con tamaños adecuados para que las sumas y productos (que aparecen abajo) están definidos. Entonces, se cumplen las siguientes propiedades:

i) $A(B + C) = AB + AC$

(Ley distributiva por la derecha)

ii) $(B + C)A = BA + CA$

(Ley distributiva por la izquierda)

Motivación: El lado izquierdo de esas igualdades requiere un producto matricial menos.

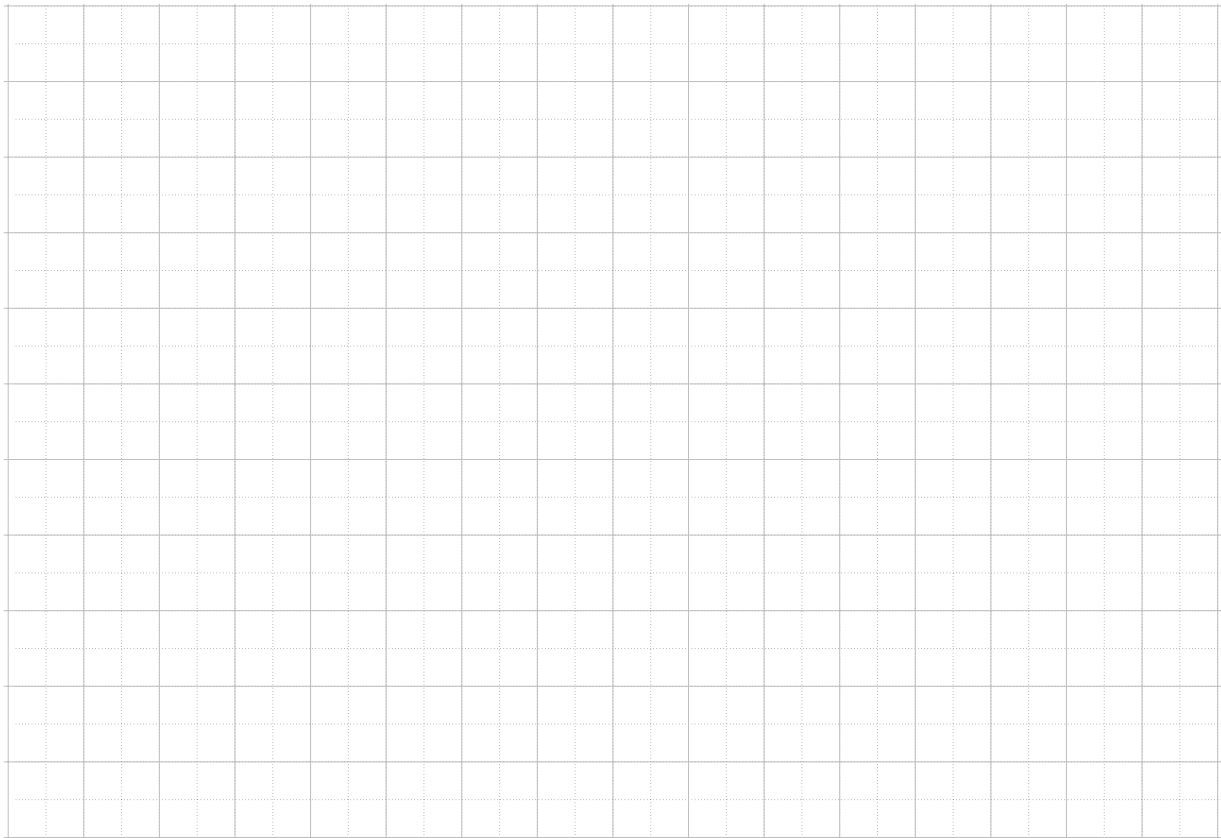
Para convencerte, calcula los dos lados para las siguientes matrices:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 2 & 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 3 \\ 0 & -2 & 1 \\ 2 & 4 & -3 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & -2 \\ -2 & 4 & 1 \end{bmatrix}$$

y cuenta los productos puntos que hiciste de ambos lados:

$$A(B + C) =$$

$$AB + AC =$$

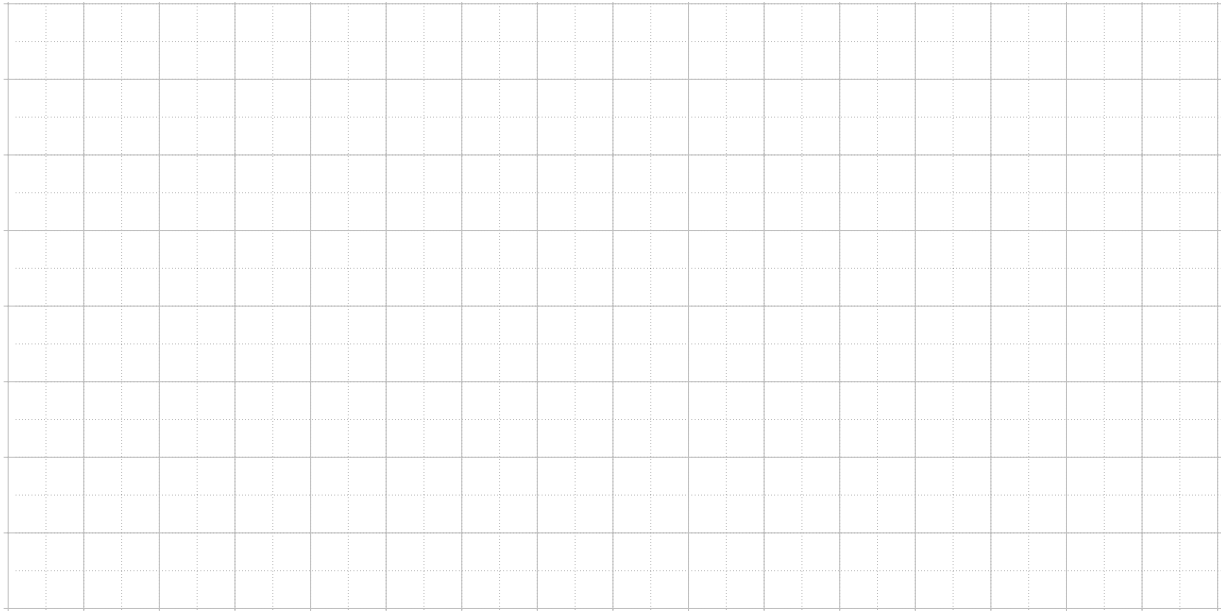


Ejercicio: para i) demuestre el caso especial: $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$.

Ayuda: véase la prueba del lema 1.2.

Ejemplo: para ii) demuestre el caso especial: $(B + C)\vec{u} = B\vec{u} + C\vec{u}$.

Usamos la definición de combinaciones lineales y propiedades de sumas de vectores:

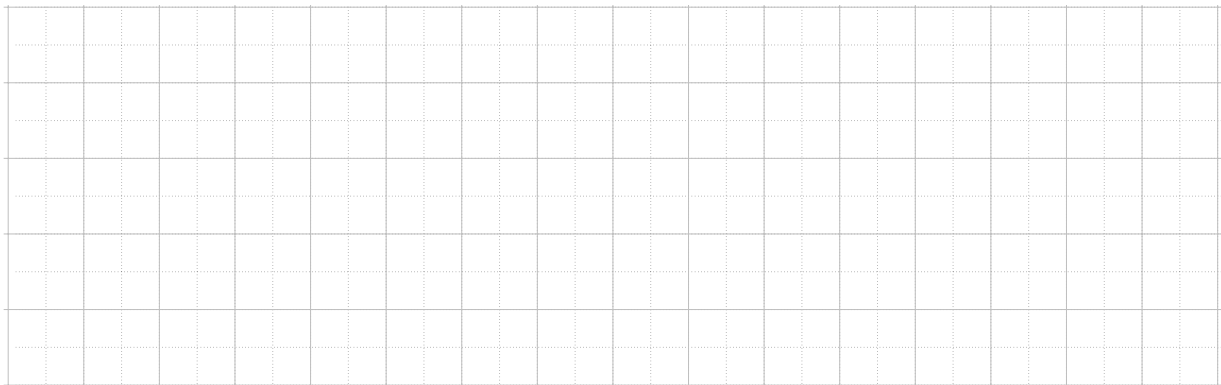


Demostración del teorema 2.2. En general, para tener la suma $B + C$ definida B & C deben tener el mismo número de columnas y renglones.

Primero, mostramos i). La “regla de oro” del producto $A(B + C)$ pide que B, C tienen n renglones, ya que A tiene n columnas. En conclusión, B & C son de tamaño $n \times p$ donde p es su número de columnas. Para deducir la igualdad re-usamos la propiedad $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$ y la definición 2.1:

$$\begin{aligned}
 AB + AC &= \begin{bmatrix} AB_{[:1]} & AB_{[:2]} & \dots & AB_{[:p]} \end{bmatrix} + \begin{bmatrix} AC_{[:1]} & AC_{[:2]} & \dots & AC_{[:p]} \end{bmatrix} \\
 &= \begin{bmatrix} AB_{[:1]} + AC_{[:1]} & AB_{[:2]} + AC_{[:2]} & \dots & AB_{[:p]} + AC_{[:p]} \end{bmatrix} \\
 &= \begin{bmatrix} A(B_{[:1]} + C_{[:1]}) & A(B_{[:2]} + C_{[:2]}) & \dots & A(B_{[:p]} + C_{[:p]}) \end{bmatrix} = A(B + C).
 \end{aligned}$$

Ahora, para mostrar ii), debes completar los argumentos usando la “regla de oro” para el producto $(B + C)A$, la igualdad $(B + C)\vec{u} = B\vec{u} + C\vec{u}$ y la definición 2.1:



Eso concluye la demostración. □

Hecho III: Leyes asociativas



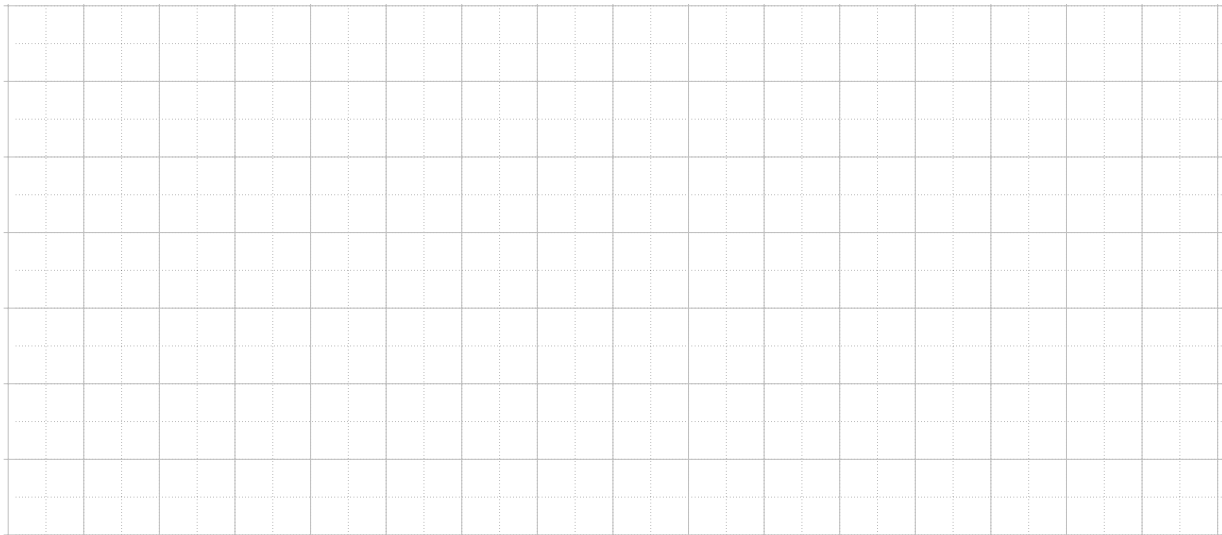
Teorema 2.3. Sea γ un escalar y sean A, B, C matrices de tamaños adecuados para definir los siguientes productos. Entonces, se cumplen las propiedades:

i) $\gamma(AB) = (\gamma A)B = A(\gamma B)$ *(multiplicación por escalar flexible)*

ii) $(AB)C = A(BC)$ *(Ley asociativa del producto)*

Motivación: Usar uno de los dos lados de ii) requiere menos productos puntos y causa matrices temporales más pequeñas, por ejemplo calcula ambos lados de ii) para A, B, C dados por:

$$A = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$



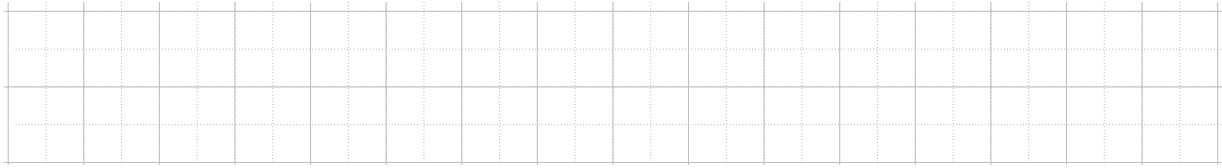
Demostración de i). Se deja como ejercicio.

Ayuda: Empiece con el caso especial $\alpha(A\vec{u}) = (\alpha A)\vec{u} = A(\alpha\vec{u})$ con las técnicas de páginas 35–36. $\square$

Demostración de ii). Paso 1. P.d.: Si $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$ y $\vec{x} \in \mathbb{R}^{p \times 1}$, entonces

$$(AB)\vec{x} = A(B\vec{x}).$$



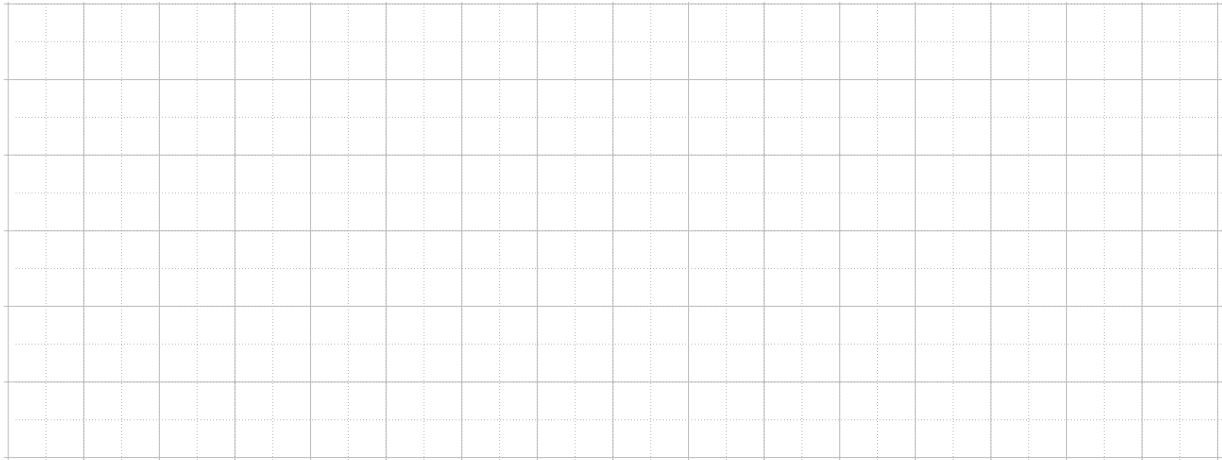


Por la regla de oro para productos de matrices, si $A \in \mathbb{R}^{m \times n}$, entonces $B \in \mathbb{R}^{n \times p}$ y $C \in \mathbb{R}^{p \times q}$. Con esas dimensiones los productos de i) son definidos. Usamos el paso 1 y la definición 2.1 para concluir: $BC = \begin{bmatrix} BC_{[:,1]} & BC_{[:,2]} & \dots & BC_{[:,q]} \end{bmatrix}$ y

$$\begin{aligned} A(BC) &= A \begin{bmatrix} BC_{[:,1]} & BC_{[:,2]} & \dots & BC_{[:,q]} \end{bmatrix} \\ &= \begin{bmatrix} A(BC_{[:,1]}) & A(BC_{[:,2]}) & \dots & A(BC_{[:,q]}) \end{bmatrix} \\ &= \begin{bmatrix} (AB)C_{[:,1]} & (AB)C_{[:,2]} & \dots & (AB)C_{[:,q]} \end{bmatrix} \\ &= (AB) \begin{bmatrix} C_{[:,1]} & C_{[:,2]} & \dots & C_{[:,q]} \end{bmatrix} = (AB)C. \end{aligned}$$

Con eso hemos demostrado ii). □

Ejemplo: Simplifica la expresión: $A(B - C)D + ABD - A(2B + C)D + 2ACD$.



Ejercicio: Usando lema 2.1, teoremas 2.2 y 2.3 demuestre que $\gamma(B + C) = \gamma B + \gamma C$.

Ejercicio: Simplifica la expresión: $A(2B + C) + (2A - B)C + 2A(B - 2C)$.



Definición. Dados una matriz cuadrada A y un número natural positivo k definimos

$$A^k = \underbrace{A * A * A * \dots * A}_{k \text{ veces}}.$$



Definición. Se dice que dos matrices A & B **conmutan** cuando $AB = BA$.

Debido al producto conmutativo $ba = ab$ para $a, b \in \mathbb{R}$, se tiene $(a + b)(a - b) = a^2 - b^2$.

Ejemplo: Demuestra que A & B conmutan si y solo si $(A + B)(A - B) = A^2 - B^2$.

Demostración.



□

2.2. Propiedades de matrices transpuestas

Recordamos las siguientes definiciones:

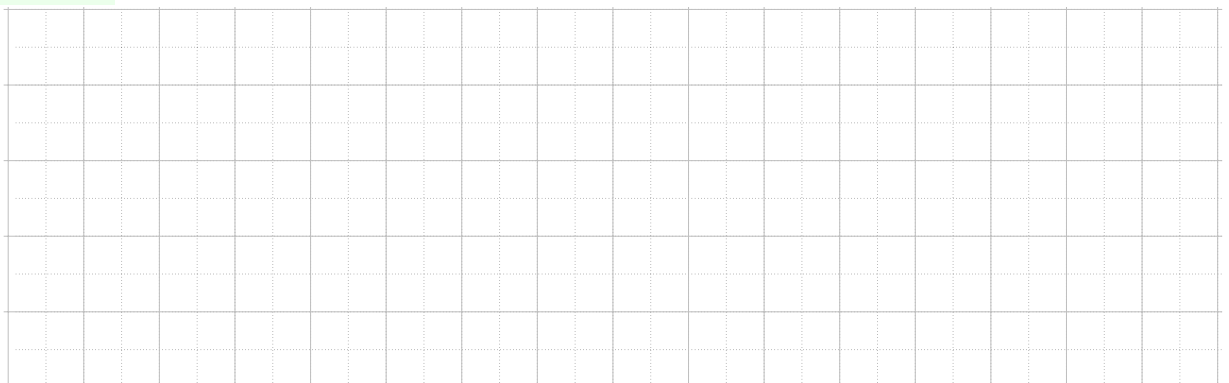
En página 54 vimos la definición de la transpuesta $A^T \in \mathbb{R}^{n \times m}$ de una matriz $A \in \mathbb{R}^{m \times n}$.

En página 55 vimos la definición de vectores canónicos $\vec{e}_j$.

Para mostrar las propiedades serán útiles los siguientes productos entre vectores canónicos y una matriz. Si A es una matriz de tamaño $m \times n$ y el vector canónico es

- a) $\vec{e}_j \in \mathbb{R}^{n \times 1}$, entonces $A\vec{e}_j \in \mathbb{R}^{m \times 1}$ es la j -ésima columna de A .
- b) $\vec{e}_i^T = [0, \dots, 0, 1, 0, \dots, 0] \in \mathbb{R}^{1 \times m}$, entonces $\vec{e}_i^T A \in \mathbb{R}^{1 \times n}$ es el i -ésimo renglón de A .

Ejemplos:



Propiedades de matrices transpuestas frecuentemente usadas:



Teorema 2.4. Sean A y B matrices cuyos tamaños son tales que las operaciones indicadas están definidas y sea $\gamma \in \mathbb{R}$ un escalar. Entonces

- a) $(A^\top)^\top = A$
- b) $(A + B)^\top = A^\top + B^\top$
- c) $(\gamma A)^\top = \gamma A^\top$
- d) $(AB)^\top = B^\top A^\top$
- e) $(A^k)^\top = (A^\top)^k$ para naturales $k \geq 2$.

Demostración.

Las propiedades **a)**, **b)** y **c)** son buenos ejercicios para practicar.

Ahora **d)**. Primero demostramos el caso especial $A \in \mathbb{R}^{m \times n}$ y $\vec{u} \in \mathbb{R}^{n \times 1}$:

$$(A\vec{u})^\top = \vec{u}^\top A^\top \quad (\star)$$

Usamos la definición de combinaciones lineales, **b)** y **c)** y teoremas 2.3 y 2.2 para llegar a:

$$\begin{aligned} (A\vec{u})^\top &= (u_1 A_{[1]} + u_2 A_{[2]} + \dots + u_n A_{[n]})^\top \\ &= u_1 (A_{[1]})^\top + u_2 (A_{[2]})^\top + \dots + u_n (A_{[n]})^\top \\ &= u_1 (\vec{e}_1^\top A^\top) + u_2 (\vec{e}_2^\top A^\top) + \dots + u_n (\vec{e}_n^\top A^\top) \\ &= (u_1 \vec{e}_1^\top) A^\top + (u_2 \vec{e}_2^\top) A^\top + \dots + (u_n \vec{e}_n^\top) A^\top \\ &= (u_1 \vec{e}_1^\top + u_2 \vec{e}_2^\top + \dots + u_n \vec{e}_n^\top) A^\top = \begin{bmatrix} u_1 & u_2 & \dots & u_n \end{bmatrix} A^\top = \vec{u}^\top A^\top. \end{aligned}$$

Para **d)** en general fijamos $A \in \mathbb{R}^{m \times n}$ & $B \in \mathbb{R}^{n \times p}$, ya que así los productos están definidos. Luego, sabemos que el i -ésimo renglón del producto AB es $\vec{e}_i^\top(AB) = (\vec{e}_i^\top A)B$ por teorema 2.3, es decir, $(AB)_{[i]} = A_{[i]}B$ y el producto AB puede ser escrito de las dos formas siguientes

$$\begin{bmatrix} AB_{[1]} & AB_{[2]} & \dots & AB_{[p]} \end{bmatrix} = AB = \begin{bmatrix} A_{[1]}B \\ A_{[2]}B \\ \vdots \\ A_{[m]}B \end{bmatrix}.$$

Esas igualdades y la definición de la transpuesta dan:

$$(AB)^\top = \begin{bmatrix} AB_{[1]} & AB_{[2]} & \dots & AB_{[p]} \end{bmatrix}^\top = \begin{bmatrix} (AB_{[1]})^\top \\ (AB_{[2]})^\top \\ \vdots \\ (AB_{[p]})^\top \end{bmatrix} \stackrel{(\star)}{=} \begin{bmatrix} B_{[1]}^\top A^\top \\ B_{[2]}^\top A^\top \\ \vdots \\ B_{[p]}^\top A^\top \end{bmatrix} = B^\top A^\top.$$

2. Álgebra matricial

Para demostrar $e)$ usamos algo que se conoce como inducción sobre k empezando con $k = 2$.

Paso Base:

Sea $k = 2$, entonces $(A^2)^\top = (A * A)^\top \stackrel{d)}{=} A^\top * A^\top = (A^\top)^2$, lo cual confirma la igualdad.

Notar que el producto $A * A$ requiere (por la regla de oro) que A es cuadrada.

Hipótesis inductiva:

Supongamos que el resultado se cumple para un $k \geq 2$, es decir, $(A^k)^\top = (A^\top)^k$.

Paso inducción:

Tomando la hipótesis de inducción aumentamos k a $k + 1$:

Empecemos con el lado izquierdo:

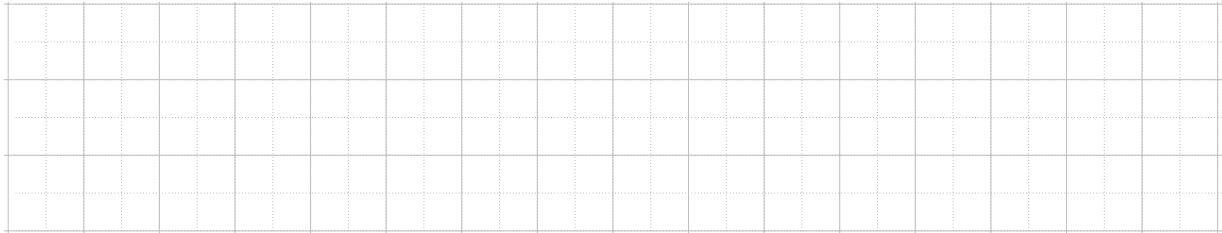
$$(A^{k+1})^\top = (A^k * A)^\top \stackrel{d)}{=} A^\top (A^k)^\top = A^\top (A^\top)^k = (A^\top)^{k+1}.$$

Con esto se termina la demostración. □

2.3. Matrices especiales



Definición. Una matriz cuadrada es **diagonal** si entradas distinto de cero se encuentran únicamente en su diagonal principal.

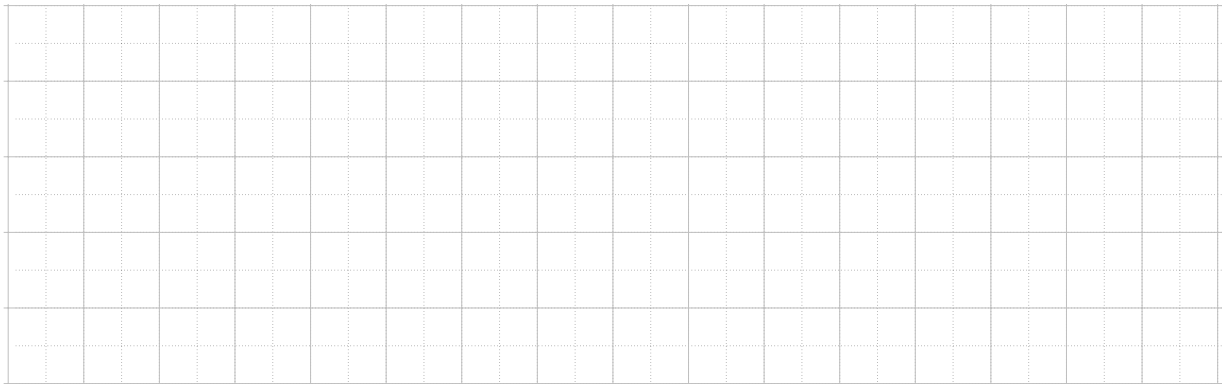


Definición. Una matriz A es **triangular inferior** si $(i < j \implies a_{ij} = 0)$.

En palabras: Entradas arriba de la diagonal principal son cero.

Una matriz A es **triangular superior** si $(i > j \implies a_{ij} = 0)$.

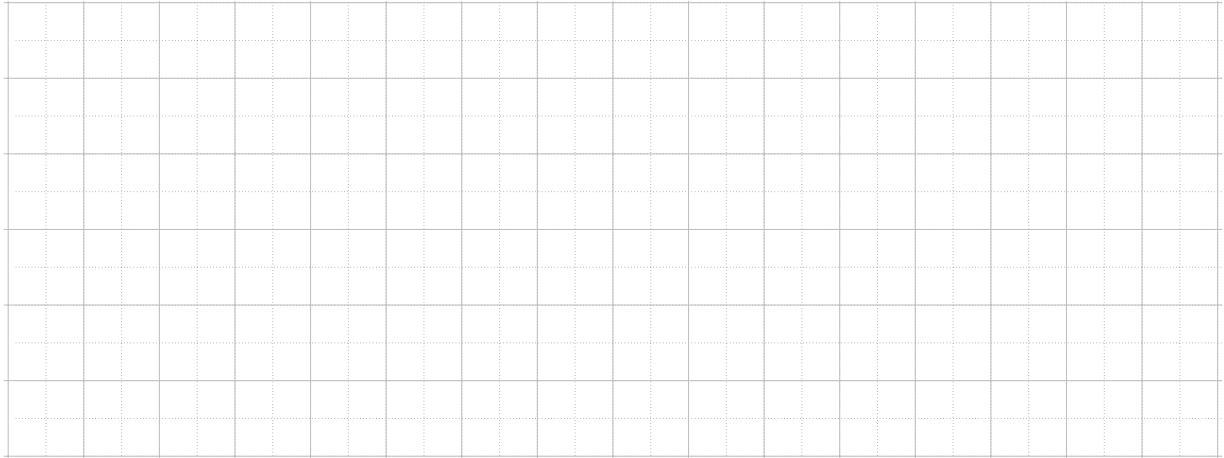
En palabras: Entradas abajo de la diagonal principal son cero.



Notar que A es triangular inferior si y solo si $A^\top$ es triangular superior.

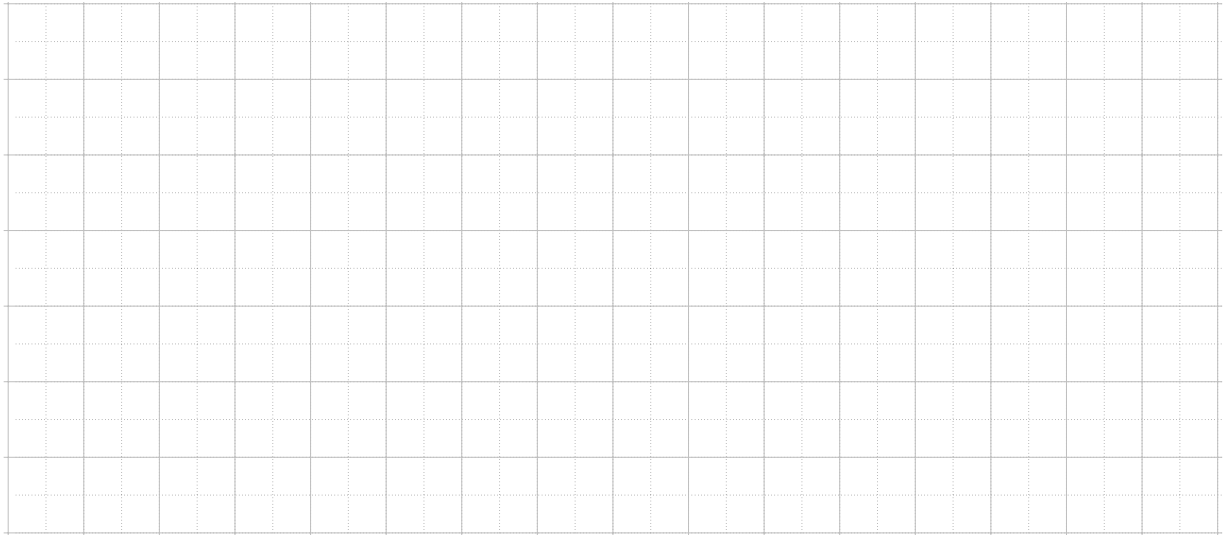
 | **Definición.** Se dice que una matriz A es **simétrica** si $A = A^T$ (alternativa: $A + A^T = 2A$).

Ejemplos:



 | **Definición.** Se dice que una matriz A es **antisimétrica** si $A = -A^T$ (alternativa: $A + A^T = \mathcal{O}$).

Ejemplos:



 | Cualquier matriz cuadrada se puede escribir como su parte simétrica y antisimétrica:

$$A = \frac{1}{2}[A + A^T] + \frac{1}{2}[A - A^T]. \quad (2.1)$$

Ejercicio: Demuestre que $A + A^T$ es una matriz simétrica y que $A - A^T$ es antisimétrica.

2.4. Motivación del siguiente tema

Aún no lo hemos definido, pero hemos visto ejemplos de un concepto común.

Ejemplo 1: El lema 1.2 demuestra que si $A\vec{u} = \vec{0}$ y $A\vec{v} = \vec{0}$ entonces $A(\alpha\vec{u} + \beta\vec{v}) = \vec{0}$ para $\alpha, \beta \in \mathbb{R}$.

Ejemplo 2: Si A, B son matrices simétricas, entonces γA es simétrica y $A + B$ es simétrica.



Ejemplo 3: Si A, B son matrices triangulares inferiores, entonces γA es triangular inferior y $A + B$ es triangular inferior.



El concepto se conoce como:

cerradura bajo sumas y multiplicaciones por escalar

y es fundamental para definir un espacio vectorial.

En “geometría vectorial” el concepto de vector se asocia con ubicación y desplazamiento.

El próximo tema cambiará nuestra idea de vectores al caso más general: Una vez definido, un *espacio vectorial* será un conjunto de objetos junto con una suma entre ellos y una multiplicación de un objeto por un escalar. Los objetos serán denominados *vectores*, pero podrían ser vectores, matrices, polinomios o funciones continuas.

3. Espacios vectoriales

Algunas propiedades matemáticas que vimos son válidas para objetos de distintas categorías. Por eso, la definición de un espacio vectorial es un poquito abstracta.



Definición 3.1. Sean $\mathcal{V} \neq \emptyset$ un conjunto de objetos,

$+: \mathcal{V} \times \mathcal{V} \rightarrow \mathcal{V}$ una suma de dos objetos de $\mathcal{V}$ que da un objeto en $\mathcal{V}$ y sea

$\cdot: \mathbb{R} \times \mathcal{V} \rightarrow \mathcal{V}$ un producto de escalar por un objeto de $\mathcal{V}$ que da un objeto en $\mathcal{V}$.

Entonces, $(\mathcal{V}, +, \cdot)$ se denomina **espacio vectorial** y sus objetos serán llamados **vectores**, cuando el conjunto y las operaciones cumplen lo siguiente:

1. Para cualesquiera dos elementos $\vec{a}, \vec{b} \in \mathcal{V}$, se tiene $\vec{a} + \vec{b} \in \mathcal{V}$ y

1.a) Existe un elemento especial $\vec{0} \in \mathcal{V}$ que se llama **elemento neutro** (a veces *nulo*) y cumple $\vec{a} + \vec{0} = \vec{a}$ para todo $\vec{a} \in \mathcal{V}$.

1.b) Para cada elemento $\vec{a} \in \mathcal{V}$ existe un elemento negativo $-\vec{a} \in \mathcal{V}$ el **aditivo inverso** tal que $\vec{a} + (-\vec{a}) = \vec{0}$.

1.c) La suma es **asociativa**, es decir para cada $\vec{a}, \vec{b}, \vec{c} \in \mathcal{V}$ se cumple

$$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}.$$

1.d) La suma es **conmutativa**, es decir para cada $\vec{a}, \vec{b} \in \mathcal{V}$ se cumple $\vec{a} + \vec{b} = \vec{b} + \vec{a}$.

2. Para cada $\vec{a} \in \mathcal{V}$ y un número real r se tiene que $r\vec{a} \in \mathcal{V}$, siempre $r\vec{a} = r \cdot \vec{a}$. Además

2.a) Para todo $\vec{a} \in \mathcal{V}$ se cumple que $1 \cdot \vec{a} = \vec{a}$, es decir 1 es el neutro de multiplicación por escalar.

2.b) Para todo $r \in \mathbb{R}$ y elementos $\vec{a}, \vec{b} \in \mathcal{V}$ se cumple la propiedad **distributiva**:

$$r(\vec{a} + \vec{b}) = r\vec{a} + r\vec{b}$$

2.c) Para todo $r, s \in \mathbb{R}$ y elemento $\vec{a} \in \mathcal{V}$ se cumple la propiedad **distributiva**:

$$(r + s)\vec{a} = r\vec{a} + s\vec{a}$$

2.d) Para todo $r, s \in \mathbb{R}$ y cada elemento $\vec{a} \in \mathcal{V}$ se cumple la propiedad **asociativa**. Esto es

$$r(s\vec{a}) = (rs)\vec{a}.$$

3.1. Ejemplos

Existen varios ejemplos de espacios vectoriales. A continuación veremos los 3 más comunes. Para ver si un conjunto y las operaciones $+$, $\cdot$ forman un espacio vectorial debemos verificar las 10 condiciones dadas en la definición 3.1.

3.1.1. El espacio vectorial $\mathbb{R}^n$

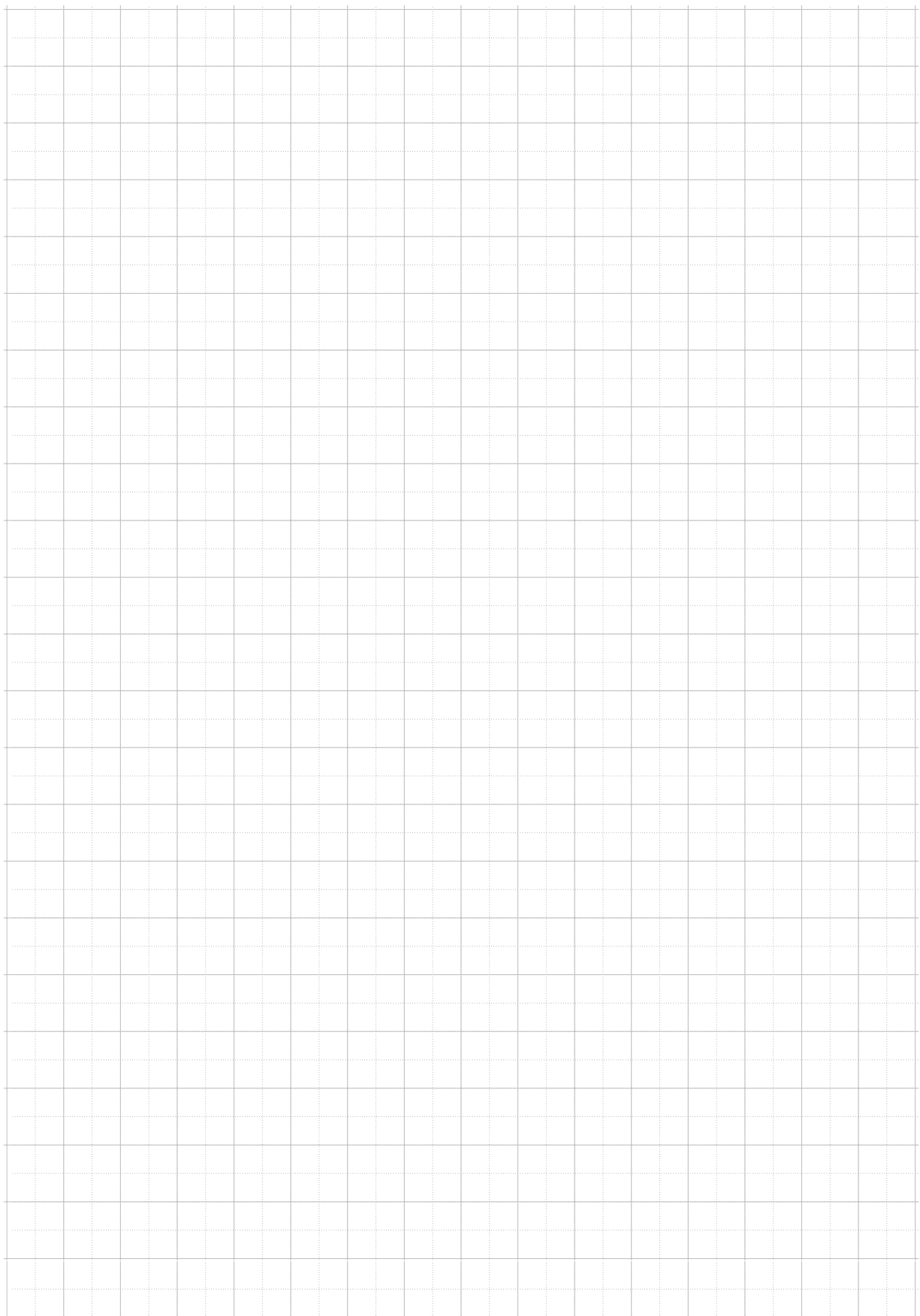
A continuación probaremos que el conjunto de todos los vectores en $\mathbb{R}^n$ junto con la suma (de vectores) y multiplicación por escalar forma un espacio vectorial.

Solución: Tomamos las operaciones $+$, $\cdot$ que conocemos y el conjunto

$$\mathcal{V} := \mathbb{R}^n = \left\{ \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} : v_1, v_2, \dots, v_n \in \mathbb{R} \right\}$$

Demostración. Verificamos las condiciones de la definición 3.1:





Dado que se cumplen todas las propiedades de la definición 3.1, es decir, 1), 1. $\{a, b, c, d\}$, 2) y 2. $\{a, b, c, d\}$, concluimos que $(\mathbb{R}^n, +, \cdot)$ es un espacio vectorial. $\square$

3.1.2. El espacio vectorial de matrices reales

El conjunto de las matrices con coeficientes reales (junto con la suma $+$ y el producto con escalar) es un espacio vectorial. Tomamos las operaciones $+, \cdot$ que conocemos y el conjunto

$$\mathcal{V} := \mathbb{R}^{m \times n} = \left\{ \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} : a_{ij} \in \mathbb{R} \right\}$$

Demostración. Verificamos las condiciones de la definición 3.1:

1) Para cada 2 matrices $A, B \in \mathbb{R}^{m \times n}$ se tiene que la suma es:

1.a) La matriz

1.b) “ $-A$ ” ejercicio.

1.c)

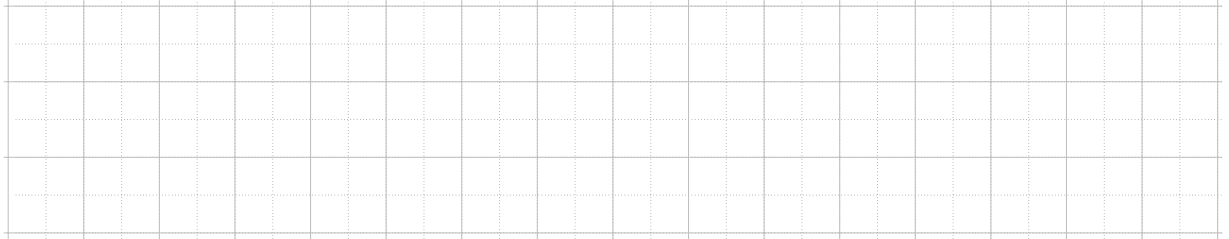
1.d)



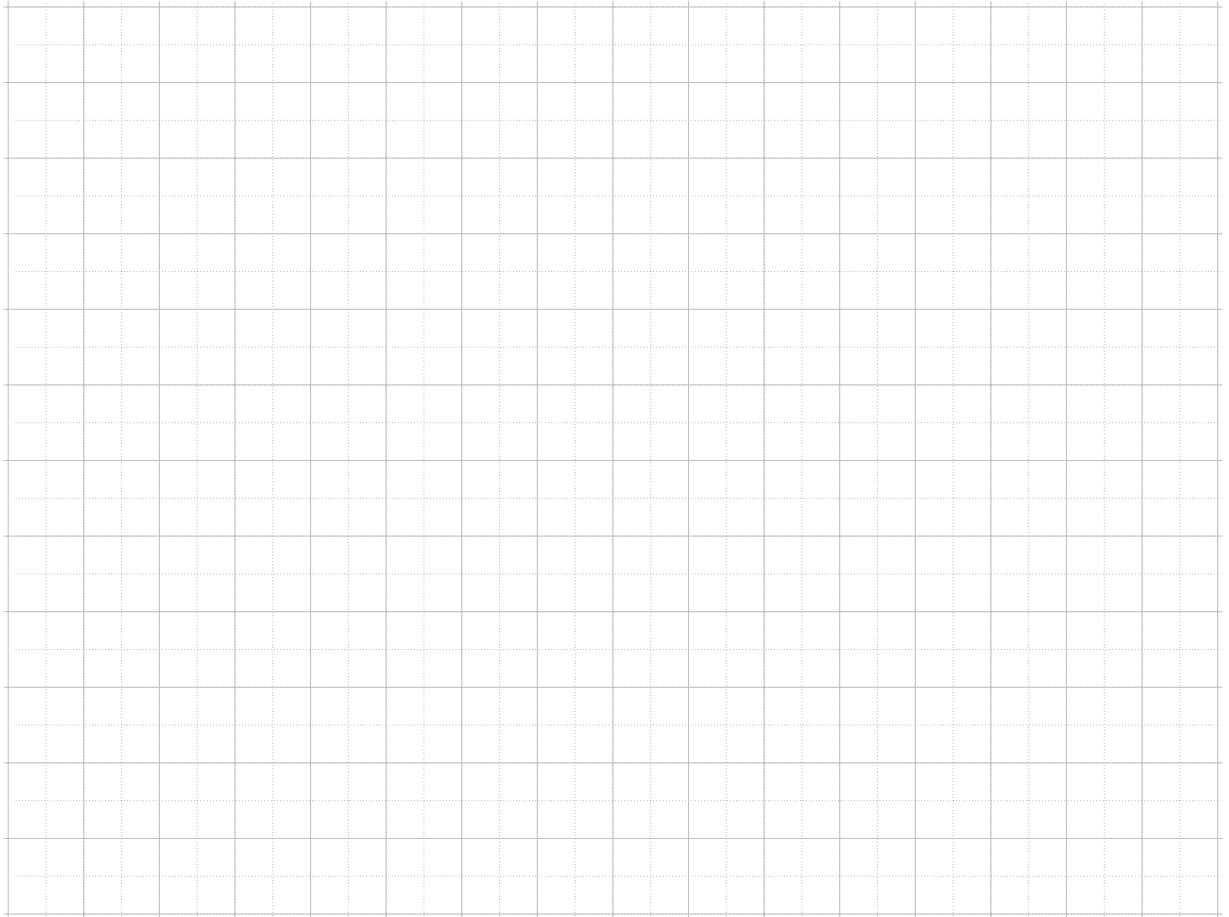
2)



2.a)

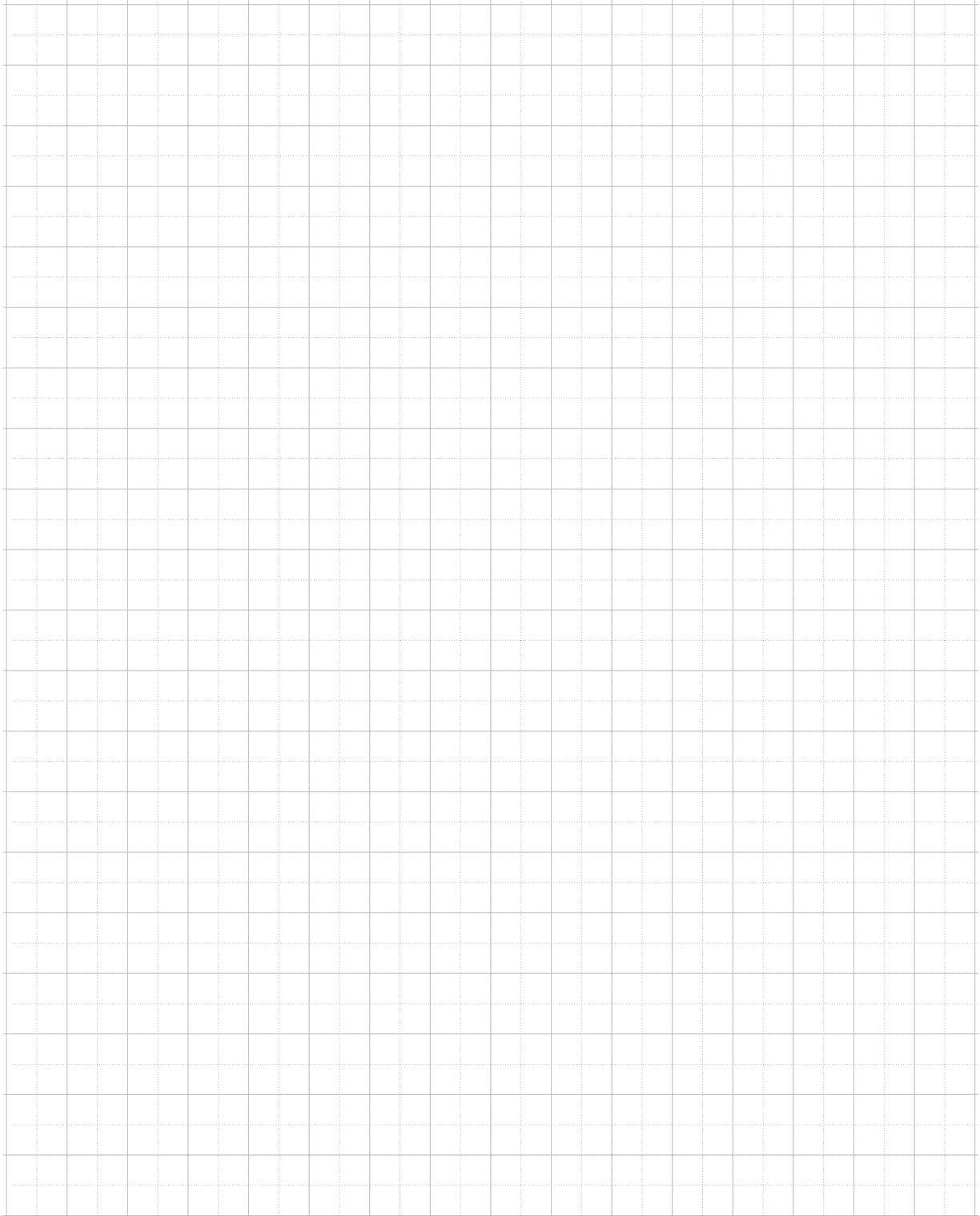


2.b)



2.c) ejercicio.

2.d)



Dado que se cumplen todas las propiedades de la definición 3.1, concluimos que $(\mathbb{R}^{m \times n}, +, \cdot)$ es un espacio vectorial. □



3.1.3. El espacio vectorial de polinomios hasta cierto grado

Veremos el siguiente caso especial:

El conjunto de todos los polinomios de grado ≤ 2 con coeficientes en $\mathbb{R}$ y con operaciones $+, \cdot$ definidas adecuadamente forma un espacio vectorial.

Para verificar si la definición 3.1 se cumple, tenemos que definir el conjunto $\mathcal{V}$ y las operaciones. Sea $\mathcal{V} := \mathcal{P}_2 = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ el conjunto de todos los polinomios de grado ≤ 2 con coeficientes en $\mathbb{R}$. Sean $u, w \in \mathcal{P}_2$, entonces $u(x) = a_1x^2 + b_1x + c_1$ & $w(x) = a_2x^2 + b_2x + c_2$ y la suma se define como:

$$(u + w)(x) := u(x) + w(x) = (a_1x^2 + b_1x + c_1) + (a_2x^2 + b_2x + c_2) \quad \text{para todo } x \in \mathbb{R} \quad (+)$$

y el producto con $\gamma \in \mathbb{R}$ se define por

$$(\gamma \cdot u)(x) := \gamma u(x) = \gamma(a_1x^2 + b_1x + c_1) \quad \text{para todo } x \in \mathbb{R} \quad (\times)$$

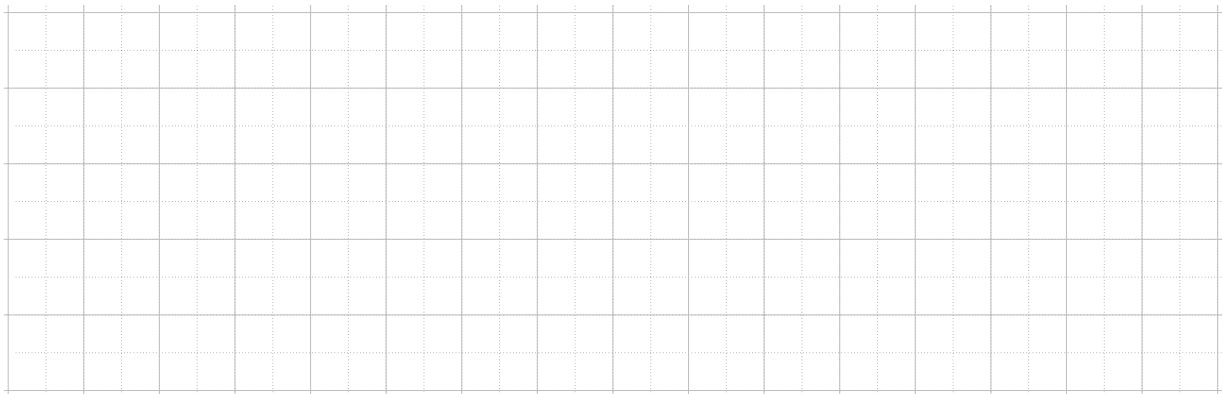
Ahora justificamos que $(\mathcal{P}_2, +, \cdot \overset{\text{mult.}}{\text{esc.}})$ es un espacio vectorial.

Demostración. Verificamos las condiciones de la definición 3.1:

1) ¿La suma es un polinomio?



1a) ¿Cuál elemento de $\mathcal{P}_2$ puede servir como neutro?



1b) ¿Cuál elemento de $\mathcal{P}_2$ puede servir inverso aditivo?



1d) Verificar si la suma $+$ es conmutativa



1c) Verificar si la suma $+$ es asociativa



2) Sea $\gamma \in \mathbb{R}$ y $u \in \mathcal{P}_2$. Por demostrar: $\gamma u \in \mathcal{P}_2$:



2a) Por demostrar: $1 \cdot u = u$:

2b) A continuación solamente se probará una de 2 propiedades distributivas, pero si en alguno de los exámenes le es requerido probar si algo es un espacio vectorial, entonces usted tendrá que probar todas las propiedades. Por demostrar: Para $r, s \in \mathbb{R}$ se tiene $(r + s)u = ru + su$:

2c) Por demostrar: $(rs)u = r(su)$:

Dado que se cumplen todas las propiedades de la definición 3.1, concluimos que $(\mathcal{P}_2, +, \cdot, \text{mult.}_{esc.})$ es un espacio vectorial. $\square$



Usando las mismas operaciones $(f + g)(x) = f(x) + g(x)$ & $(r \cdot f)(x) = r f(x)$ y las ideas en la demostración anterior se puede justificar lo siguiente:

- Los polinomios $\mathcal{P}_n$ de grado $\leq n$ con coeficientes en $\mathbb{R}$ y con $+, \cdot$ forman un espacio vectorial.
- Las funciones continuas con las operaciones $+, \cdot$ forman un espacio vectorial. Aquí, se ocupan resultados de cálculo, *e.g.* “la suma de funciones continuas es una función continua”.

3.2. Propiedades adicionales en espacios vectoriales

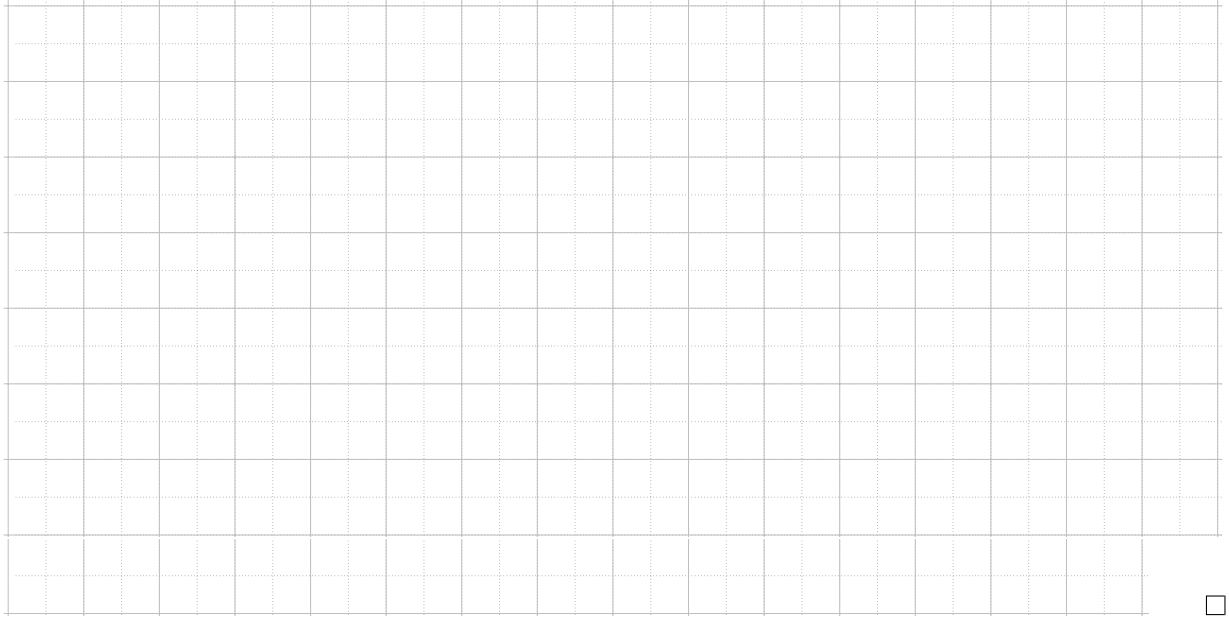
Dadas las propiedades mencionadas en la definición 3.1 los siguientes hechos serán validos:

Hecho 3.2.1:

Sea $(\mathcal{V}, +, \cdot, \cdot_{\text{esc.}}^{\text{mult.}})$ un espacio vectorial y $\vec{0} \in \mathcal{V}$ su elemento neutro (de la suma).

Entonces, para cada $\vec{v} \in \mathcal{V}$ se tiene $0 \cdot \vec{v} = \vec{0}$ donde 0 es el cero en $\mathbb{R}$.

Demostración. Sea $\vec{v} \in \mathcal{V}$. Ocuparemos propiedades del espacio vectorial (abstracto) para mostrar lo requerido:



□

Hecho 3.2.2 “cancelación”:

Sea $(\mathcal{V}, +, \cdot)$ un espacio vectorial. Demuestre que para $\vec{x}, \vec{y}, \vec{z} \in \mathcal{V}$ se cumple lo siguiente:

Si $\vec{x} + \vec{y} = \vec{x} + \vec{z}$, entonces $\vec{y} = \vec{z}$.

Demostración. Sean $\vec{x}, \vec{y}, \vec{z} \in \mathcal{V}$ que cumplen $\vec{x} + \vec{y} = \vec{x} + \vec{z}$, entonces



□

Hecho 3.2.3 “unicidad del neutro”:

Sea $(\mathcal{V}, +, \cdot)$ un espacio vectorial, y sean $\vec{x}, \vec{y} \in \mathcal{V}$.

Demuestre que si $\vec{x} + \vec{y} = \vec{x}$, entonces $\vec{y} = \vec{0}$, i.e. el elemento neutro.

Demostración. Sean $\vec{x}, \vec{y} \in \mathcal{V}$ tales que se cumple $\vec{x} + \vec{y} = \vec{x}$.

□

Hecho 3.2.4 “un producto con escalar es cero cuando uno de los factores lo es”:

Sea $(\mathcal{V}, +, \cdot)$ un espacio vectorial. Para cada $\vec{v} \in \mathcal{V}$ y $\gamma \in \mathbb{R}$ se tiene que

$$\gamma \vec{v} = \vec{0} \quad \Longleftrightarrow \quad \gamma = 0 \quad \text{ó} \quad \vec{v} = \vec{0}.$$

Demostración. Como es una proposición de la forma “ $\Longleftrightarrow$ ” mostraremos primero la suficiencia.

□

Hechos 3.2.5 “secundarios”:

Demuestre que si $(\mathcal{V}, +, \cdot)$ es un espacio vectorial, entonces

- a) Para cada $\vec{v} \in \mathcal{V}$ se tiene que $(-1)\vec{v} = -\vec{v}$.
- b) Para cada $\vec{v}, \vec{w} \in \mathcal{V}$ y todo $\gamma \in \mathbb{R}$ con $\gamma \neq 0$. Si $\gamma\vec{v} = \gamma\vec{w}$, entonces $\vec{v} = \vec{w}$.
- c) Para cada $\vec{v} \in \mathcal{V}$ y todo $\gamma, \delta \in \mathbb{R}$ con $\vec{v} \neq \vec{0}$. Si $\gamma\vec{v} = \delta\vec{v}$, entonces $\gamma = \delta$.

Ejercicio: Demuestre a), b) y c).

Ayuda: Cada caso se muestra usando la definición 3.1 y uno de los hechos anteriores.



Resumen.

La definición 3.1 “solo” pide verificar 10 propiedades.

Esas 10 permiten demostrar los 7 hechos adicionales anteriores.

$\therefore$ Para usar las $10 + 7 = 17$ propiedades, solo necesitamos verificar las 10 de la definición.

Todas esas propiedades aparecen en los números reales.

3.3. Subespacios vectoriales

Similar al concepto de “subconjunto” existe el siguiente concepto importante.



Definición 3.2. Sea $(\mathcal{V}, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ un espacio vectorial y sea $S \subseteq \mathcal{V}$ un subconjunto no vacío. Se dice que $(S, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ es un **sub-espacio vectorial** cuando es un espacio vectorial con las mismas operaciones.

El siguiente teorema nos permitirá decidir con mayor facilidad si un subconjunto $S \subseteq \mathcal{V}$ con las operaciones (del espacio vectorial “grande”) es ó no es sub-espacio vectorial.



Teorema 3.1. Sea $(\mathcal{V}, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ un espacio vectorial y sea $S \subseteq \mathcal{V}$ un subconjunto no vacío. Entonces, la tupa $(S, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ es un sub-espacio vectorial si y solo si

- i) $\vec{0} \in S$ “el neutro pertenece a S ”,
- ii) $\forall \vec{u}, \vec{v} \in S$ se cumple que $\vec{u} + \vec{v} \in S$ “la suma se queda en S ”,
- iii) $\forall \vec{v} \in S, \gamma \in \mathbb{R}$ se tiene que $\gamma\vec{v} \in S$ “el producto con escalar se queda en S ”.

Demostración. Es una proposición de la forma “ $\iff$ ”.

Mostraremos primero la dirección “ $\implies$ ”. Supongamos que $(S, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ es un sub-espacio vectorial. Entonces, por definición 3.2 $(S, +, \cdot \overset{\text{mult.}}{\underset{\text{esc.}}{}})$ también es un espacio vectorial. Por ser un espacio vectorial la definición 3.1 nos da la propiedad 1) que confirma ii), la propiedad 2) que confirma iii) y la propiedad 1a) que da el elemento neutro $\vec{0} \in S$ el cual es el neutro de $\mathcal{V}$ dado que $S \subseteq \mathcal{V}$ y puesto que $\vec{0}$ es único neutro en $\mathcal{V}$ por hecho 3.2.3.

Ahora, mostramos la dirección “ $\Leftarrow$ ”: Podemos usar i), ii), iii), y los hipótesis generales $S \subseteq \mathcal{V}$ y que $(\mathcal{V}, +, \cdot)$ es espacio vectorial. Para ser sub-espacio vectorial $(S, +, \cdot)$ debe ser un espacio vectorial y $S \subseteq \mathcal{V}$, es decir, solo falta verificar las condiciones de la definición 3.1 (página 67):

- 1) Dados $\vec{u}, \vec{v} \in S$ se tiene por hipótesis ii) que $\vec{u} + \vec{v} \in S$.
- 1a) Por i) $\vec{0} \in S$. Además, dado que $S \subseteq \mathcal{V}$ la propiedad $\vec{0} + \vec{v} = \vec{v}$ para $\vec{v} \in S$ se hereda del espacio vectorial $(\mathcal{V}, +, \cdot)$.
- 1b) Por demostrar $-\vec{v} \in S$:
Tomamos $\vec{v} \in S$, entonces por iii) y hecho 3.2.5a sabemos que $-\vec{v} = (-1)\vec{v} \in S$.
Además, dado que $S \subseteq \mathcal{V}$ la propiedad $\vec{v} + (-\vec{v}) = \vec{0}$ para $\vec{v} \in S$ se hereda del espacio vectorial $(\mathcal{V}, +, \cdot)$.
- 1c) Dado que $S \subseteq \mathcal{V}$ la asociatividad de la suma $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$ se hereda del espacio vectorial $(\mathcal{V}, +, \cdot)$.
- 1d) Dado que $S \subseteq \mathcal{V}$ la conmutatividad de la suma $\vec{u} + \vec{v} = \vec{v} + \vec{u}$ se hereda del espacio vectorial $(\mathcal{V}, +, \cdot)$.
- 2) Dados $\vec{u} \in S$ y $\gamma \in \mathbb{R}$ se tiene por hipótesis iii) que $\gamma\vec{u} \in S$.

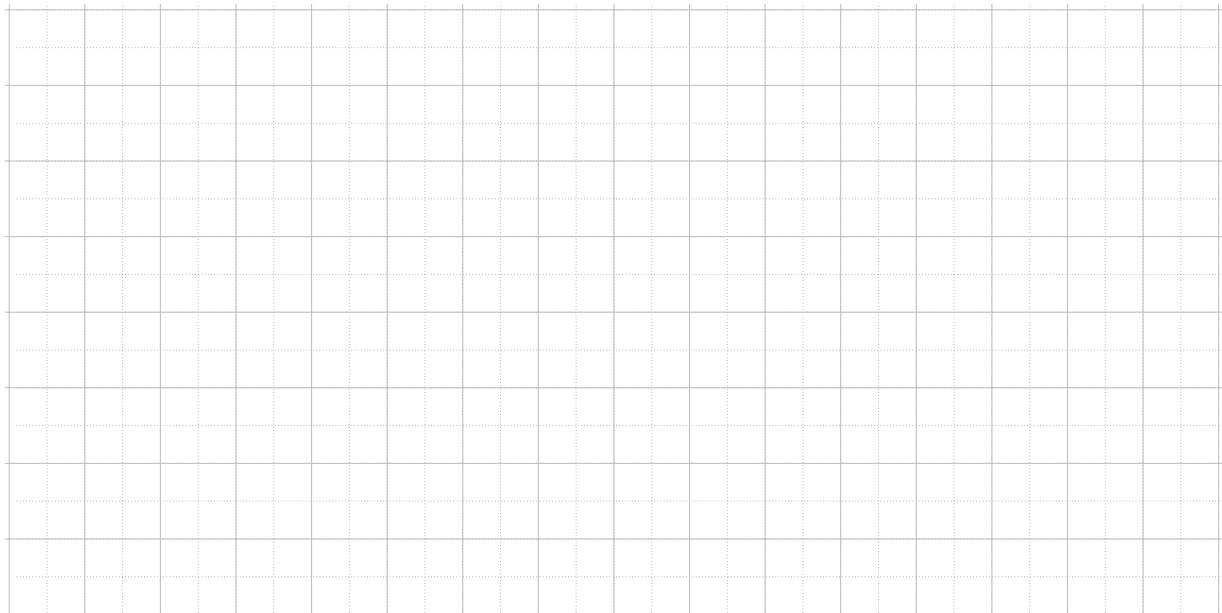
Ejercicio: Termina esa prueba demostrando lo que falta $(2. \{a, b, c, d\})$. □

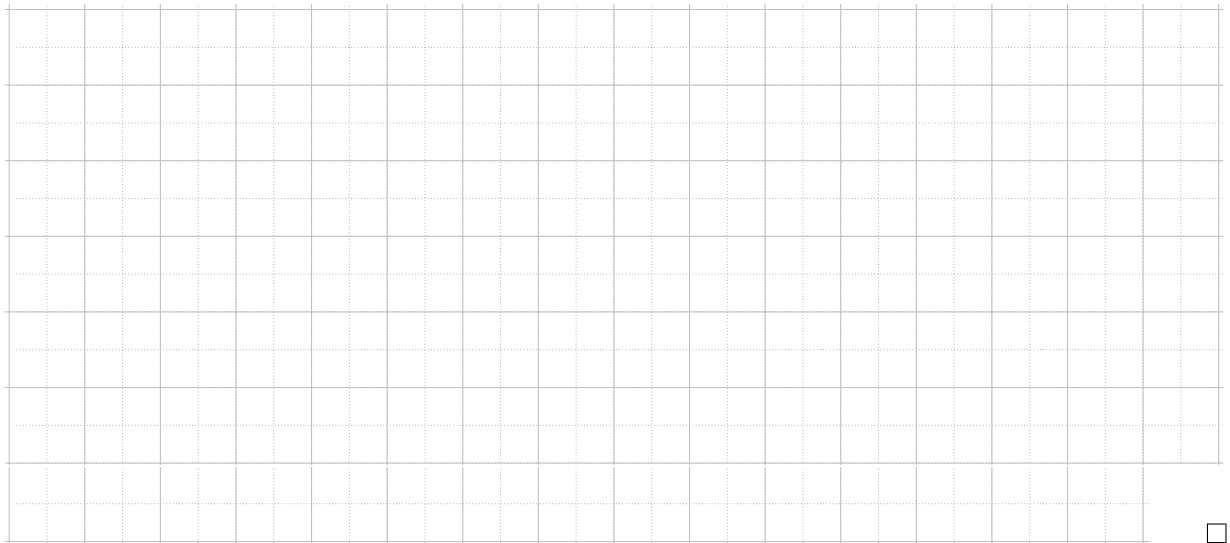
3.3.1. Ejemplo 1

En $\mathbb{R}^3$ consideremos al conjunto $S = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : 5x_1 + 7x_2 + 11x_3 = 0 \right\}$.

Demuestre que $(S, +, \cdot)$ es un subespacio vectorial de $(\mathbb{R}^3, +, \cdot)$.

Demostración.





3.3.2. Ejemplo 2

Se puede comprobar que el conjunto de polinomios $\mathcal{P}_2$ (de grado ≤ 2) bajo la suma & la multiplicación por escalar es un subespacio vectorial de $\mathcal{P}_n$ (polinomios de grado $\leq n$) cuando $2 \leq n$.

3.4. Sub-espacios asociados a una matriz

A continuación definimos los subespacios más importantes en matemáticas aplicadas.

Ejemplo visual:





Definición 3.3. El **núcleo (ó kernel ó espacio nulo)** de $A \in \mathbb{R}^{m \times n}$ es el conjunto de vectores en $\mathbb{R}^n$ que son enviados (a través de A) al cero. Se usan las siguientes notaciones:

$$\text{Null}(A) = \text{Ker}(A) := \left\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0} \right\}.$$

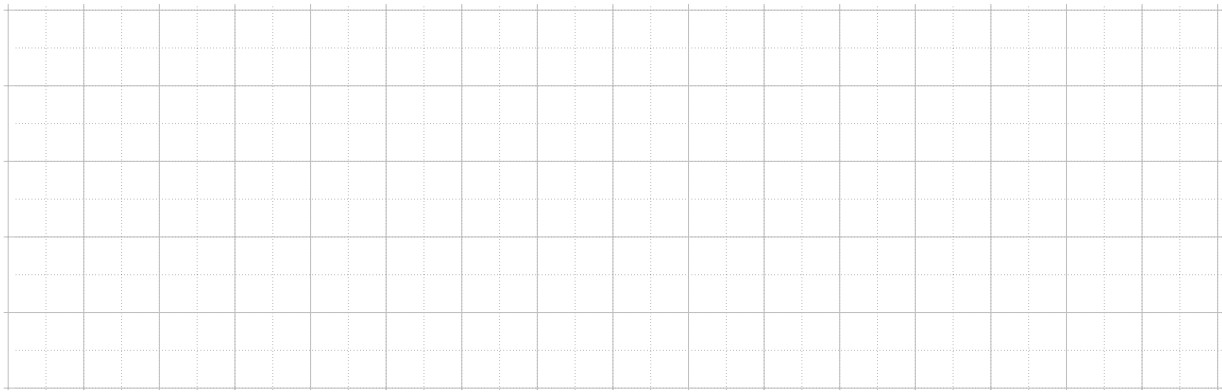
La **imagen** de $A \in \mathbb{R}^{m \times n}$ es el subconjunto de vectores en $\mathbb{R}^m$ que fueron obtenidos como productos de A con vectores en $\mathbb{R}^n$:

$$\text{Img}(A) := \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}.$$

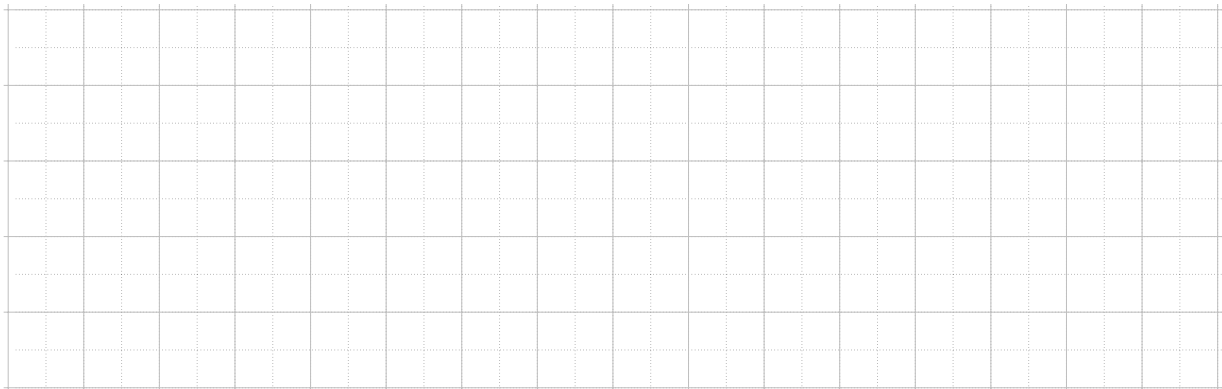
Vale la pena reconocer las siguientes conexiones:



$\{\vec{0}_n\} \subseteq \text{Ker}(A) \subseteq \mathbb{R}^n$ y consiste de las soluciones del sistema homogéneo $A\vec{x} = \vec{0}$.



$\{\vec{0}_m\} \subseteq \text{Img}(A) \subseteq \mathbb{R}^m$ y consiste de las posibles combinaciones lineales de las columnas de A .



El nombre “espacio nulo de A ” se debe al siguiente lema.



Lema 3.2. Sea A una matriz de tamaño $m \times n$. Entonces

- i) El conjunto $\text{Img}(A) = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\}$ es un sub-espacio vectorial de $\mathbb{R}^m$.
- ii) El conjunto $\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0}\}$ es un sub-espacio vectorial de $\mathbb{R}^n$, por lo tanto, el nombre alternativo “espacio nulo”.

3. Espacios vectoriales

Demostración.



Ejercicio: El número ii) se deja de tarea. Sugerencia: aprovecha el lema 1.2.

□

4. Combinaciones lineales

En esta sección profundizamos nuestro conocimiento de combinaciones lineales e introducimos los conceptos: “independencia lineal”, “espacios generados”, “bases” y “dimensión”.

Introducción en el plano.

Empezaremos nuestro estudio en $\mathbb{R}^2$, es decir en

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

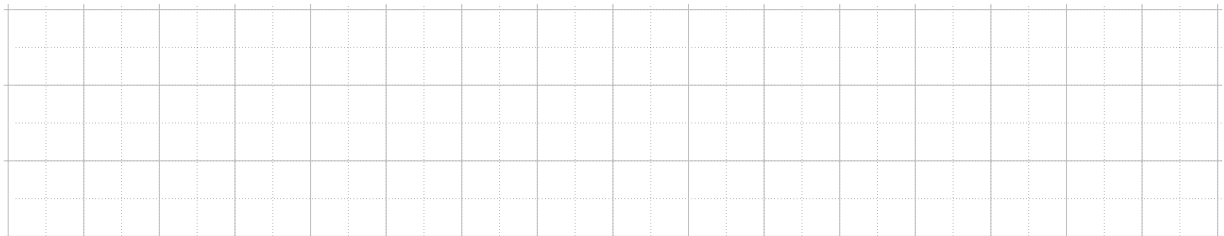
Supongamos que tenemos dos vectores $\vec{u}, \vec{w} \in \mathbb{R}^2$ y escalares $\alpha, \beta \in \mathbb{R}$.

Podemos visualizar “cuatro casos” de las combinaciones lineales $\alpha\vec{u} + \beta\vec{w}$.

Ejemplo (visualizamos): Sean $\vec{u} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$ & $\vec{w} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

Caso 1) Si $\alpha = 0$ & $\beta = 0$, entonces obtenemos

“el vector cero”



Caso 2) Si $\alpha \in \mathbb{R}$ & $\beta = 0$, entonces obtenemos $\alpha\vec{u}$

“ α -veces la dirección de $\vec{u}$ ”



Caso 3) $\alpha = 1$ & $\beta = 3$

“suma de vectores con $\alpha, \beta > 0$ ”



Caso 4) Si $\alpha = 1$ & $\beta = -1$, entonces obtenemos

“resta de vectores” (dibujo importante)



Observación: Si quiero más que una recta, entonces necesito dos vectores. Más aún, la combinación lineal $\alpha\vec{u} + \beta\vec{w}$ permite llegar a cada punto en el plano solo cuando $\vec{w} \neq \alpha\vec{u}$ para $\alpha \in \mathbb{R}$ y $\vec{u} \neq \vec{0}$, es decir, cuando $\alpha\vec{u}$ genera una recta y $\vec{w}$ no es parte de esa recta.

Un concepto expuesto aquí es: “ $\vec{u}$ y $\vec{w}$ contribuyen direcciones diferentes”.

Se dice que un vector $\vec{u}$ es *independiente* si $\vec{u} \neq \vec{0}$.

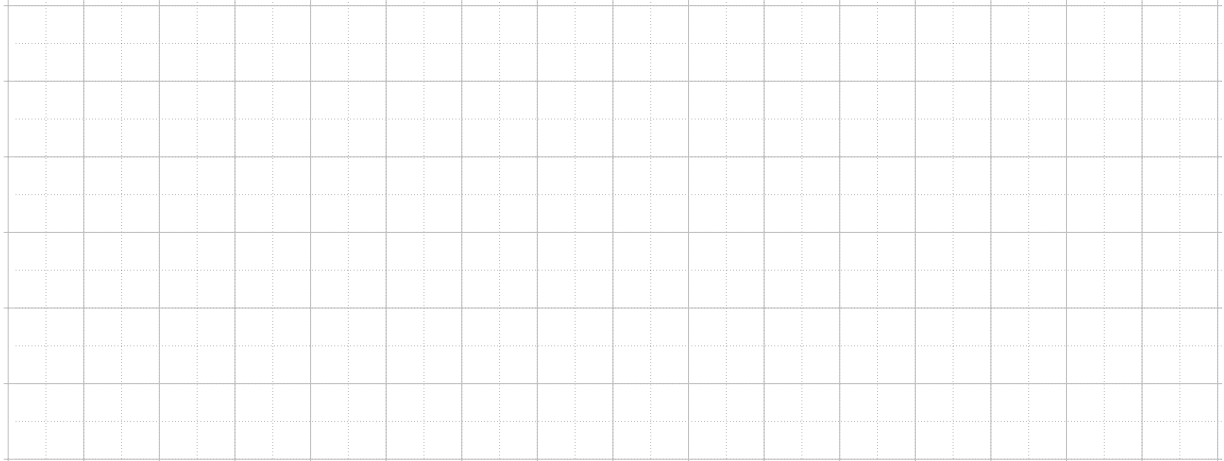
Se dice que dos vectores $\vec{u}, \vec{w}$ son *independientes* si $\vec{u} \neq \vec{0}$ y $\vec{w} \neq \alpha\vec{u}$ para $\alpha \in \mathbb{R}$.

Introducción en el espacio.

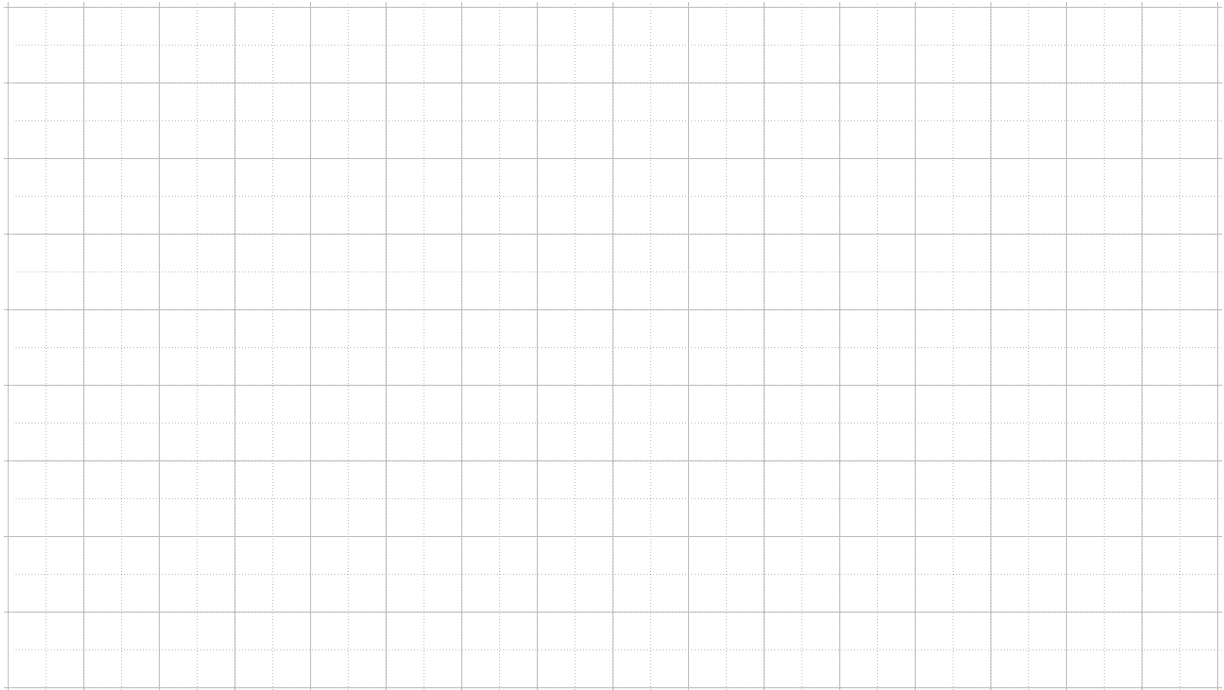
Sea $\alpha\vec{u} + \beta\vec{v} + \gamma\vec{w}$ una combinación lineal de vectores $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$ escalados con $\alpha, \beta, \gamma \in \mathbb{R}$.

Esta combinación lineal puede tomar tres formas:

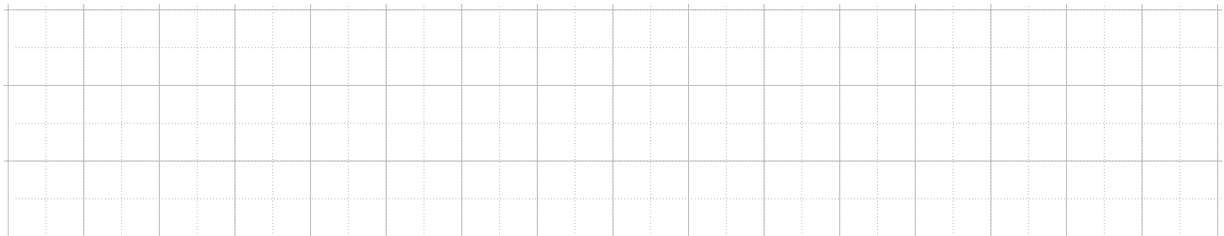
Caso I: Si $\vec{u} \neq \vec{0}$ y $\vec{v}, \vec{w} \in \{\alpha\vec{u}: \alpha \in \mathbb{R}\}$, entonces tenemos una recta.



Caso II: Si $\vec{u}, \vec{v}$ son independientes pero $\vec{w} \in \{\alpha\vec{u} + \beta\vec{v}: \alpha, \beta \in \mathbb{R}\}$, entonces tenemos un plano.



Caso III: Si $\vec{u}, \vec{v}$ son independientes y $\vec{w} \notin \{\alpha\vec{u} + \beta\vec{v}: \alpha, \beta \in \mathbb{R}\}$, entonces se llena $\mathbb{R}^3$.



Se dice que tres vectores $\vec{u}, \vec{v}, \vec{w}$ son *independientes* si $\vec{w} \neq \alpha\vec{u} + \beta\vec{v}$ para $\alpha, \beta \in \mathbb{R}$ y adicionalmente $\vec{u}, \vec{v}$ son independientes.

En general



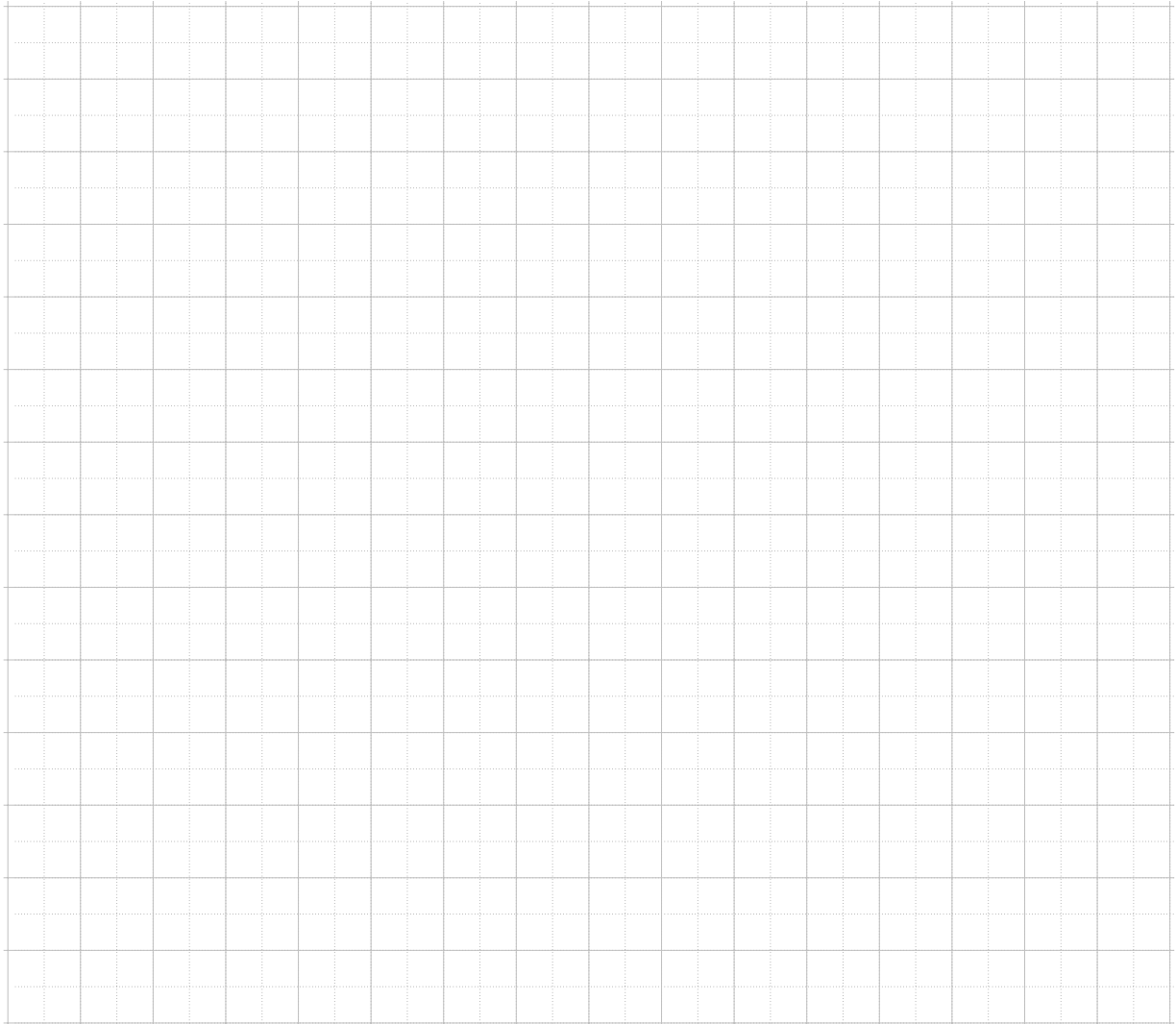
Definición 4.1. Sea $(\mathcal{V}, +, \cdot_{\text{esc.}}^{\text{mult.}})$ un espacio vectorial. Se dice que **un vector** $\vec{w} \in \mathcal{V}$ es **combinación lineal** de los vectores $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \in \mathcal{V}$ si existen escalares $x_1, x_2, \dots, x_p \in \mathbb{R}$ tales que $\vec{w}$ se puede expresar como

$$\vec{w} = x_1 \vec{v}_1 + x_2 \vec{v}_2 + \dots + x_p \vec{v}_p$$

donde los escalares $x_1, x_2, \dots, x_p$ reciben el nombre de **coeficientes de una combinación lineal**.

Ejemplos: Considere los vectores $\vec{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ & $\vec{v}_2 = \begin{bmatrix} 1/2 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3$.

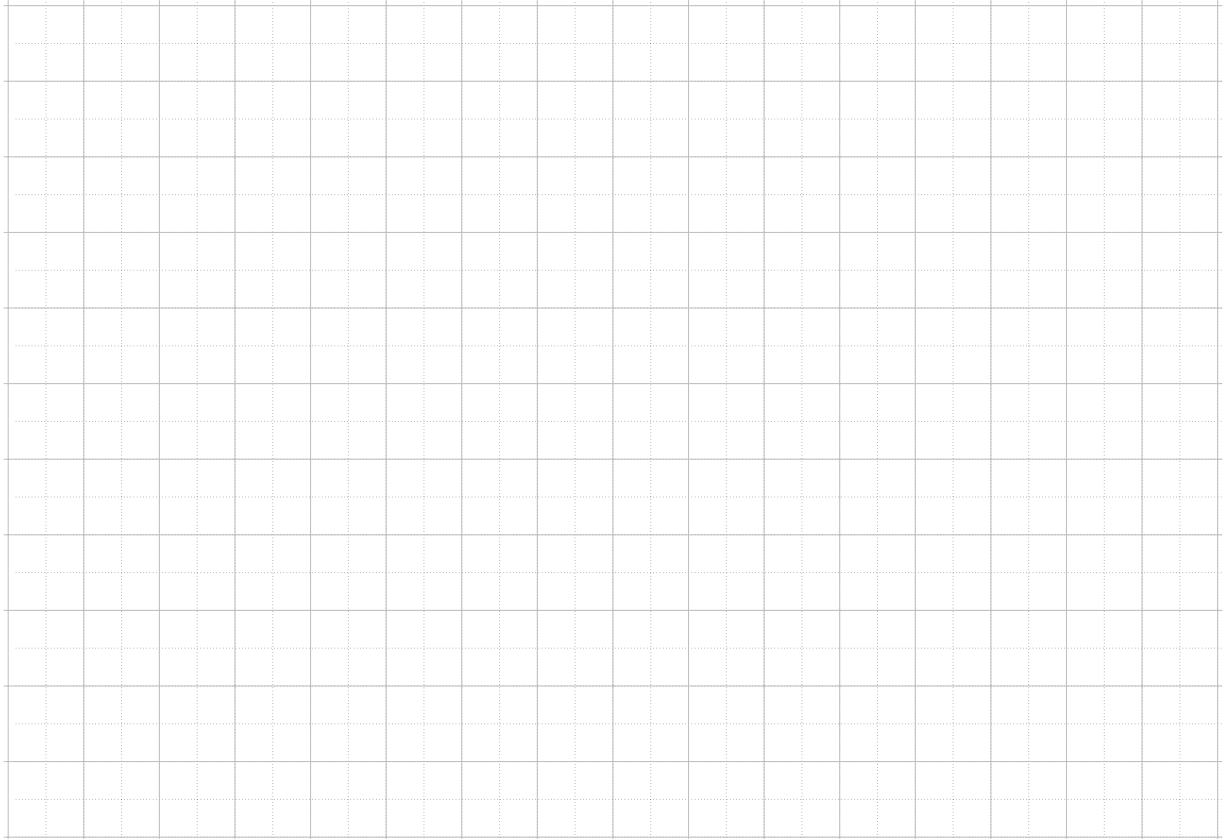
¿Para cuáles valores $\gamma \in \mathbb{R}$ el vector $\vec{w} = \begin{bmatrix} 1 \\ 2 \\ \gamma \end{bmatrix}$ es una combinación lineal de $\vec{v}_1$ & $\vec{v}_2$?



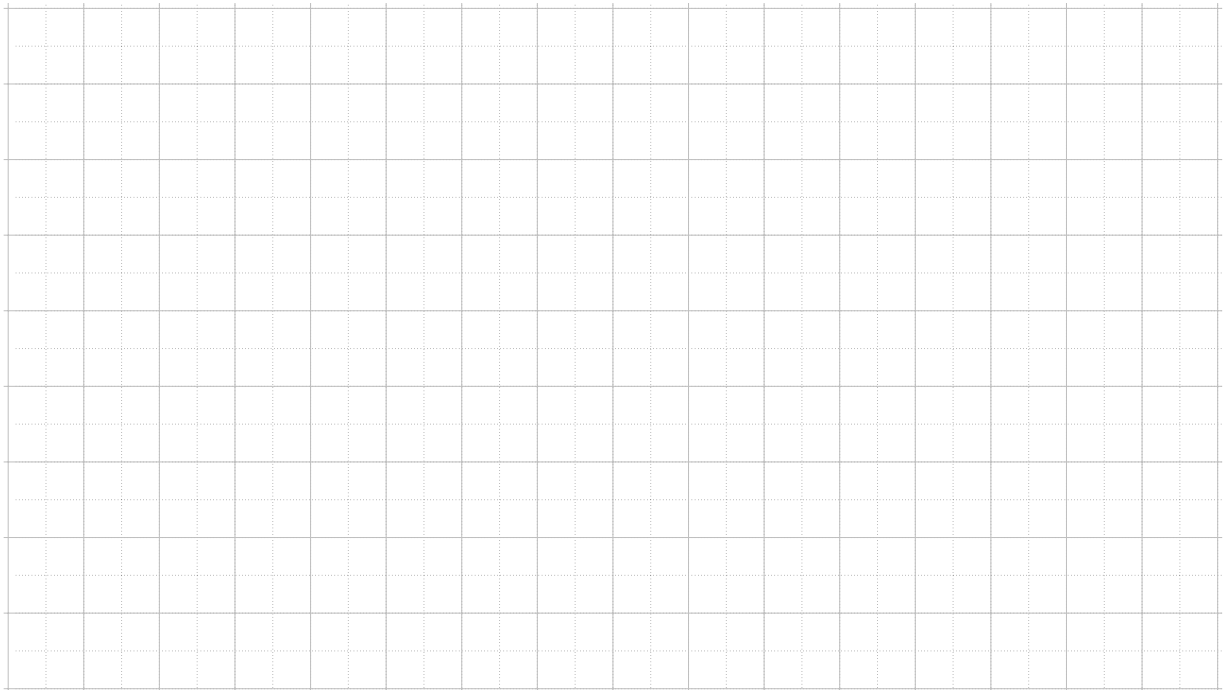
Ejercicio: Considere los vectores $\vec{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$ & $\vec{v}_2 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \in \mathbb{R}^3$. Compruebe si los vectores $\vec{u} = \begin{bmatrix} 4 \\ -2 \\ 5 \end{bmatrix}$ & $\vec{w} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ son combinación lineal de los vectores $\vec{v}_1$ & $\vec{v}_2$.



I) Tenemos que ver si existen escalares $x_1, x_2 \in \mathbb{R}$ tales que $\vec{u} = x_1\vec{v}_1 + x_2\vec{v}_2$, es decir



II) Tenemos que ver si existen escalares $x_1, x_2 \in \mathbb{R}$ tales que $\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2$, es decir



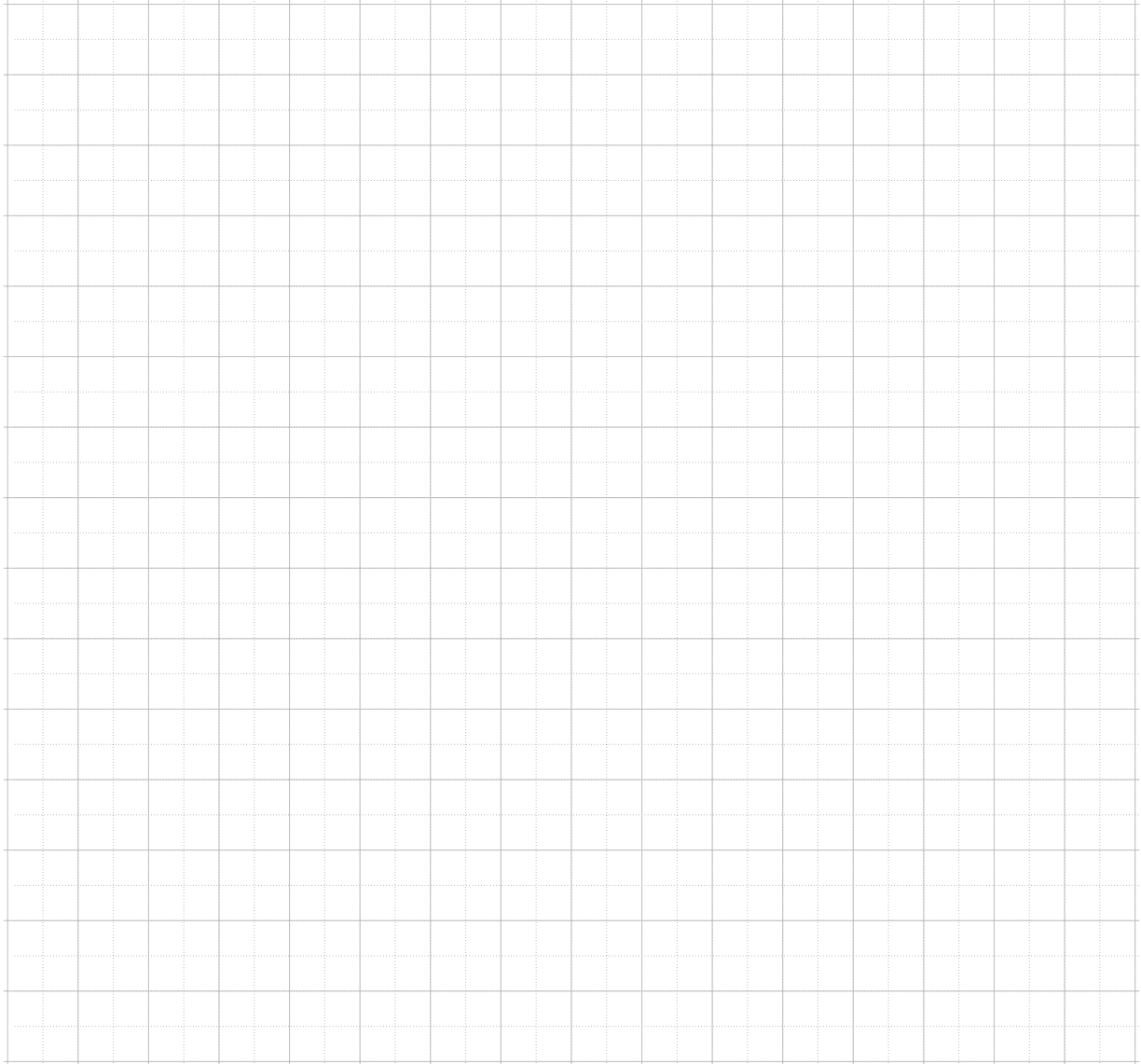
Resultado y conclusión: El sistema es inconsistente. Por eso, concluimos que el vector $\vec{w}$ NO se puede escribir como combinación lineal de $\vec{v}_1$ & $\vec{v}_2$. Otra formulación: No se pueden encontrar escalares $x_1, x_2 \in \mathbb{R}$ que satisfagan $\vec{w} = x_1\vec{v}_1 + x_2\vec{v}_2$.

4. Combinaciones lineales

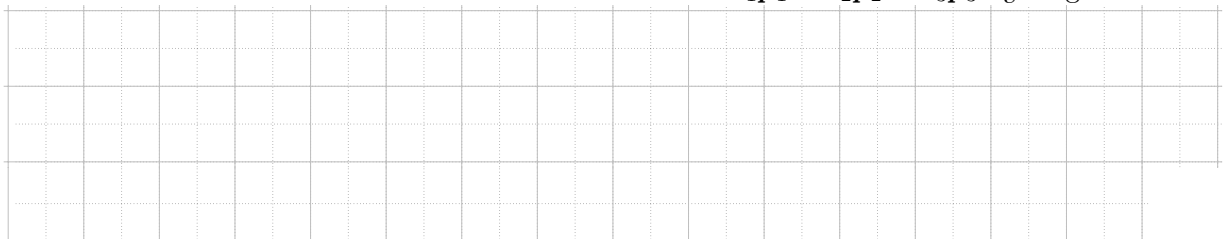
Ejemplo: Expresar el polinomio $\vec{w} = t^2 + 4t - 3$ del espacio vectorial $(\mathbb{P}_2, +, \cdot)$ como combinación lineal de los polinomios $\vec{p}_1 = t^2 - 2t + 5$, $\vec{p}_2 = 2t^2 - 3t$ y $\vec{p}_3 = t + 3$.

Sol: Supongamos que podemos escribir $\vec{w} = x_1\vec{p}_1 + x_2\vec{p}_2 + x_3\vec{p}_3$, donde $x_1, x_2, x_3 \in \mathbb{R}$:

$$\vec{w} = x_1\vec{p}_1 \quad + x_2\vec{p}_2 \quad + x_3\vec{p}_3$$



Prueba: Del sistema anterior obtenemos la combi. lineal $x_1\vec{p}_1 + x_2\vec{p}_2 + x_3\vec{p}_3$. ¿Es igual a $\vec{w}$?



4.1. Conjuntos generados



Definición 4.2. Sea $(\mathcal{V}, +, \cdot, \text{mult.}_{\text{esc.}})$ un espacio vectorial. Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \in \mathcal{V}$. Entonces, el **conjunto generado** por $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ denotado por $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$ es el conjunto de todas las combinaciones lineales de los vectores, *i.e.*

$$\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\} = \left\{ x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p \mid x_1, x_2, \dots, x_p \in \mathbb{R} \right\}.$$

Las siguientes notaciones representan al mismo conjunto:

$$\text{gen}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\} = \langle \vec{v}_1, \vec{v}_2, \dots, \vec{v}_p \rangle = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}.$$

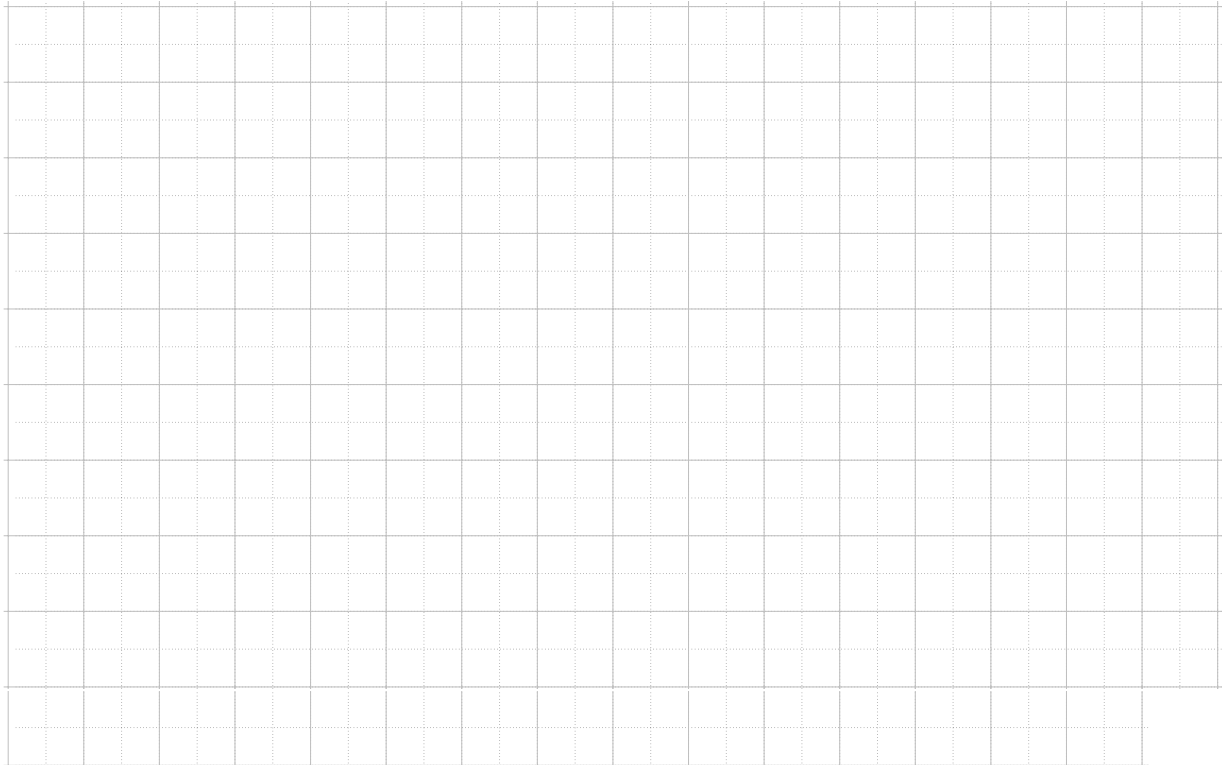
Ejemplo: Sea $\vec{v} \in \mathbb{R}^n$, entonces $\text{span}\{\vec{v}\} = \{x_1\vec{v} \mid x_1 \in \mathbb{R}\}$.

Si $\vec{v} = \vec{0}$, entonces



Si $\vec{v} \neq \vec{0}$, entonces el $\text{span}\{\vec{v}\} = \{x_1\vec{v} \mid x_1 \in \mathbb{R}\}$ es una recta infinita en $\mathcal{V}$.

Esa recta pasa por el origen (ya que $0\vec{v} = \vec{0}$) & va en dirección de $\vec{v}$ ó $-\vec{v}$.



Por definición tenemos $\vec{v} \in \text{span}\{\vec{w}\}$ si y solo si existe $x_1 \in \mathbb{R}$ tal que $\vec{v} = x_1\vec{w}$.



Si $\mathcal{V} = \mathbb{R}^n$ entonces por definición $\vec{v}_1, \dots, \vec{v}_p \in \mathbb{R}^n$ y $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \mathbb{R}^n$.

4. Combinaciones lineales

Ejemplo: Sean $\vec{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$ & $\vec{v}_2 = \begin{bmatrix} 1/2 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3$. Describe $\text{span}\{\vec{v}_1, \vec{v}_2\}$.

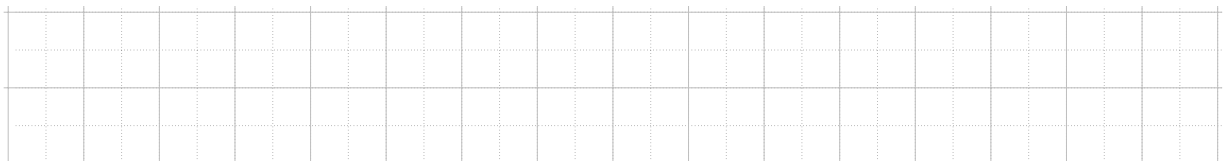
Sol:

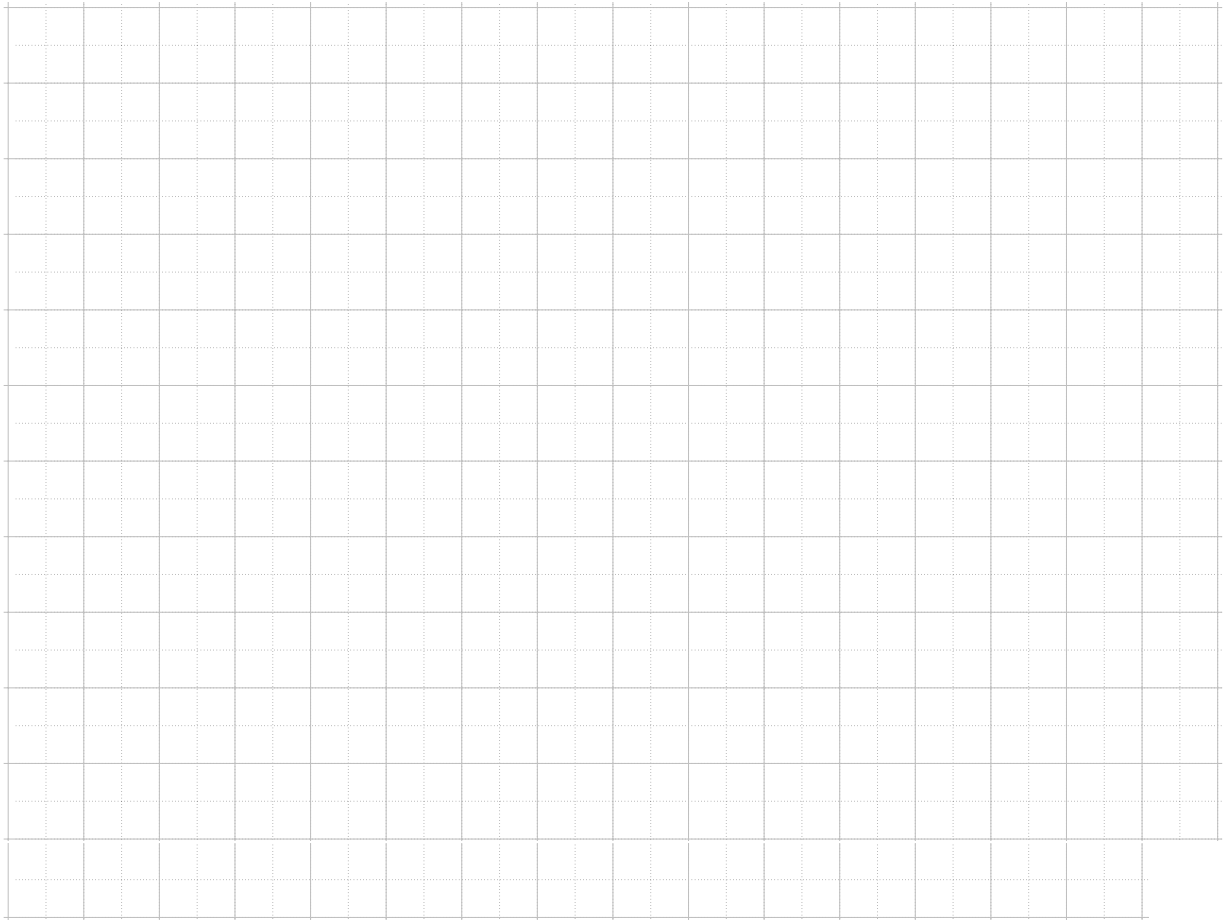


Ejemplo: Sean $\vec{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix}$ & $\vec{v}_2 = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^3$. Describa al conjunto $\text{span}\{\vec{v}_1, \vec{v}_2\}$.

Sol: Buscamos todos los puntos dentro de $\text{span}\{\vec{v}_1, \vec{v}_2\} = \left\{ x_1 \begin{bmatrix} 2 \\ 0 \\ 5 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} \mid x_1, x_2 \in \mathbb{R} \right\}$, es decir, buscamos cuales condiciones deben cumplir. Sabemos

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \text{span}\{\vec{v}_1, \vec{v}_2\} \iff \begin{pmatrix} a \\ b \\ c \end{pmatrix} = x_1 \begin{pmatrix} 2 \\ 0 \\ 5 \end{pmatrix} + x_2 \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}$$

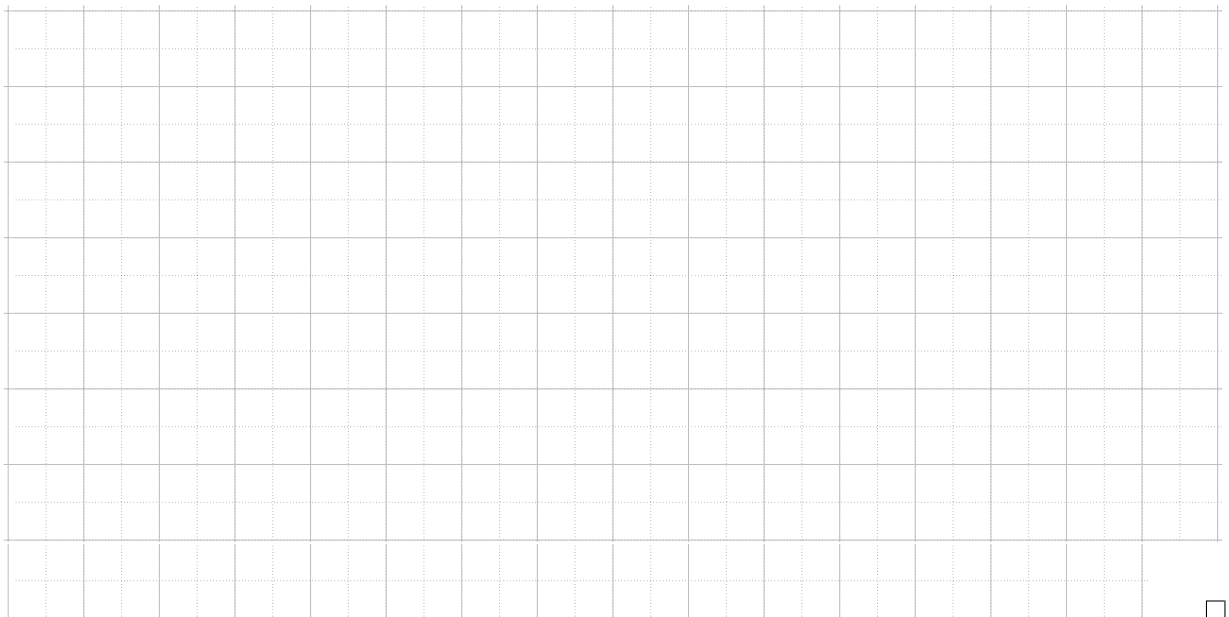




IMPORTANTE : Si $\vec{x}_*$ es solución de un sistema $A\vec{x} = \vec{b}$, entonces

$$\vec{b} \in \text{span} \{A_{[:,1]}, A_{[:,2]}, \dots, A_{[:,n]}\} = \text{Img}(A)$$

Demostración.



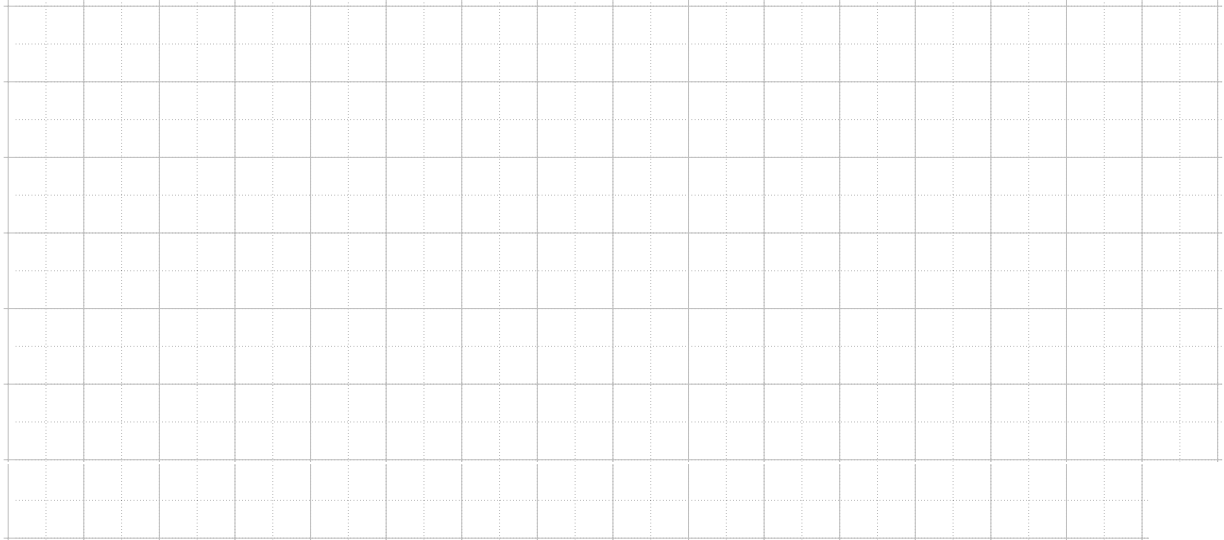
4. Combinaciones lineales

Los siguientes ejemplos motivan el uso de “elementos de $\mathcal{V}$ ” para definir un conjunto generado.

Ejemplo: En página 88 vimos que un polinomio $\vec{w}$ de grado 2 era combinación lineal de $\vec{p}_1, \vec{p}_2, \vec{p}_3$. El resultado era: $\vec{w} = t^2 + 4t - 3 = -3\vec{p}_1 + 2\vec{p}_2 + 4\vec{p}_3$, por lo cual $\vec{w} \in \text{span}\{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$, es decir “ $\vec{w}$ pertenece al conjunto generado por $\vec{p}_1, \vec{p}_2, \vec{p}_3$.”

Ejercicio: Considere los vectores $\vec{w} = 1 + t + 4t^2$  $+ 5t + t^2$ en $(\mathbb{P}_2, +, \cdot)$ y determine si $\vec{w}$ & $\vec{v}$ llacen en $\text{span}\{1 + 2t - t^2, 3 + 5t + 2t^2\}$.

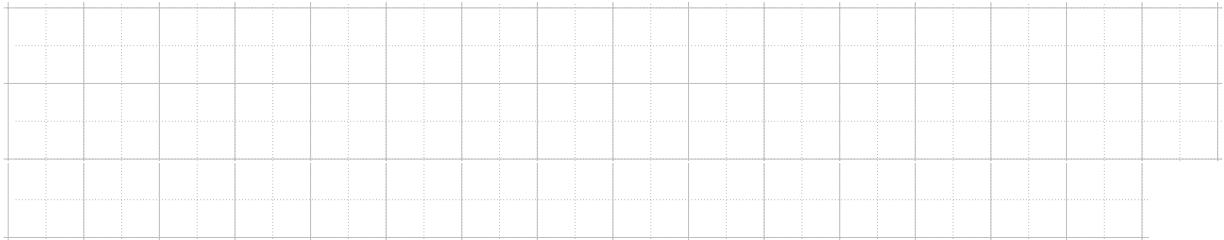
Inicio de solución:



Ejemplo: Demuestre que el conjunto $\mathbb{R}^{2 \times 2}$ de matrices 2×2 es generado por el conjunto

$$\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

Lo anterior se debe a que cualquier matriz de (2×2) de la forma $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ se puede escribir como combinación lineal de las matrices arriba, puesto que



Ejemplo: Demuestre que $\mathbb{P}_n$ es conjunto de polinomios de grado (a los más) n es generado por el conjunto $\{1, x, x^2, \dots, x^n\}$.

Solución: Todo polinomio $\vec{p} \in \mathbb{P}_n$ de grado $\leq n$ se puede escribir como

$$\vec{p} = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n = a_0 \cdot 1 + a_1x + a_2x^2 + a_3x^3 + \dots + a_nx^n$$

con escalares $a_0, a_1, \dots, a_n \in \mathbb{R}$, lo cual es una combinación lineal de elementos de $\{1, x, x^2, \dots, x^n\}$. Con eso hemos confirmado la igualdad $\text{span}\{1, x, x^2, \dots, x^n\} = \mathbb{P}_n$.

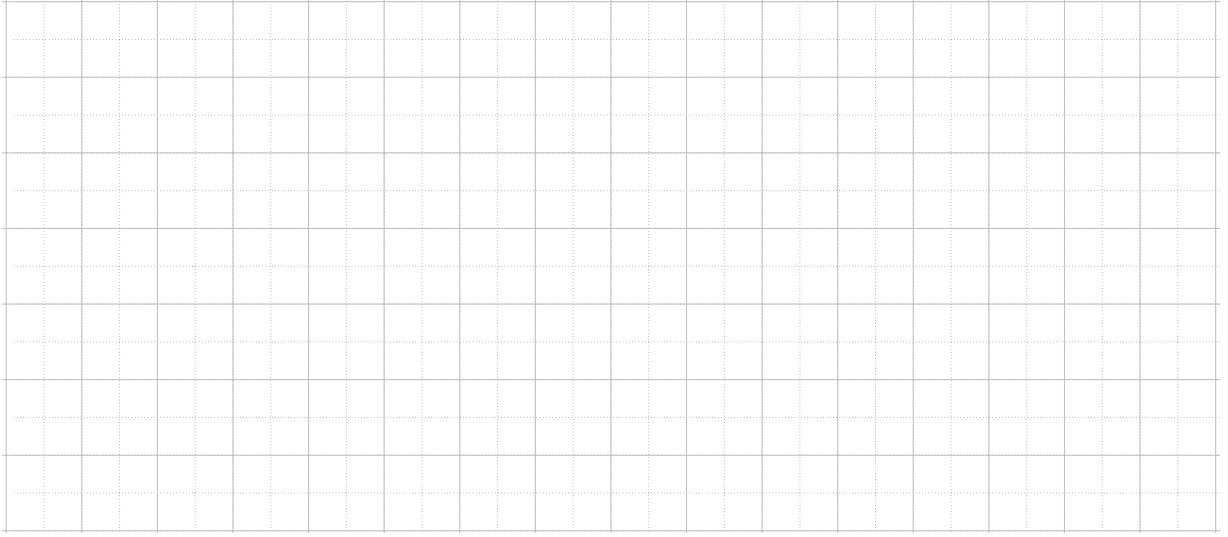


Lema 4.1. Sea $(\mathcal{V}, +, \cdot \text{mult.}_{\text{esc.}})$ un espacio vectorial. Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p, \vec{v}_{p+1} \in \mathcal{V}$ “vectores”.
Entonces:

- i) $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\} \subseteq \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p, \vec{v}_{p+1}\}$
- ii) Si $\vec{v}_{p+1} \in \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$, entonces $\text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\} = \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p, \vec{v}_{p+1}\}$
- iii) Sea $S := \text{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p\}$, entonces $\vec{0} \in S$ y $(S, +, \cdot \text{mult.}_{\text{esc.}})$ es un espacio vectorial.

Demostración.

- i) Sea $\vec{w} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$, entonces



- ii) Por i) ya sabemos $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \text{span}\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}\}$.

Resta demostrar que $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} \supseteq \text{span}\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}\}$.

Sea $\vec{w} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}\}$, entonces existen escalares $x_1, \dots, x_p \in \mathbb{R}$ tales que

$$\vec{w} = x_1 \vec{v}_1 + \dots + x_p \vec{v}_p + x_{p+1} \vec{v}_{p+1}$$

Por otro lado, por la hipótesis $\vec{v}_{p+1} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$ existen escalares $a_1, \dots, a_p \in \mathbb{R}$ tales que

$$\vec{v}_{p+1} = a_1 \vec{v}_1 + \dots + a_p \vec{v}_p$$

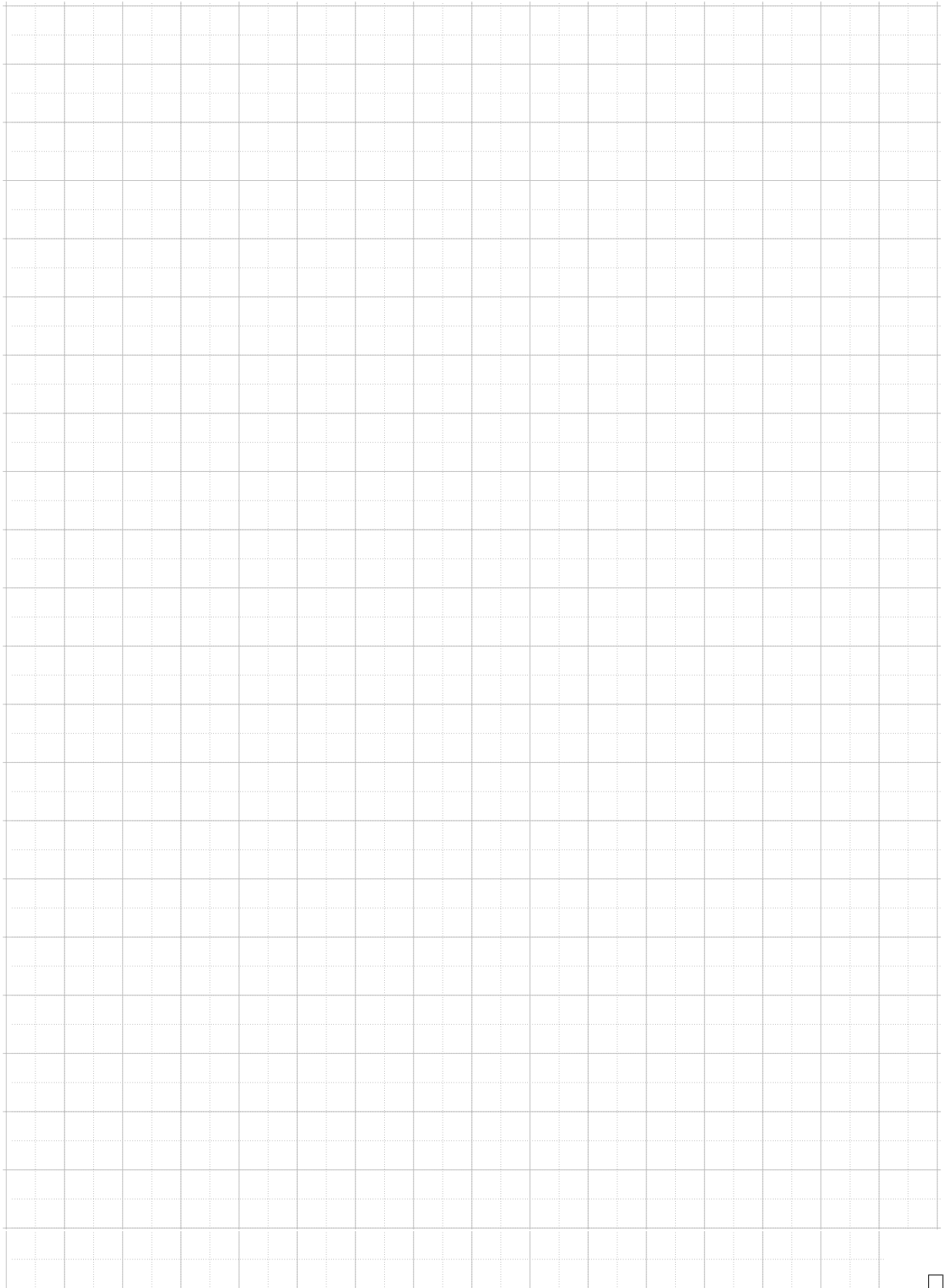
Sustituyendo esa identidad en la anterior de $\vec{w}$ y usando definición 3.1 obtenemos

$$\begin{aligned} \vec{w} &= x_1 \vec{v}_1 + \dots + x_p \vec{v}_p + x_{p+1} (a_1 \vec{v}_1 + \dots + a_p \vec{v}_p) \\ &\stackrel{2b)}{=} x_1 \vec{v}_1 + \dots + x_p \vec{v}_p + x_{p+1} (a_1 \vec{v}_1) + \dots + x_{p+1} (a_p \vec{v}_p) \\ &\stackrel{2c)}{=} x_1 \vec{v}_1 + \dots + x_p \vec{v}_p + (x_{p+1} a_1) \vec{v}_1 + \dots + (x_{p+1} a_p) \vec{v}_p \\ &\stackrel{1d), 2b)}{=} (x_1 + x_{p+1} a_1) \vec{v}_1 + \dots + (x_p + x_{p+1} a_p) \vec{v}_p \\ &= \gamma_1 \vec{v}_1 + \gamma_2 \vec{v}_2 + \dots + \gamma_p \vec{v}_p \quad \text{con} \quad \gamma_i = (x_i + x_{p+1} a_i) \in \mathbb{R}. \end{aligned}$$

Concluimos que $\vec{w} \in \text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$ ya que es combinación lineal de estos vectores.

4. Combinaciones lineales

iii) Dado que $(\mathcal{V}, +, \cdot)$ es espacio vectorial, aprovechamos el teorema 3.1.



Ejemplo: El conjunto generado permite cambiar elementos.

Sean $\vec{u}$ & $\vec{v}$ dos vectores en el espacio vectorial $(\mathcal{V}, +, \cdot)$. Demuestre que

$$\text{span}\{\vec{u}, \vec{v}\} = \text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\}.$$

Demostración. “ $\subseteq$ ”: Por demostrar $\text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\} \subseteq \text{span}\{\vec{u}, \vec{v}\}$.



“ $\supseteq$ ”: Por demostrar $\text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\} \supseteq \text{span}\{\vec{u}, \vec{v}\}$. Esa parte se simplifica si justificamos que $\vec{u}, \vec{v}$ son elementos de $\text{span}\{\vec{u} + 2\vec{v}, \vec{u} - \vec{v}\}$. Por eso determinamos sus combinaciones lineales:

$$\begin{aligned}\vec{u} &= x_1(\vec{u} + 2\vec{v}) + x_2(\vec{u} - \vec{v}) = (x_1 + x_2)\vec{u} + (2x_1 - x_2)\vec{v} \\ \& \quad \vec{v} &= x_3(\vec{u} + 2\vec{v}) + x_4(\vec{u} - \vec{v}) = (x_3 + x_4)\vec{u} + (2x_3 - x_4)\vec{v}.\end{aligned}$$

La primera igualdad se cumple para $x_1 = 1/3, x_2 = 2/3$ y la segunda para $x_3 = 1/3, x_4 = -1/3$. Ahora, sea $\vec{w} = \alpha\vec{u} + \beta\vec{v} \in \text{span}\{\vec{u}, \vec{v}\}$, entonces



□



Debido al ejemplo anterior podemos escribir cada punto en $\mathbb{R}^2$ en distintas coordenadas. Por ejemplo, en $\text{span}\left\{\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$ o en $\text{span}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\}$.

4.2. Dependencia e independencia lineal

Esta sección generaliza la “independencia de vectores” que vimos en las clases 10 y 11.

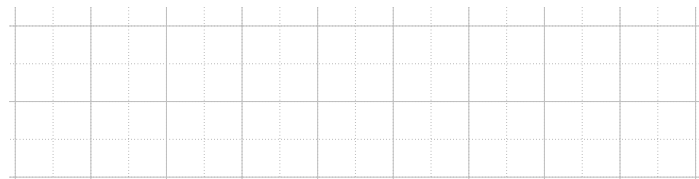
Figura A:



¿Se puede escribir el vector $\vec{c}$ en términos del vector $\vec{a}$?

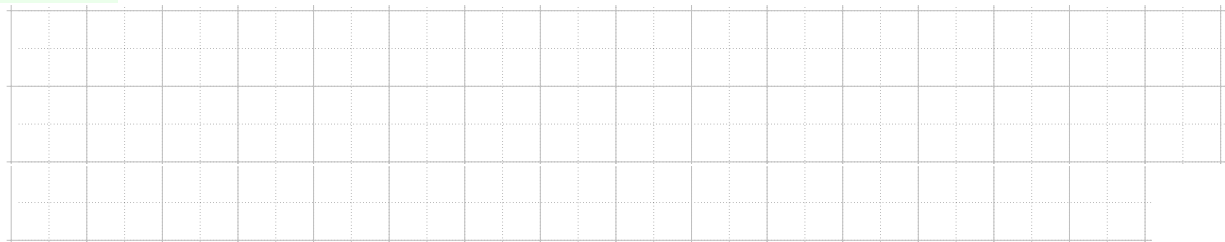


¿El vector $\vec{d}$ es combinación lineal de $\vec{a}$ & $\vec{b}$?



Definición 4.3. Los vectores $\vec{v}_1, \dots, \vec{v}_p$ serán **linealmente dependientes (l.d.)** cuando al menos uno de estos vectores se puede escribir como combinación lineal de los otros. En el caso opuesto, los vectores serán **linealmente independientes (l.i.)**.

Ejemplos:



Observaciones:

- Cuando varios vectores son l.d., pueden ocurrir dos casos: i) Cada vector (del espacio) es combinación lineal de ellos, ii) existen vectores que no son combinación lineal de ellos.
- Para los vectores en el plano (véase Fig. A) se cumple:
 - a) A lo más dos vectores en un plano son linealmente independientes.
 - b) Un vector siempre se puede poner en combinación lineal de dos vectores linealmente independientes.
- Para los vectores en el espacio (véase Fig. B) se cumple:
 - a) A lo más tres vectores en el espacio son linealmente independientes.
 - b) Un vector siempre se puede poner en combinación lineal de tres vectores linealmente independientes: $\vec{d} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{c}$.

La formulación de la definición 4.3 sugiere que verificar dependencia lineal es una tarea de combinatoria. El siguiente teorema propone “la forma” más usada de verificar la propiedad de ser linealmente (in)dependiente.

 **Teorema 4.2** (Independencia Lineal). *Los vectores $\vec{v}_1, \dots, \vec{v}_p$ son linealmente independientes (l.i.) si y solamente si la ecuación $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p = \vec{0}$ tiene una única solución, la trivial ($x_1 = x_2 = \dots = x_p = 0$).*

Demostración. “ $\Rightarrow$ ” Sean $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_p$ vectores l.i.



“ $\Leftarrow$ ” Por hipótesis tenemos que $x_1 = x_2 = \dots = x_p = 0$ es la única solución de la combinación lineal $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p = \vec{0}$.

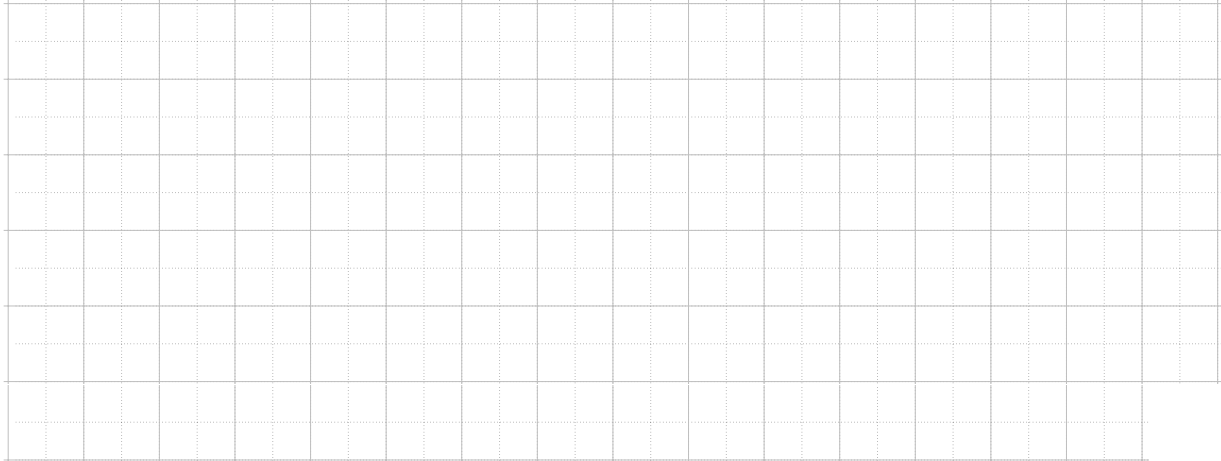
Supongamos que $\vec{v}_1, \dots, \vec{v}_p$ no son linealmente independientes, entonces uno de ellos se puede poner en combinación lineal de los otros. Sin pérdida de generalidad (w.l.o.g.) podemos considerar a $\vec{v}_1$ y ponerlo en combinación lineal de los otros, es decir, existen escalares $a_2, a_3, \dots, a_p$ tales que

$$\vec{v}_1 = a_2\vec{v}_2 + a_3\vec{v}_3 + \dots + a_p\vec{v}_p \quad \Longleftrightarrow \quad \vec{0} = -\vec{v}_1 + a_2\vec{v}_2 + a_3\vec{v}_3 + \dots + a_p\vec{v}_p.$$

La segunda igualdad da una segunda solución a la ecuación $x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_p\vec{v}_p = \vec{0}$ con $x_1 = -1 \neq 0$, por lo cual, contradice que la única solución es $x_1 = x_2 = \dots = x_p = 0$. Se concluye que los vectores son linealmente independientes. $\square$

Ejemplo 1: Decida si los vectores $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ & $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ son linealmente independientes.

Solución: Por el teorema 4.2 es suficiente ver si la solución del sistema $x_1\vec{v}_1 + x_2\vec{v}_2 = \vec{0}$ es única.



Ejemplo 2: Decida si los vectores $\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$, $\begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix}$ & $\begin{pmatrix} 2 \\ -4 \\ 7 \end{pmatrix}$ son linealmente independientes.

Solución: Se ocupa el teorema 4.2 (Independencia Lineal) para mostrar si los vectores son linealmente independiente ó dependiente. Sean $x_1, x_2, x_3 \in \mathbb{R}$ tales que:

$$x_1 \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} + x_2 \begin{pmatrix} 4 \\ 0 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 2 \\ -4 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$





IMPORTANTE : El teorema 4.2 interpretado para vectores $\{\vec{v}_1, \dots, \vec{v}_p\} \subset \mathbb{R}^m$: Los vectores $\{\vec{v}_1, \dots, \vec{v}_p\}$ son linealmente independientes si y solo si el sistema $A\vec{x} = \vec{0}$ con la matriz cuyas columnas son los vectores $\vec{v}_1, \dots, \vec{v}_p$ tiene una única solución (la trivial), es decir el siguiente sistema debe tener únicamente la solución cero:

$$\begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_p \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \vec{0}$$

Es decir, el rango de A debe ser $r = p$, lo cual solo puede suceder cuando $p \leq m$.

Ejemplo 3: Determine si el conjunto de polinomios $\{1 - x, 2 + x + 2x^2, 1 + x^2\}$ es linealmente independiente.

Solución: Por el teorema 4.2 es suficiente ver si $\alpha = \beta = \gamma = 0$ es la única solución de

$$\alpha(1 - x) + \beta(2 + x + 2x^2) + \gamma(1 + x^2) = \vec{0}$$

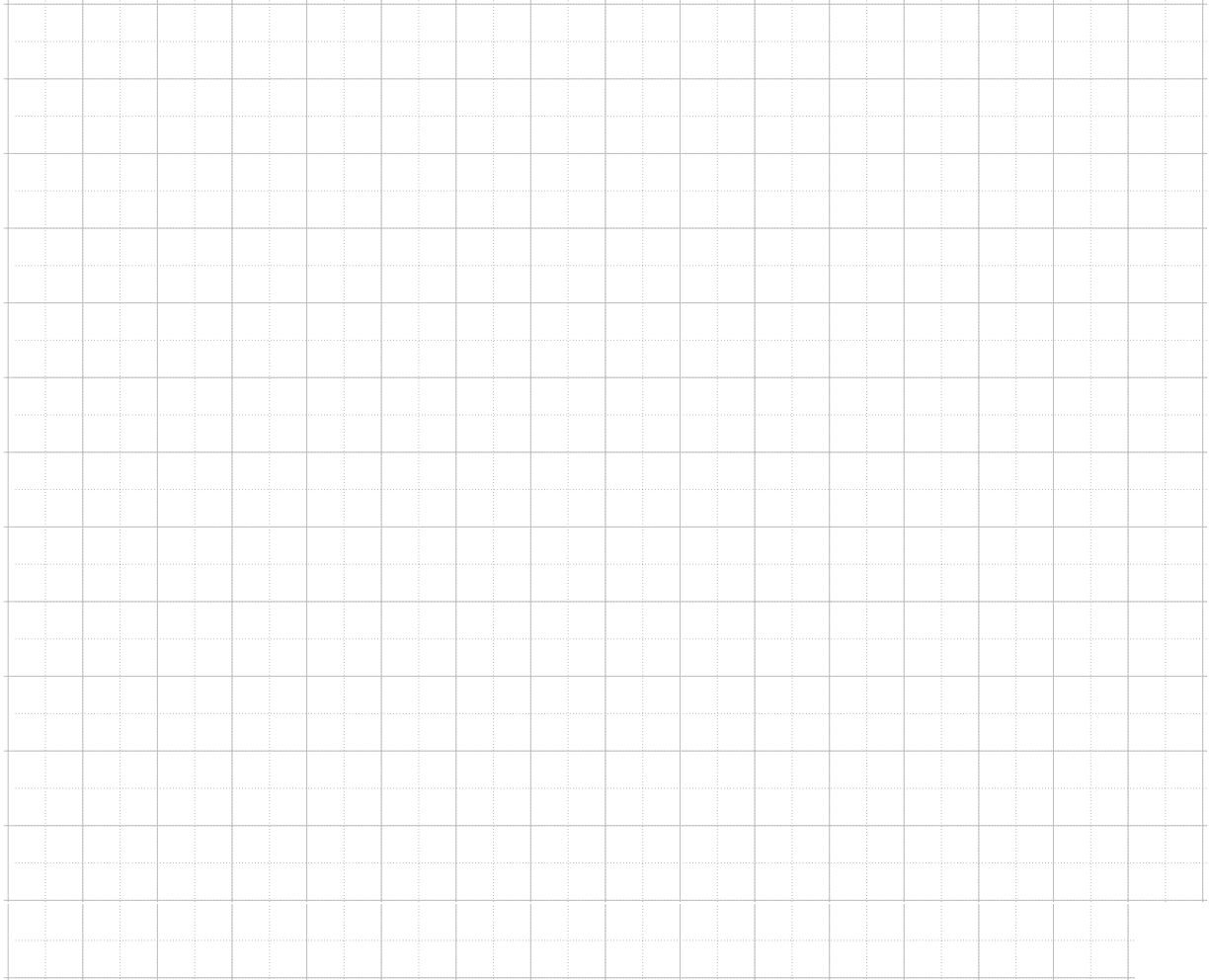


4. Combinaciones lineales

Ejemplo 4: Determine si las siguientes matrices son linealmente independientes:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Solución: Por el teorema 4.2 de la independencia lineal, tenemos que mostrar que $\alpha A + \beta B + \gamma C = 0$ implica $\alpha = \beta = \gamma = 0$.



Ejemplo 5: Determine si el conjunto de funciones $\{\sin x, \cos x\}$ es linealmente independiente (en el espacio vectorial de las funciones reales continuas en $[0, 2\pi)$).

Sol.: Sean $\vec{u} = \sin x$ & $\vec{v} = \cos x$. Entonces, por el teorema 4.2 de la independencia lineal, tenemos que $\alpha\vec{u} + \beta\vec{v} = \vec{0}$ debe implicar $\alpha = \beta = 0$. Dado que $\alpha\vec{u} + \beta\vec{v} = \vec{0} = 0(x)$ equivale a

$$\alpha(\sin x) + \beta(\cos x) = 0 \quad \text{para todo} \quad x \in [0, 2\pi),$$

podemos tomar puntos particulares en el intervalo $[0, 2\pi)$.

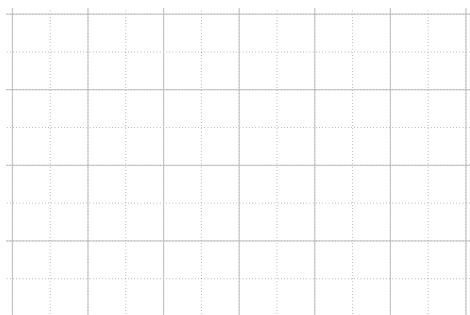
Tomando $x = 0$, obtenemos $0 = \alpha(\sin 0) + \beta(\cos 0) = \alpha \cdot 0 + \beta \cdot 1 = \beta$ y tomando $x = \pi/2$, obtenemos $0 = \alpha(\sin \pi/2) + \beta(\cos \pi/2) = \alpha \cdot 1 + \beta \cdot 0 = \alpha$. Concluimos que $\alpha = \beta = 0$ y por el teorema 4.2 sigue que las funciones son linealmente independientes.

4.3. Bases y dimensión

Comenzaremos esta sección con el siguiente ejemplo:

Un cabezal de un *plotter* S se mueve a través de una barra hacia arriba y hacia abajo, también hacía un lado y hacia otro, como se visualiza en la figura. Conteste las siguientes preguntas:

Figura:



a) ¿Con estos movimientos se llegan a todos los puntos del plano, *i.e.* se cubre todo el plano?



b) ¿Habrá una ventaja cuando se permiten movimientos adicionales en otro ángulo (distinto a 90°) ?



c) ¿Cuántas direcciones de movimientos son necesarias en un “aparato” con el fin de que se cubran todos los puntos del espacio?



Las preguntas y respuestas anteriores motivan la siguiente definición de bases.

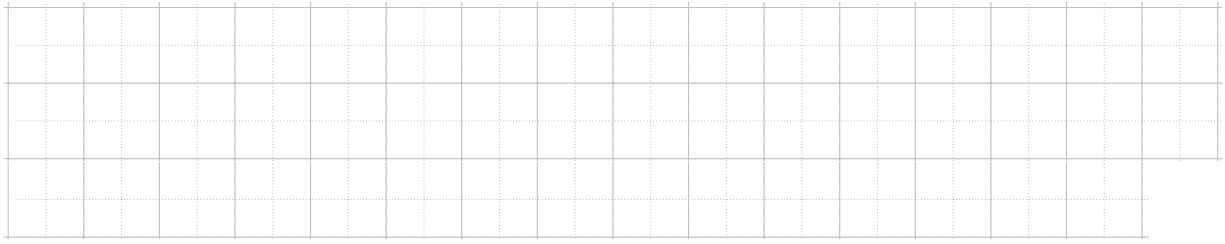


Definición 4.4. Sea $(\mathcal{V}, +, \cdot)$ un espacio vectorial. Un conjunto de vectores $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\} \subset \mathcal{V}$ es una **Base** de $\mathcal{V}$ cuando se cumplen las dos condiciones:

- los vectores $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ son linealmente independientes
- y $\mathcal{V} = \text{span}\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$, es decir, cada vector $\vec{w} \in \mathcal{V}$ es combinación lineal de estos vectores.

Ejemplo 1: Compruebe que los vectores $\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ & $\vec{b}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ forman una base de $\mathbb{R}^2$.

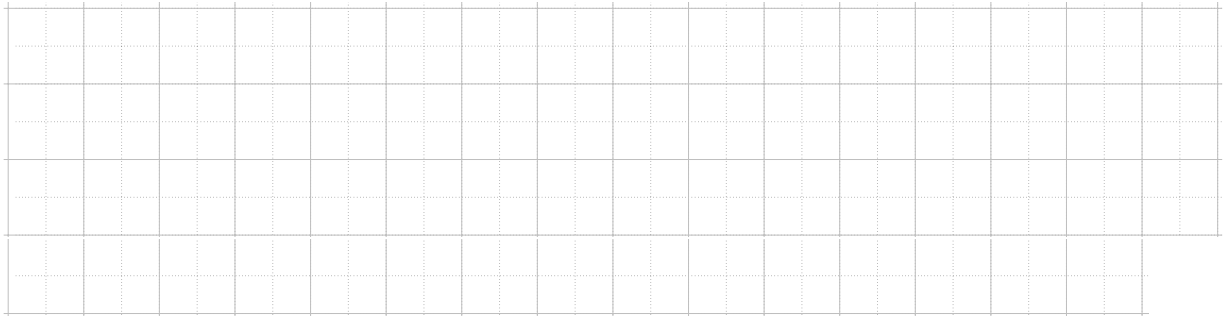




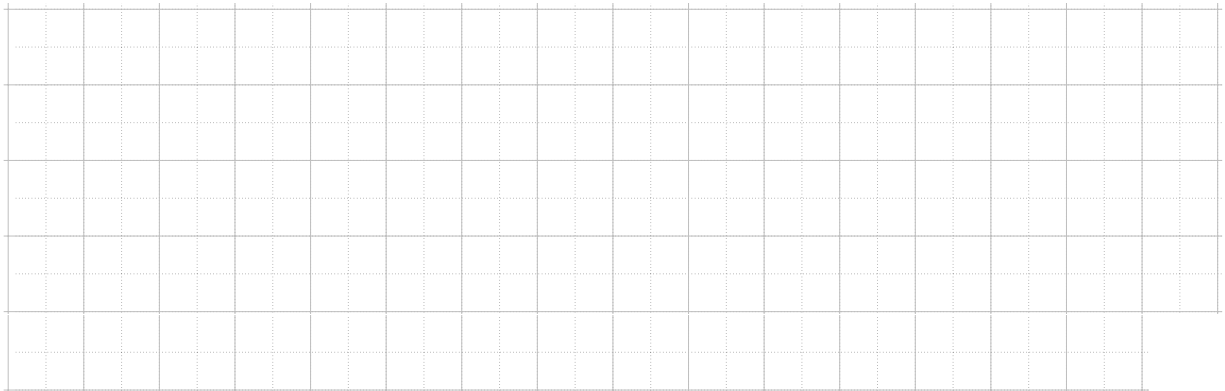
Ejercicio: “Las bases no son únicas.” Pruebe que los siguientes conjuntos son bases de $\mathbb{R}^2$:

$$\mathcal{B}_1 = \left\{ \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad \mathcal{B}_2 = \left\{ \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 1.5 \\ 2 \end{bmatrix} \right\} \quad \& \quad \mathcal{B}_3 = \left\{ \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

Ejemplo 2: ¿Los vectores $\{\vec{e}_1, \vec{e}_2, \vec{u}\}$ con $\vec{u} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \in \mathbb{R}^2$ forman una base de $\mathbb{R}^2$?



Ejemplo 3: ¿Los vectores $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ & $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$ forman una base de $\mathbb{R}^3$?



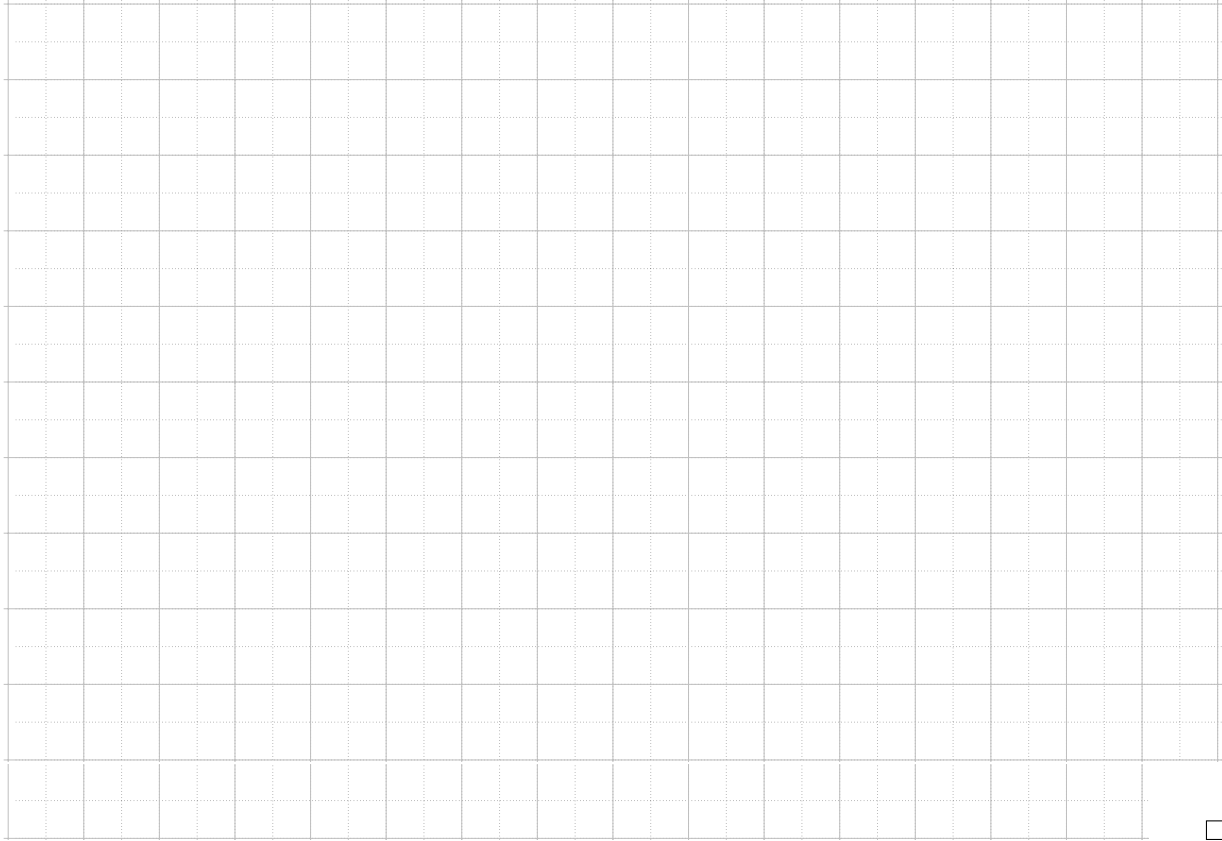
Interpretación: Una base de $\mathcal{V}$ es un sub-conjunto finito de tamaño mínimo que a la vez permite generar todos los vectores en $\mathcal{V}$.



Corolario 4.3. Los siguientes vectores (vectores canónicos) forman una base de $\mathbb{R}^n$:

$$\left\{ \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \vec{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \right\}.$$

Demostración. El razonamiento es similar al ejemplo 1.



□

Más adelante demostraremos que dos bases del mismo $\mathcal{V}$ tienen el mismo número de elementos. Por eso, la siguiente definición tiene sentido.

 | **Definición 4.5.** Al número de vectores de una base de $\mathcal{V}$ se le denomina **dimensión de $\mathcal{V}$** ($\dim \mathcal{V}$).

Ejemplos:

- La dimensión del plano $\mathbb{R}^2$ es dos, ya que $\mathcal{B} = \{\vec{e}_1, \vec{e}_2\}$ es una base.
- La dimensión del espacio $\mathbb{R}^3$ es tres, ya que $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es una base.
- La dimensión de $\mathbb{R}^n$ es n , ya que $\mathcal{B} = \{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ es una base.
- La dimensión de matrices $\mathbb{R}^{2 \times 2}$ es 4, ya que una base es

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

- La dimensión de los polinomios de grado a lo mas n es $n+1$, ya que el conjunto $\{1, x, x^2, \dots, x^n\}$ tiene $n+1$ elementos, es linealmente independiente y genera esos polinomios.

Esos ejemplos dejan realizar que encontrar una base y la dimensión de espacios vectoriales “completos” no es un gran desafío. Lo que interesa en teoría, aplicaciones y en optimización es la dimensión de sub-espacios.

4.3.1. Bases y dimensión de sub-espacios

Sean $(\mathcal{V}, +, \cdot)$ un espacio vectorial y $(S, +, \cdot)$ con $S \subseteq \mathcal{V}$ un sub-espacio vectorial.

Por la definición 3.2 la tupla $(S, +, \cdot)$ también es un espacio vectorial.



Conclusión: Por las definiciones 4.4 & 4.5 el conjunto S tiene una base y una dimensión.

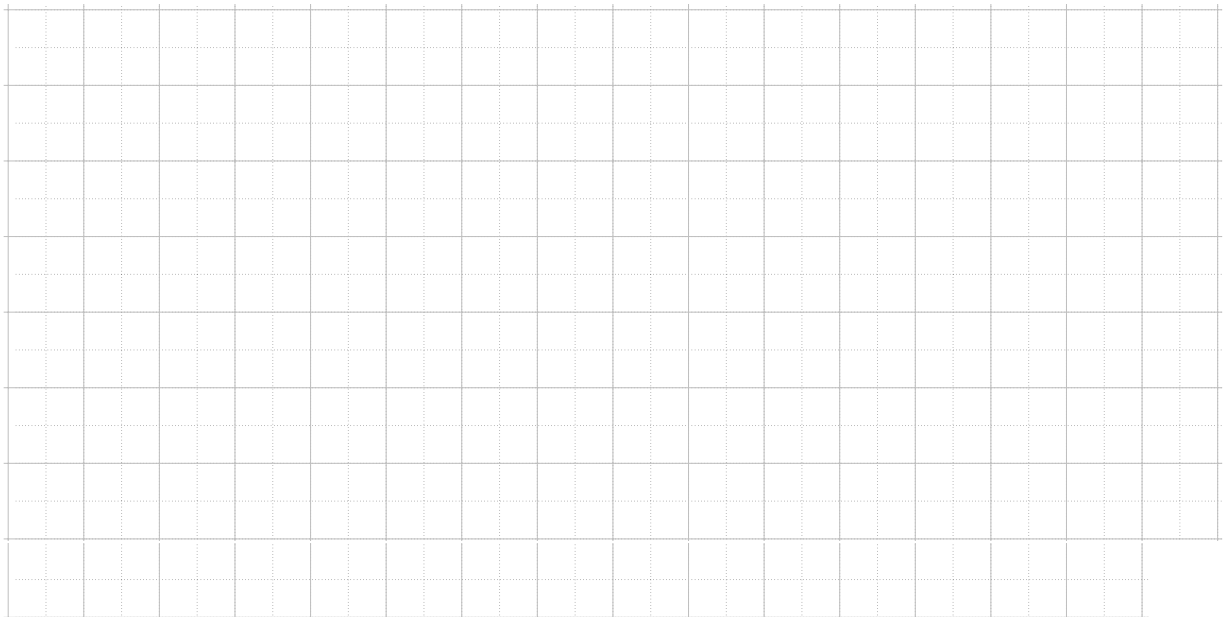
En general, cuando S es subconjunto estricto de $\mathcal{V}$ ($S \subsetneq \mathcal{V}$) tendremos $\dim S < \dim \mathcal{V}$.

Ejemplo 1: Los vectores $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ & $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^3$ forman una base de $S = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : x_3 = 0 \right\}$.



El lema 3.2 dice que el conjunto $\text{Null}(A)$, de las soluciones $\vec{x}$ de un sistema homogéneo $A\vec{x} = \vec{0}$ con una matriz $A \in \mathbb{R}^{m \times n}$, es un (sub-)espacio vectorial de $\mathbb{R}^n$. Por lo tanto, tiene una base y dimensión.

Ejemplo 2: Encuentre una base y la dimensión del conjunto de soluciones del siguiente sistema homogéneo: $x_1 + 3x_2 - 5x_3 = 0$.

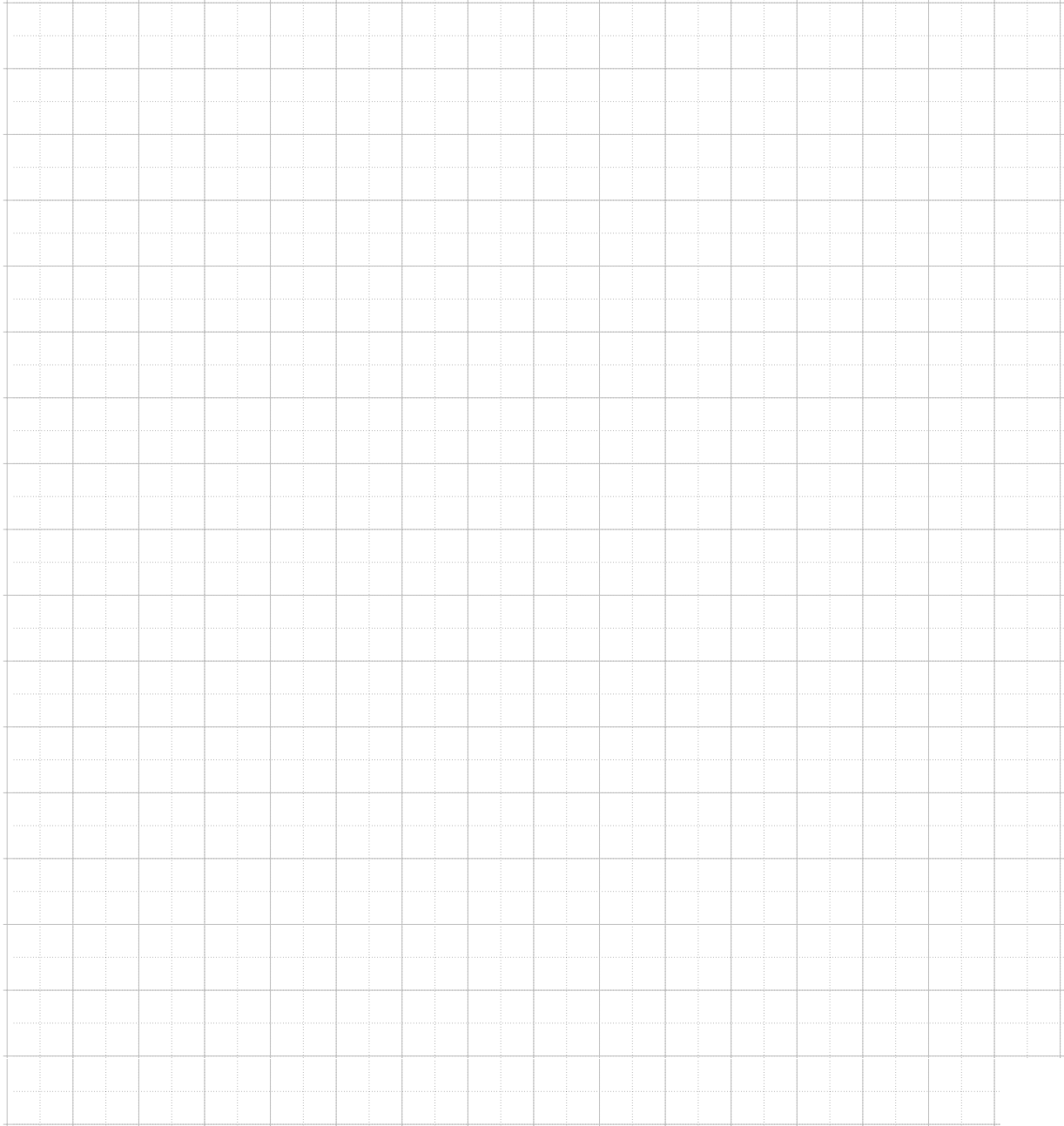


Ejemplo 3: (Determinar una base y dimensión de un espacio nulo)

Encuentre una base y la dimensión del conjunto de soluciones del siguiente sistema:

$$x_1 + 2x_2 + 7x_3 + 9x_4 - 6x_5 = 0$$

$$2x_1 + 5x_2 + 3x_3 - x_4 - 2x_5 = 0$$



Valioso: Sea r el rango de A de $m \times n$, entonces en los ejemplos anteriores se reconoce la siguiente igualdad:

$$\text{“número de variables libres”} = \dim(\text{Null}(A)) = n - r.$$

La segunda igualdad es conocida como el “Teorema del Rango”.

4. Combinaciones lineales

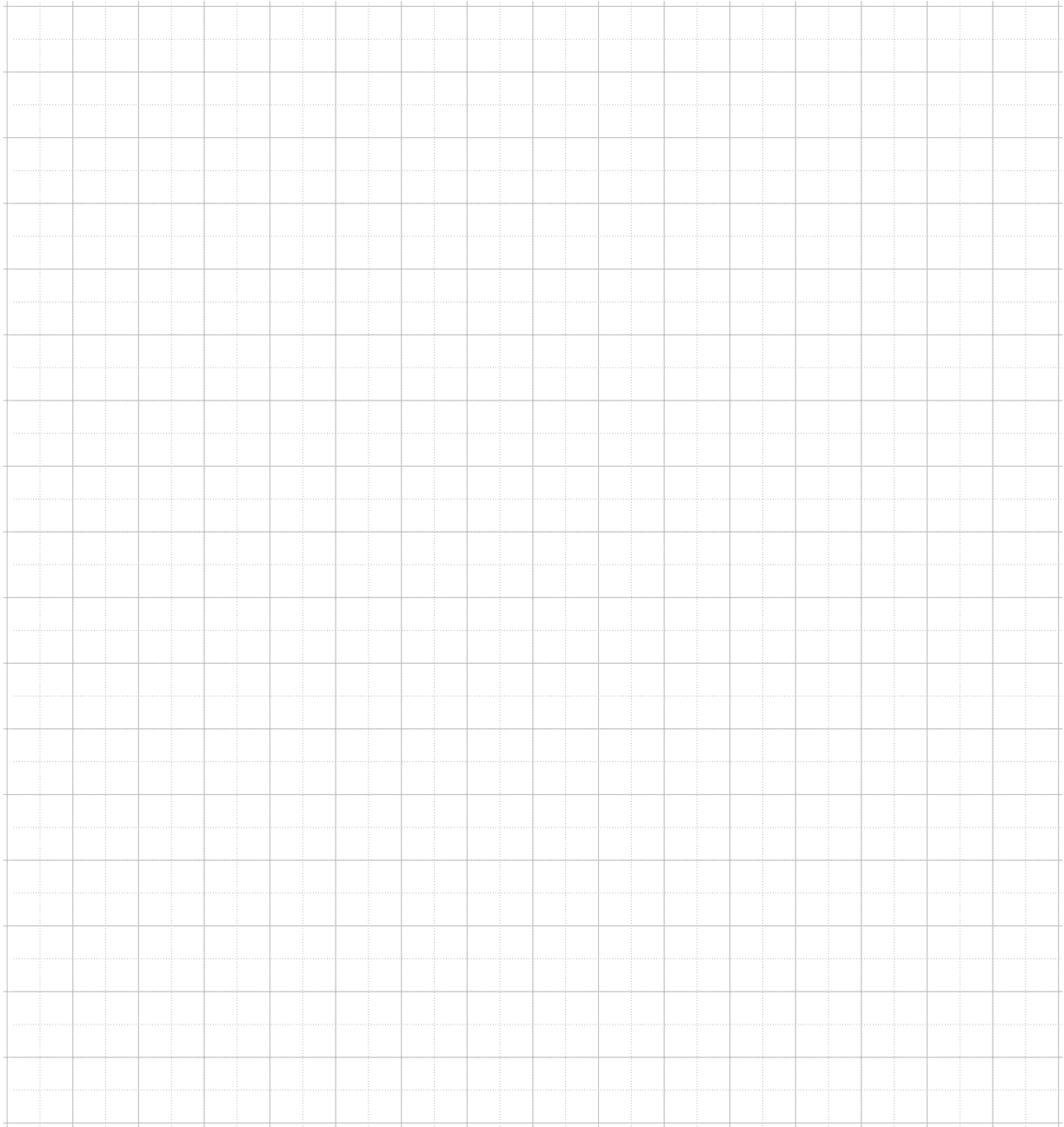


Valioso: Las columnas de una matriz $A \in \mathbb{R}^{m \times n}$ son linealmente independientes si y solo si el sistema $A\vec{x} = \vec{0}$ tiene solución única (ver teorema 4.2). Eso solo ocurre si $r = \text{rango}(A) = n$.

Conclusión: Las variables básicas indican columnas lin. indep., mientras cada variable libre identifica un columna lin. dependiente.

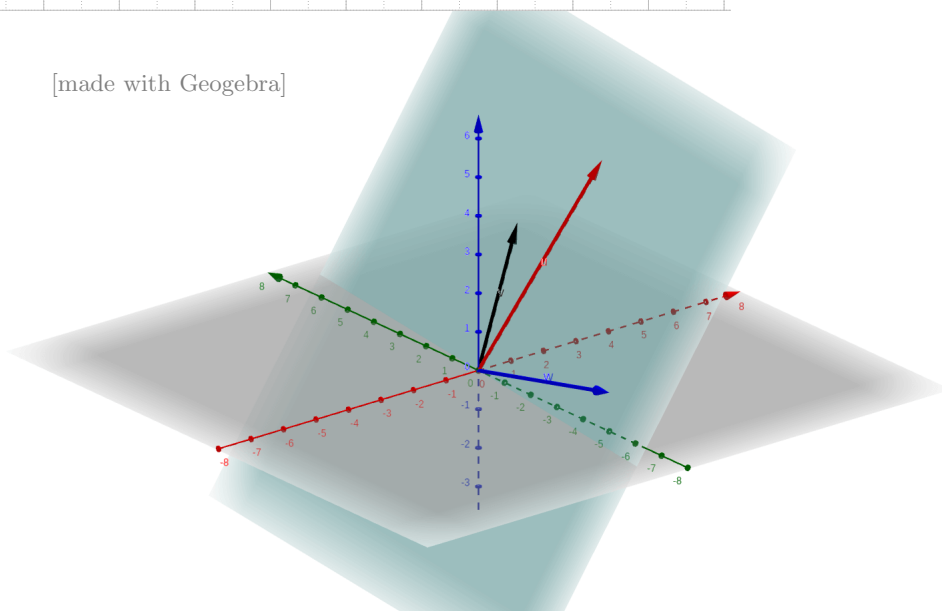
Ejemplo 4: Encuentre una base del conjunto $\text{span}\{\vec{u}, \vec{v}, \vec{w}\} = \text{Img}(A)$ donde

$$\vec{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \quad \vec{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} \quad \& \quad A = \begin{bmatrix} 3 & 2 & 0 \\ -1 & 1 & -5 \\ 5 & 3 & 1 \end{bmatrix}.$$





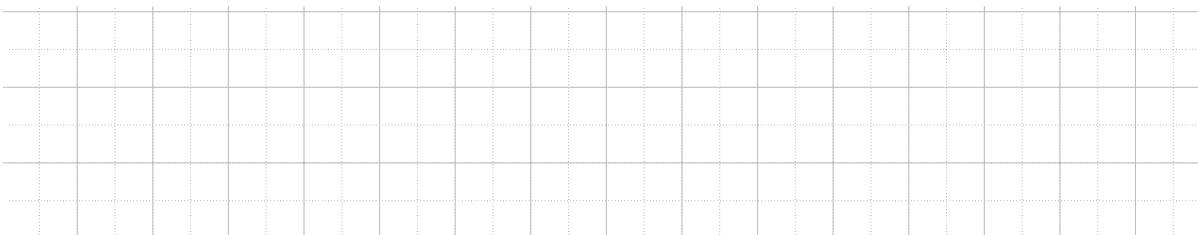
[made with Geogebra]



Geométricamente significa que los vectores $\vec{u}, \vec{v}, \vec{w}$ yacen todos en el mismo plano. Cada pareja de estos vectores podría servir como base del plano. Pero, los unos principales en la FER de A identifican $\vec{u}, \vec{v}$.



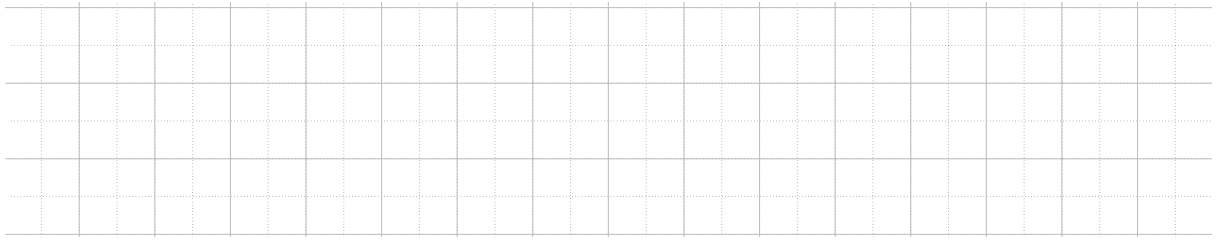
Definición (Espacio Columna). Sea A una matriz de $m \times n$. El sub-espacio $\text{Img}(A)$ de $\mathbb{R}^m$, generado por las columnas de A , también se denota por $\text{Col}(A)$:



4. Combinaciones lineales



Definición (Espacio Renglón). Sea A una matriz de $m \times n$. Al sub-espacio de $\mathbb{R}^n$, generado por los renglones de A , lo denotaremos por $\text{Renglón}(A)$:

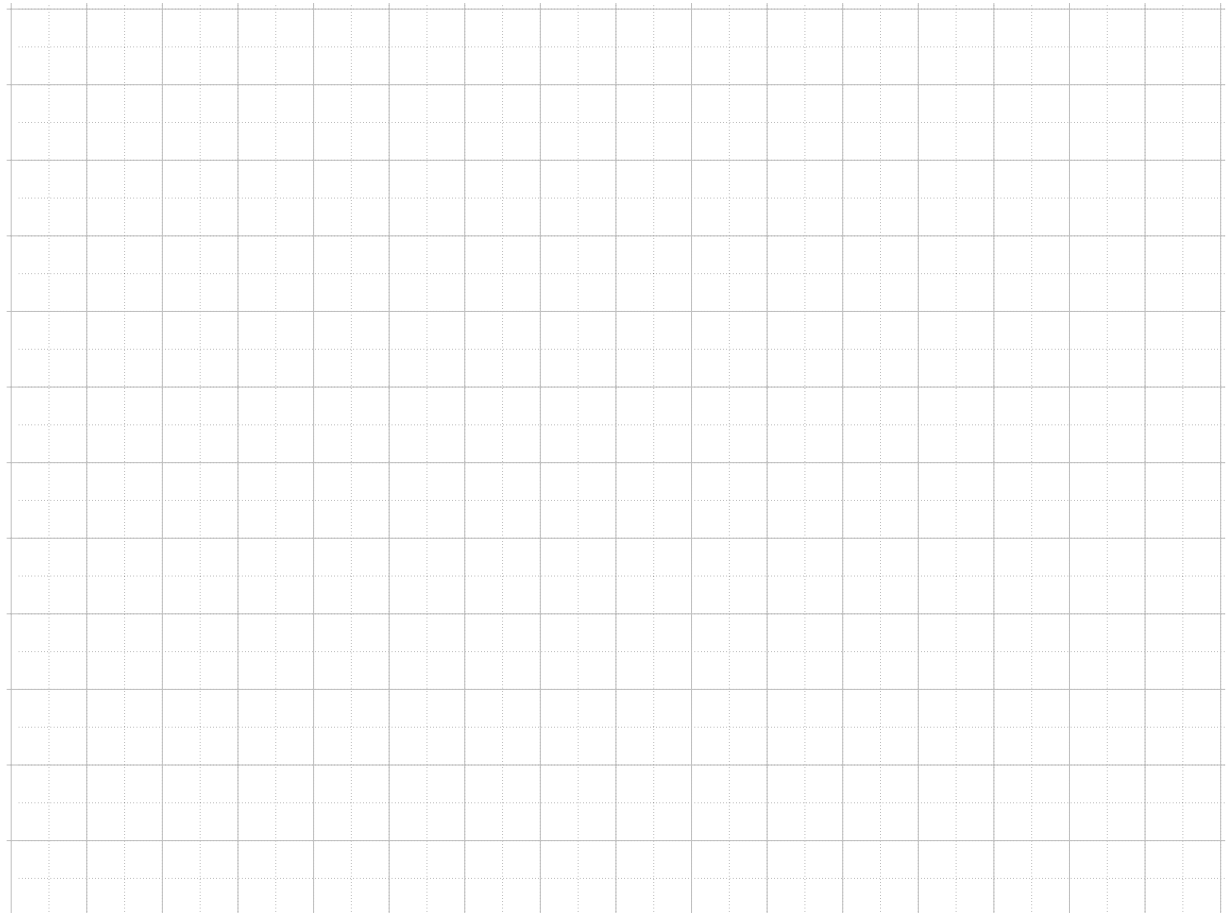


Las operaciones elementales son combinaciones lineales particulares de renglones y como son reversibles, tenemos $\text{Renglón}(\text{FER}(A)) = \text{Renglón}(A)$. Más aún, los renglones de la FER distintos de cero forman una base, ya que son linealmente independientes y generan los renglones originales.

Ejemplo 5: Encuentre una base para el espacio $\text{Renglón}(A)$ donde

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$$

Solución: Encontrar la FER y definir la base (y dimensión).





Resultado: Solo cuentan los renglones distintos de cero, es decir, la base y el espacio renglón de este ejemplo son:

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 4 \end{pmatrix} \right\} \quad \text{y} \quad \text{Renglón}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 4 \end{pmatrix} \right\} \subsetneq \mathbb{R}^5.$$

Por definición, el número de elementos de la base $\mathcal{B}$ es la dimensión del espacio. En este caso, la dimensión de $\text{Renglón}(A)$ es igual a $|\mathcal{B}| = 3$. Además, como en el ejemplo anterior podemos usar que los unos principales indican las columnas (originales) L.I, es decir, también hemos encontrado

$$\text{Col}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \\ 1 \end{pmatrix} \right\} \subsetneq \mathbb{R}^4. \quad (4.1)$$



Lema 4.4 (importante). Sea $A \in \mathbb{R}^{m \times n}$ una matriz con m renglones y n columnas y rango r , entonces el espacio $\text{Col}(A) \subseteq \mathbb{R}^m$, el espacio $\text{Renglón}(A) \subseteq \mathbb{R}^n$ y

$$\dim(\text{Renglón}(A)) = r = \dim(\text{Col}(A)).$$



4. Combinaciones lineales

Demostración. La FER de A tiene r unos principales y contiene combinaciones lineales de renglones de A . Entonces, tiene r renglones distintos de cero que son linealmente independientes. Dado que los renglones originales son combinaciones lineales de estos renglones, cada renglón de A que fue convertido en cero en la FER (de A) es combinación lineal (no-trivial) de los r renglones L.I. Por cada tal renglón, Lema 4.1(b) permite escribir el conjunto generado sin ese renglón, es decir, se reduce el número de renglones en el conjunto generador a los r vectores L.I. Por lo tanto, esos renglones forman una base de $\text{Renglón}(A)$ y $\dim(\text{Renglón}(A)) = r$.

Por otro lado, los r unos principales identifican r columnas linealmente independientes y usar 1 como valor de una variable libre, permite escribir la columna asociada como combinación lineal (no-trivial) de las r columnas L.I. Entonces, Lema 4.1(b) permite escribir el conjunto generado sin esas columnas y solo quedan las r columnas L.I. en el conjunto generador. Por lo tanto, esas columnas forman una base de $\text{Col}(A)$ y $\dim(\text{Col}(A)) = r$. $\square$



Los dos ejemplos anteriores (4,5) mostraron como encontrar bases de $\text{Col}(A)$ y de $\text{Renglón}(A)$.

A veces el siguiente truco puede reducir el número de operaciones elementales.

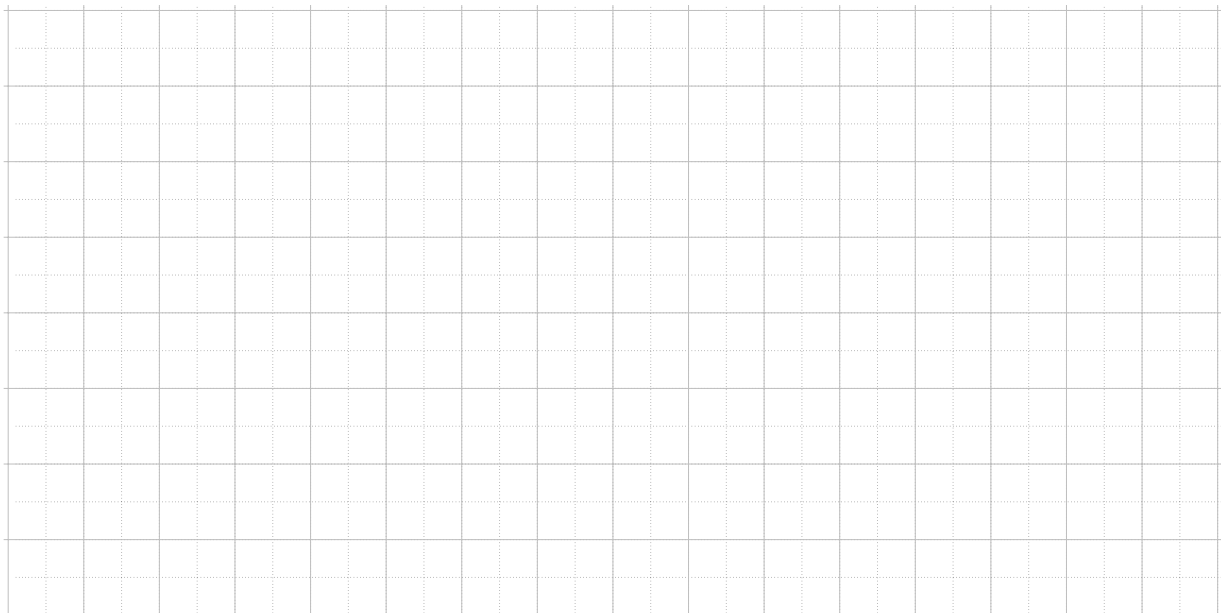
Las columnas de A son los renglones de A^T , por lo cual

$$\text{Col}(A) = \text{Renglón}(A^T) \quad \text{y} \quad \text{Renglón}(A) = \text{Col}(A^T).$$

Ejemplo 6: Encuentre otra base para el espacio $\text{Col}(A)$ usando $\text{Renglón}(A^T) = \text{Col}(A)$.

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$$

Solución: Encontrar la FER de A^T y definir la base (y dimensión).





Resultado: Hemos encontrado otra base que en (4.1) con la misma dimensión 3 :

$$\text{Col}(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} \right\} \subsetneq \mathbb{R}^4. \quad (4.2)$$

Más tarde usaremos Lema 4.1(b) para ver que los conjuntos en (4.1) y (4.2) son iguales.

4.3.2. Resultados generales e importantes



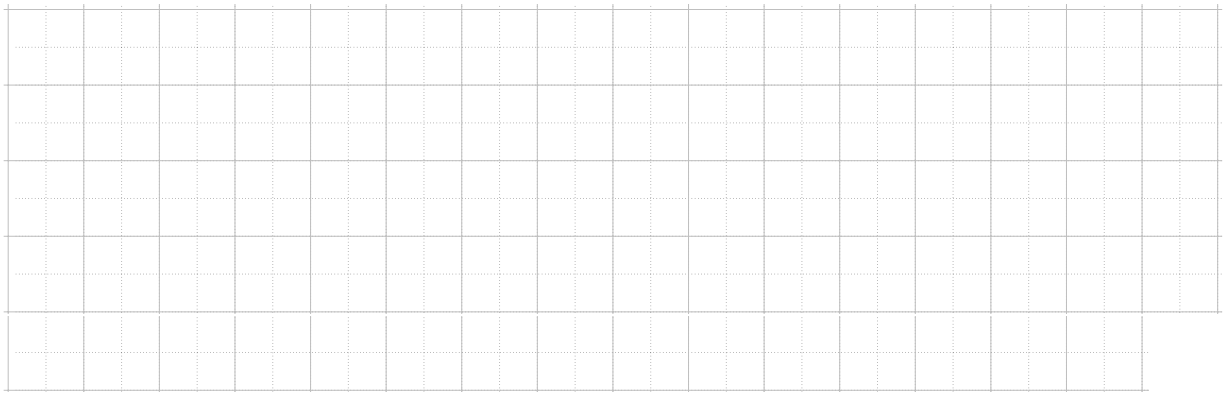
Lema 4.5 (Sistemas homogéneos). Si $[A|\vec{0}]$ es un sistema homogéneo de m ecuaciones con n variables, donde $m < n$, entonces el sistema tiene soluciones no-triviales.

Demostración. Dado que es un sistema homogéneo, este tiene al menos la solución trivial $\vec{0}$ y por lo tanto es consistente. Más aún, sabemos que $\text{rango}(A) = r \leq \min\{m, n\} = m$ ya que hay a lo más tantos unos principales como ecuaciones o variables (lo que sea menor). Sabemos que hay $n - r$ variables libres y con la hipótesis concluimos

$$\text{"# de variables libres"} = n - r \geq n - m > 0,$$

es decir, el sistema tiene por lo menos una variable libre, la cual, si es distinta de cero, da soluciones no-triviales. $\square$

Ejemplos: Matrices (aumentadas) con variables libres (en páginas 102, 104, 105 y 109):



Los siguientes resultados son “teoremas” por su importancia, no por la complejidad de la prueba.

Ejemplos: El siguiente teorema dice que conjuntos con más elementos que una base son linealmente dependientes.





Teorema 4.6. Sea $(\mathcal{V}, +, \cdot)$ un (sub-)espacio vectorial con una base $\{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathcal{V}$ de $\mathcal{V}$. Si $\vec{u}_1, \dots, \vec{u}_m \in \mathcal{V}$ con $m > n$, entonces $\{\vec{u}_1, \dots, \vec{u}_m\}$ es un conjunto lin. dependiente.

Demostración. Tenemos que ver si la siguiente ecuación tiene una solución no-trivial:

$$\alpha_1 \vec{u}_1 + \alpha_2 \vec{u}_2 + \dots + \alpha_m \vec{u}_m = \vec{0}. \quad (\star)$$

Por hipótesis $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ es una base de $\mathcal{V}$. Entonces cada elemento de $\mathcal{V}$ se puede escribir en combinación lineal de los elementos de $\mathcal{B}$. Como $\vec{u}_1, \dots, \vec{u}_m \in \mathcal{V}$ podemos escribir sus elementos como sigue:

$$\left. \begin{aligned} \vec{u}_1 &= a_{11}\vec{v}_1 + a_{12}\vec{v}_2 + \dots + a_{1n}\vec{v}_n \\ \vec{u}_2 &= a_{21}\vec{v}_1 + a_{22}\vec{v}_2 + \dots + a_{2n}\vec{v}_n \\ &\vdots \\ \vec{u}_m &= a_{m1}\vec{v}_1 + a_{m2}\vec{v}_2 + \dots + a_{mn}\vec{v}_n \end{aligned} \right\} \quad (\star\star)$$

Al sustituir las ecuaciones del sistema $(\star\star)$ en la ecuación $(\star)$ obtenemos

$$\begin{aligned} \alpha_1(a_{11}\vec{v}_1 + a_{12}\vec{v}_2 + \dots + a_{1n}\vec{v}_n) + \alpha_2(a_{21}\vec{v}_1 + a_{22}\vec{v}_2 + \dots + a_{2n}\vec{v}_n) \\ + \dots + \alpha_m(a_{m1}\vec{v}_1 + a_{m2}\vec{v}_2 + \dots + a_{mn}\vec{v}_n) = \vec{0}. \end{aligned}$$

La definición 3.1 nos permite expandir los paréntesis (2b), multiplicar primero los escalares (2d), cambiar la posición de los vectores sumados (1d), evaluar primero las sumas asociados al vector $\vec{v}_i$ (1c) y factorizar los vectores $\vec{v}_i$ (2c). Como resultado la igualdad anterior da:

$$\begin{aligned} (\alpha_1 a_{11} + \alpha_2 a_{21} + \dots + \alpha_m a_{m1})\vec{v}_1 + (\alpha_1 a_{12} + \alpha_2 a_{22} + \dots + \alpha_m a_{m2})\vec{v}_2 \\ + \dots + (\alpha_1 a_{1n} + \alpha_2 a_{2n} + \dots + \alpha_m a_{mn})\vec{v}_n = \vec{0}. \end{aligned}$$

Dado que $\mathcal{B}$ es base, los vectores $\vec{v}_i$ son lin. independientes, entonces los coeficientes de la combinación lineal anterior deben ser cero. De este modo tenemos el siguiente sistema:

$$\left. \begin{aligned} \alpha_1 a_{11} + \alpha_2 a_{21} + \dots + \alpha_m a_{m1} &= 0 \\ \alpha_1 a_{12} + \alpha_2 a_{22} + \dots + \alpha_m a_{m2} &= 0 \\ &\vdots \\ \alpha_1 a_{1n} + \alpha_2 a_{2n} + \dots + \alpha_m a_{mn} &= 0 \end{aligned} \right\}$$

Este sistema lineal homogéneo tiene n ecuaciones y $m > n$ variables. En estas condiciones, el lema 4.5 dice que existe una solución no-trivial para los $\alpha_1, \dots, \alpha_m$, es decir, para la ecuación $(\star)$ existe una solución tal que $\alpha_i \neq 0$ para alguna $i \in \{1, \dots, m\}$. Entonces, por definición el conjunto de los vectores en $\{\vec{u}_1, \dots, \vec{u}_m\}$ es linealmente dependiente, lo cual termina la prueba. $\square$



Teorema 4.7 (de la Base). Sea $(\mathcal{V}, +, \cdot)$ un (sub-)espacio vectorial.

Si $\{\vec{u}_1, \dots, \vec{u}_m\}, \{\vec{v}_1, \dots, \vec{v}_n\} \subset \mathcal{V}$ son bases de $\mathcal{V}$, entonces $m = n$.

“Cualesquiera dos bases de un (sub-)espacio vectorial tienen el mismo número de vectores.”

Demostración. Sean $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ & $\mathcal{C} = \{\vec{u}_1, \dots, \vec{u}_m\}$ dos bases de $\mathcal{V}$.

Necesitamos probar que $m = n$, lo haremos en forma indirecta suponiendo $m > n$ ó $m < n$.

Caso 1 ($m > n$) : Dado que $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$ es base y $m > n$ el teorema 4.6 dice que cada conjunto con $m > n$ elementos es linealmente dependiente. Eso contradice que $\mathcal{C}$ es base.

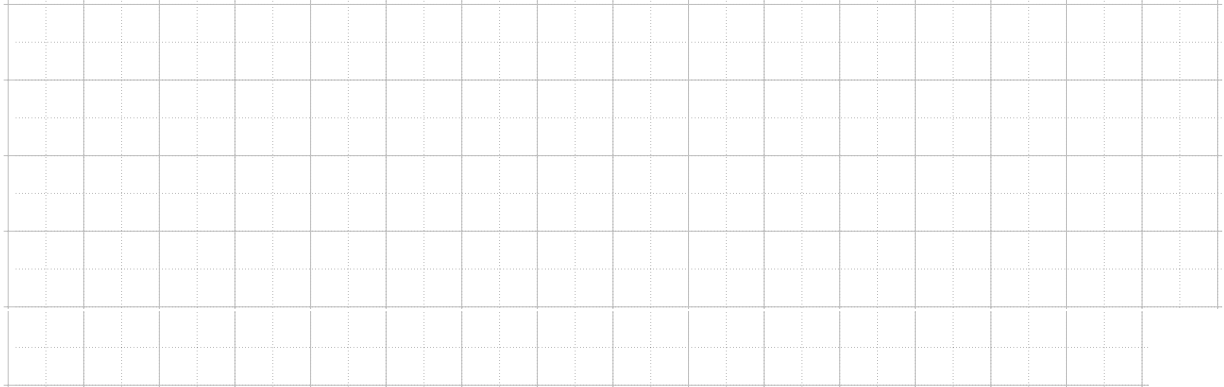
Caso 2 ($m < n$) : Dado que $\mathcal{C} = \{\vec{u}_1, \dots, \vec{u}_m\}$ es base y $n > m$ el teorema 4.6 dice que cada conjunto con $n > m$ elementos es linealmente dependiente. Eso contradice que $\mathcal{B}$ es base.

Negando los casos tenemos $m \leq n$ & $m \geq n$ lo cual da $m = n$ y termina la prueba. $\square$



Corolario 4.8. Sea $(\mathcal{V}, +, \cdot)$ un (sub-)espacio vectorial con $\mathcal{V}$ de dimensión n . Entonces, cada conjunto linealmente independiente de n vectores $\{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\} \subset \mathcal{V}$ es una base de $\mathcal{V}$.

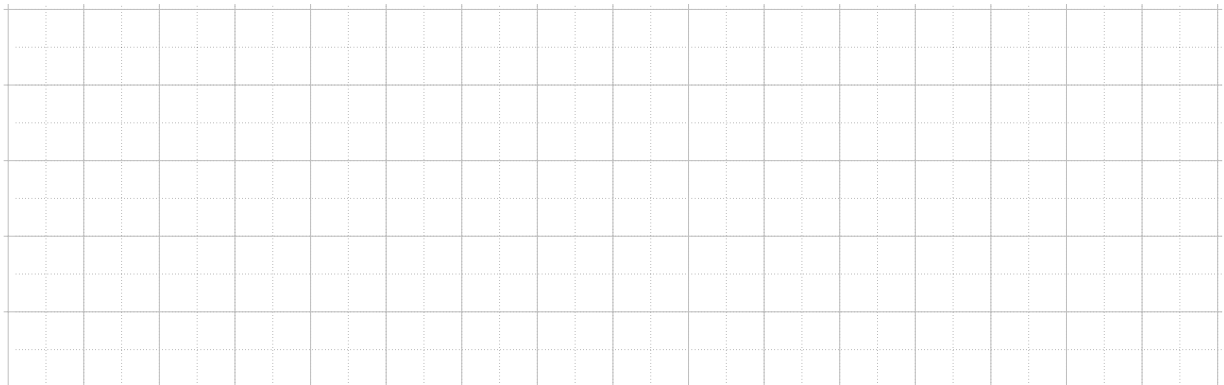
Ejemplo: Matrices distintas cuyas columnas son bases para $\mathbb{R}^4$:

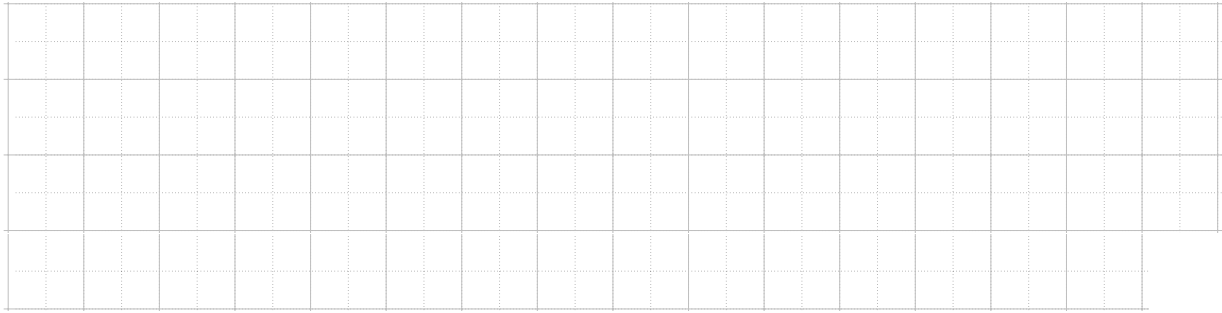


Ejemplo: Los resultados (4.1) y (4.2) muestran dos bases para $\text{Col}(A)$.

Usamos Lema 4.1(b) para mostrar igualdad entre los generados:

$$\text{span} \left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \\ 1 \end{pmatrix} \right\} \stackrel{(4.1)}{=} \text{Col}(A) \stackrel{(4.2)}{=} \text{span} \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 6 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 3 \end{pmatrix} \right\} \subsetneq \mathbb{R}^4.$$





Lema 4.9. Sea $(\mathcal{V}, +, \cdot)$ un (sub-)espacio vectorial y sea $\{\vec{v}_1, \dots, \vec{v}_p\} \subset \mathcal{V}$ un conjunto de vectores linealmente independiente. Si $\vec{w} \in \mathcal{V}$, entonces las siguientes condiciones son equivalentes:

- i) $\{\vec{w}, \vec{v}_1, \dots, \vec{v}_p\}$ es linealmente independiente.
- ii) $\vec{w}$ no pertenece a $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$.

Demostración. Ejercicio. Ayuda: Usar definiciones y el teorema 4.2. □

Interpretación: El lema 4.9 nos dice que si $\vec{w} \in \mathcal{V}$ no se puede poner en combinación lineal de un conjunto $\{\vec{v}_1, \dots, \vec{v}_p\}$ lin. independiente, entonces $\{\vec{w}, \vec{v}_1, \dots, \vec{v}_p\}$ es L.I. Por lo tanto, el teorema permite construir conjuntos L.I. más grandes. El siguiente teorema dice que este proceso termina cuando existe una base finita.

Ejemplo: La base canónica (y algo más) en $\mathbb{R}^4$:





Teorema 4.10. Sea $(\mathcal{V}, +, \cdot)$ un (sub-)espacio vectorial de dimensión n . Si $p < n$ y el conjunto $\{\vec{v}_1, \dots, \vec{v}_p\} \subset \mathcal{V}$ es lin. independiente, entonces se pueden encontrar $n - p$ vectores $\vec{v}_{p+1}, \vec{v}_{p+2}, \dots, \vec{v}_n \in \mathcal{V}$ tales que $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ es una base de $\mathcal{V}$.

Interpretación: Cualquier conjunto lin. indep. se puede extender a una base.

Demostración. Sea n la dimensión de $\mathcal{V}$ y $\{\vec{v}_1, \dots, \vec{v}_p\}$ un conjunto lin. independiente.

Si $p = n$, entonces por el corolario 4.8 el conjunto $\{\vec{v}_1, \dots, \vec{v}_p\}$ ya es base.

Si $p < n$, entonces $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} \subsetneq \mathcal{V}$ ya que la dimensión de $\mathcal{V}$ es n , lo cual (por definición 4.5) dice que una base tiene n elementos. Por lo mismo, existe un elemento $\vec{v}_{p+1} \in \mathcal{V}$ que no pertenece al espacio $\text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$. Es decir $\vec{v}_{p+1} \notin \text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$ y por la equivalencia en lema 4.9 el conjunto $\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}\}$ también es L.I. Si $p+1 < n$, entonces podemos repetir este mismo razonamiento y obtener un elemento $\vec{v}_{p+2} \notin \text{span}\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}\}$ aplicar la equivalencia del lema 4.9 para obtener que $\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}, \vec{v}_{p+2}\}$ es L.I. Esto se continua hasta obtener un conjunto de $p + (n - p)$ elementos L.I., es decir hasta obtener que $\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}, \dots, \vec{v}_n\}$ es L.I. y por lo tanto es base por el corolario 4.8. $\square$



Nota: Si tenemos un sub-espacio vectorial $(\mathcal{S}, +, \cdot)$ de un espacio vectorial $(\mathcal{V}, +, \cdot)$ y una base $\{\vec{v}_1, \dots, \vec{v}_p\}$ de $\mathcal{S}$, entonces el teorema 4.10 permite completar esta base a una base de $\mathcal{V}$, digamos $\{\vec{v}_1, \dots, \vec{v}_p, \vec{v}_{p+1}, \dots, \vec{v}_n\}$ donde $n > p$ es la dimensión de $\mathcal{V}$.

Ejemplo: Si unimos las bases de $\text{Renglón}(A)$ & $\text{Null}(A)$ tenemos una base de $\mathbb{R}^n$:



4.4. Resumen y un “mind map”

De los recientes dos temas, los siguientes hechos son de gran utilidad para seguir entendiendo:
Hacemos primero un dibujo para ubicar espacios asociados a matrices que hemos introducido.



Consecuencias :

- “La negación” del teorema 4.2:

Los vectores $\{\vec{v}_1, \dots, \vec{v}_n\}$ son linealmente dependientes si y solo si la siguiente ecuación tiene soluciones no-triviales: $\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_n \vec{v}_n = \vec{0}$.

(Una solución no-trivial es distinta de la solución cero.)

- Por el teorema 4.2 tenemos: Columnas LI. $\iff [A|\vec{0}] = [\vec{v}_1, \dots, \vec{v}_n|\vec{0}]$ tiene solución única $\iff FER(A)$ tiene n unos principales $\iff$ rango de A es n . Eso requiere $m \geq n$ ya que $n = r$ y $r \leq \min\{m, n\}$, es decir, “el sistema no tiene más variables que ecuaciones.”
- Un sistema no homogéneo (con lado derecho $\vec{b} \neq \vec{0}$) siempre tiene solución si existen tantos unos principales como ecuaciones, es decir, cuando $m = r \leq \min\{m, n\}$ lo cual permite variables libres y a la vez excluye renglones lin. dependientes.

Espacios de matrices:

Si $A \in \mathbb{R}^{m \times n}$ es una matriz de m renglones y n columnas, entonces

- Su rango r es el numero de unos principales en la FER de A
- La imagen de A son los puntos en $\mathbb{R}^m$ que se alcanzan como combinaciones lineales de columnas de A y se escribe de las tres siguientes formas:

$$\text{Img}(A) = \text{Col}(A) = \text{span} \{A_{:1}, A_{:2}, \dots, A_{:n}\} \subseteq \mathbb{R}^m.$$

- Sean $A \in \mathbb{R}^{m \times n}$ y $B = \text{FER}(A) \in \mathbb{R}^{m \times n}$.

Los r renglones de B distintas de cero forman una base de $\text{Renglón}(A) = \text{Renglón}(B) \subseteq \mathbb{R}^n$.

Los r unos principales en B identifican r columnas lin. independientes de A que forman una base de $\text{Img}(A) = \text{Col}(A)$ con $r \leq n$ columnas, donde Lema 4.1(b) confirma la igualdad entre los conjuntos generados “span”.

- Las dimensiones de su espacio renglón y columna son iguales al rango, *i.e.*

$$\dim(\text{Img}(A)) = \dim(\text{Col}(A)) = r = \dim(\text{Renglón}(A)).$$

Interpretaciones: “# columnas lin. indep.”, “# unos principales”, “# renglones lin. indep.”

- El espacio nulo $\text{Null}(A) \subseteq \mathbb{R}^n$ es el conjunto de todas las soluciones del sistema $A\vec{x} = \vec{0}$. Su dimensión es igual al “número de variables libres”, *i.e.*

$$\dim(\text{Null}(A)) = n - r.$$

Interpretaciones alternativas: “# direcciones lin. indep. cuyo producto punto con renglones de A es cero”, “# columnas lin. dep. de A ” y “# columnas de la $\text{FER}(A)$ que NO contienen un uno principal”.

La formula de la dimensión:

$$n = \dim(\text{Img}(A)) + \dim(\text{Null}(A)) \quad (4.3)$$

Formas alternativas: Interpretaciones de dimensiones se encuentran arriba.

$$\begin{aligned} n &= \dim(\text{Col}(A)) + \dim(\text{Null}(A)) \\ &= (r) + (n - r) \\ &= \dim(\text{Renglón}(A)) + \dim(\text{Null}(A)). \end{aligned}$$

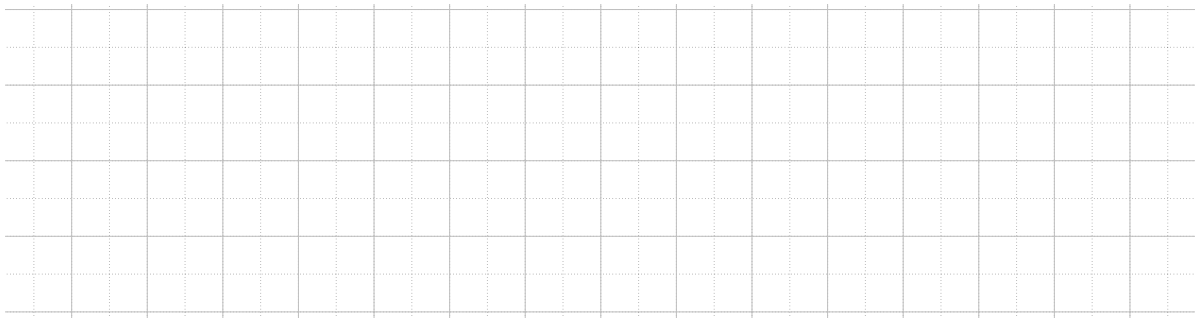


5. Matrices inversas

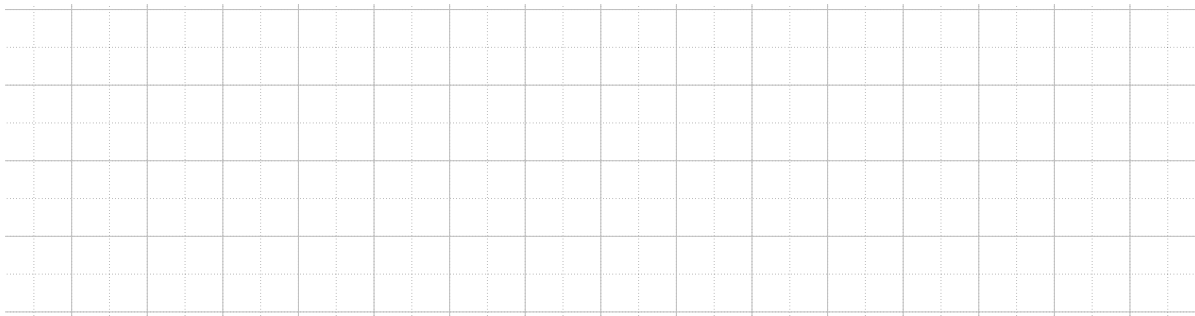
Dado un sistema lineal $A\vec{x} = \vec{b}$, esta sección es dedicada a describir como la solución $\vec{x}$ depende del lado derecho $\vec{b}$. Será necesario encontrar y definir condiciones que justifican la existencia y unicidad de soluciones del sistema cuando el lado derecho $\vec{b}$ es cualquier vector con renglones adecuadas.

Primero motivamos que A tiene que ser cuadrada, *e.g.* $n \times n$, y debe tener rango $r = n$.

Motivación 1: Si $A \in \mathbb{R}^{m \times n}$, entonces la existencia (para cualquier $\vec{b}$) requiere $r = m$:



Motivación 2: Si $A \in \mathbb{R}^{m \times n}$, entonces la unicidad requiere $r = n$:



Conclusión: Existencia y unicidad para cualquier vector $\vec{b}$ solo se puede dar para matrices A cuadradas con rango $r = n$. Es decir, su FER de A debe ser la matriz identidad.

Objetivos:

- Se requiere encontrar una segunda relación entre la solución $\vec{x}$ del sistema $A\vec{x} = \vec{b}$ y el lado derecho $\vec{b}$. En matemáticas si $\vec{b} = f(\vec{x}) = A\vec{x}$ (es una relación), entonces se usa la notación $\vec{x} = f^{-1}(\vec{b})$ donde f^{-1} describe como llegamos de $\vec{b}$ a $\vec{x}$.
- Encontrar condiciones equivalentes para la existencia y unicidad de soluciones y de la función anterior.

Intuición: “duplicar lo que pasa en los reales”

Considere la ecuación $ax = b$ con números $a, b \in \mathbb{R}$ y la incógnita x . Cuando deseamos que esta ecuación tenga solución única pedimos que $a \neq 0$, ya que así existe el “multiplicativo inverso” y se puede despejar x . Veamos lentamente como funciona ese proceso:

$$ax = b \iff a^{-1}(ax) = a^{-1}b \iff (a^{-1}a)x = (a^{-1})b \iff 1 \cdot x = a^{-1}b$$



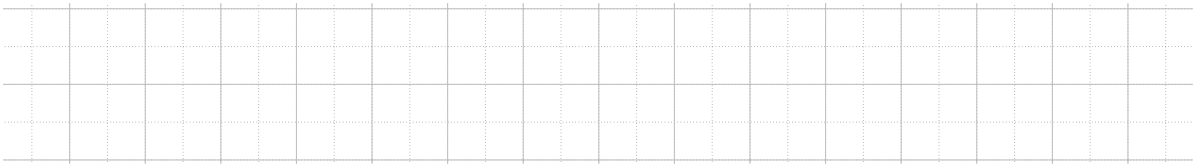
En analogía con la situación en $\mathbb{R}$, si tuviéramos una matriz “ A^{-1} ” con la propiedad $A^{-1}A = I$, entonces podríamos duplicar una dirección de arriba:

$$A\vec{x} = \vec{b} \implies A^{-1}(A\vec{x}) = A^{-1}\vec{b} \iff (A^{-1}A)\vec{x} = (A^{-1})\vec{b} \iff I\vec{x} = A^{-1}\vec{b}.$$



Para tener una equivalencia (en el primer paso) necesitamos que el producto de A & A^{-1} sea conmutativo, es decir, se ocupa $AA^{-1} = I$ para asegurar que la $\vec{x}$ es solución del sistema:

$$I\vec{x} = A^{-1}\vec{b} \implies A(I\vec{x}) = A(A^{-1}\vec{b}) \iff (AI)\vec{x} = (AA^{-1})\vec{b} \iff A\vec{x} = I\vec{b} = \vec{b}.$$



Conclusión: Si se puede resolver el sistema $A\vec{x} = \vec{b}$ con solución única, entonces debe existir una matriz B tal que $BA = I$. Además, la misma matriz debe satisfacer $AB = I$. Una vez mostrado la unicidad de B usáremos el símbolo A^{-1} , pero por el momento la siguiente definición estará suficiente.



Definición. Sea A una matriz cuadrada de $n \times n$. Decimos que A es **invertible** (ó **no singular** ó **regular**) si y solo si existe una matriz B de $n \times n$ que cumple

$$B * A = I \quad \& \quad A * B = I \tag{5.1}$$

donde I es la matriz identidad de $n \times n$. La matriz B es una **inversa** de A .

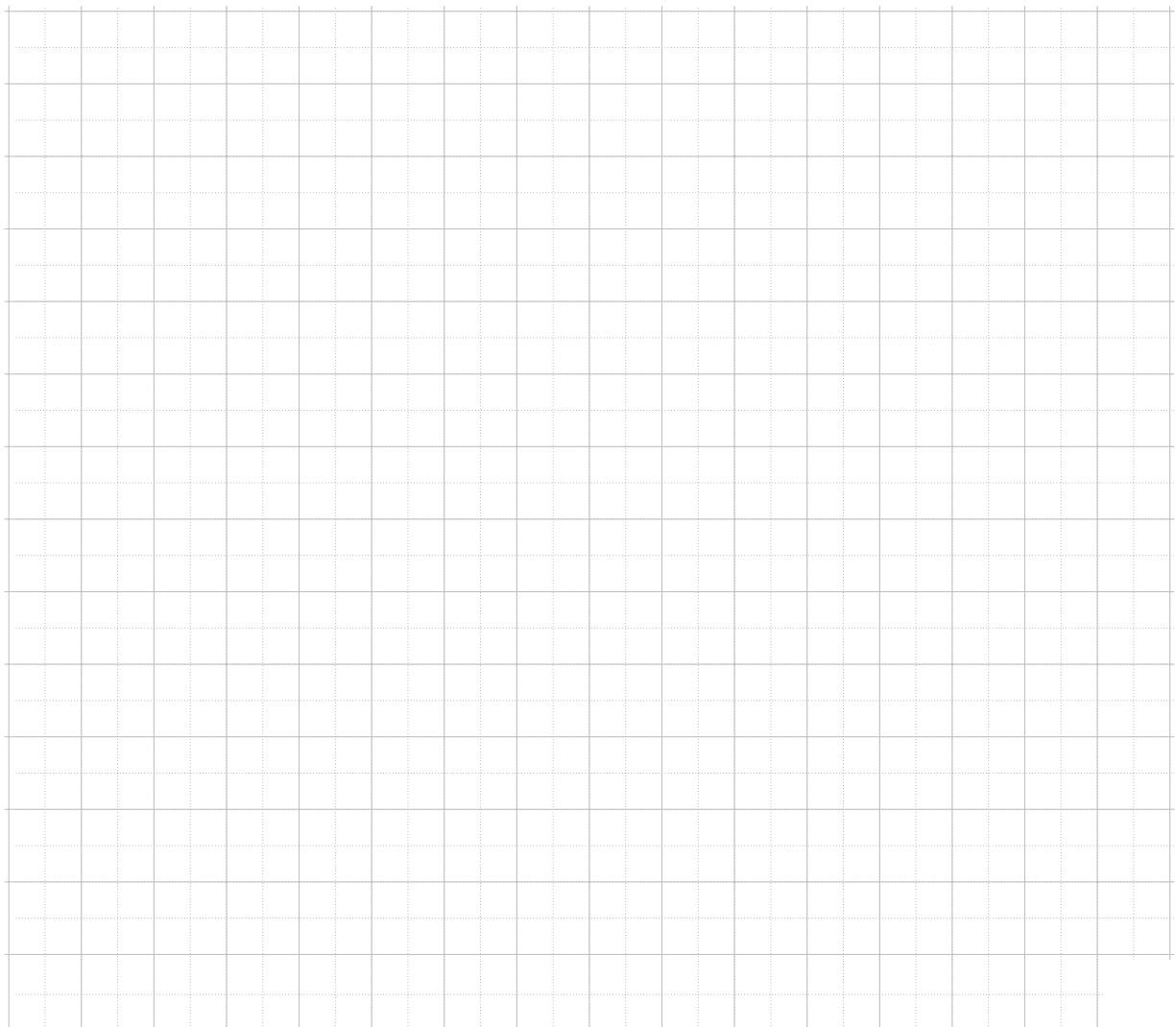
Ejemplo: Compruebe que una inversa de $A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ es $B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$.

Solución: tenemos que calcular los productos AB & BA :



Ejemplo: Demuestre que la matriz $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ no tiene inversa.

Solución: Si A tiene inversa, entonces podríamos encontrar una matriz:



El siguiente resultado confirma la conmutatividad y unicidad. Su demostración sugiere como encontrar la inversa.



Lema 5.1. Suponga, que A es una matriz de $n \times n$ y existe otra matriz B de $n \times n$ tal que $BA = I$. Entonces B satisface $AB = I$ y es única. Esa matriz B es conocida como **la** (matriz) **inversa** de A y será denotada por A^{-1} .

Demostración. Paso 1: Si $BA = I$, entonces existe C tal que $AC = I$.

Si $BA = I$, entonces el sistema lineal $A\vec{x} = \vec{0}$ tiene la solución

$$\vec{x} = I\vec{x} = (BA)\vec{x} = B(A\vec{x}) = B\vec{0} = \vec{0}.$$

Por lo tanto, las columnas de A son L.I. (teorema 4.2). Por eso, la $FER(A)$ es la identidad (pues tiene n unos principales). Como resultado, el sistema $A\vec{x} = \vec{b}$ tiene solución única para cualquier vector. En particular, para los vectores canónicos tenemos soluciones $\{\vec{x}_1, \dots, \vec{x}_n\} \subset \mathbb{R}^n$ definidas por los sistemas

$$A\vec{x}_1 = \vec{e}_1, \quad A\vec{x}_2 = \vec{e}_2, \quad \dots, \quad A\vec{x}_n = \vec{e}_n.$$

Ahora, definimos la matriz C tomando las $\vec{x}_1, \dots, \vec{x}_n$ como sus columnas y multiplicamos

$$AC = A[\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n] = [A\vec{x}_1, A\vec{x}_2, \dots, A\vec{x}_n] = [\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n] = I. \quad (\star)$$

Paso 2: Suponga ahora que $BA = I$ y $AC = I$, entonces $C = IC = (BA)C = B(AC) = BI = B$ ó $B = C$. Con el mismo argumento se muestra que B es única, ya que suponer que existen dos matrices que cumplen ambas igualdades $B_1A = I$ y $AB_2 = I$ resultan en $B_1 = B_2$. $\square$

El método de Gauss-Jordan para calcular la inversa

La idea es usar $(\star)$ que equivale al sistema $AB = I$ y resolverlo.

- *Paso 1:* Dada la matriz A de $n \times n$ se construye la matriz “super-aumentada” $[A | I]$, donde I es la matriz identidad de $n \times n$.

(Esa matriz representa el sistema $AB = I$ con la incógnita B .)

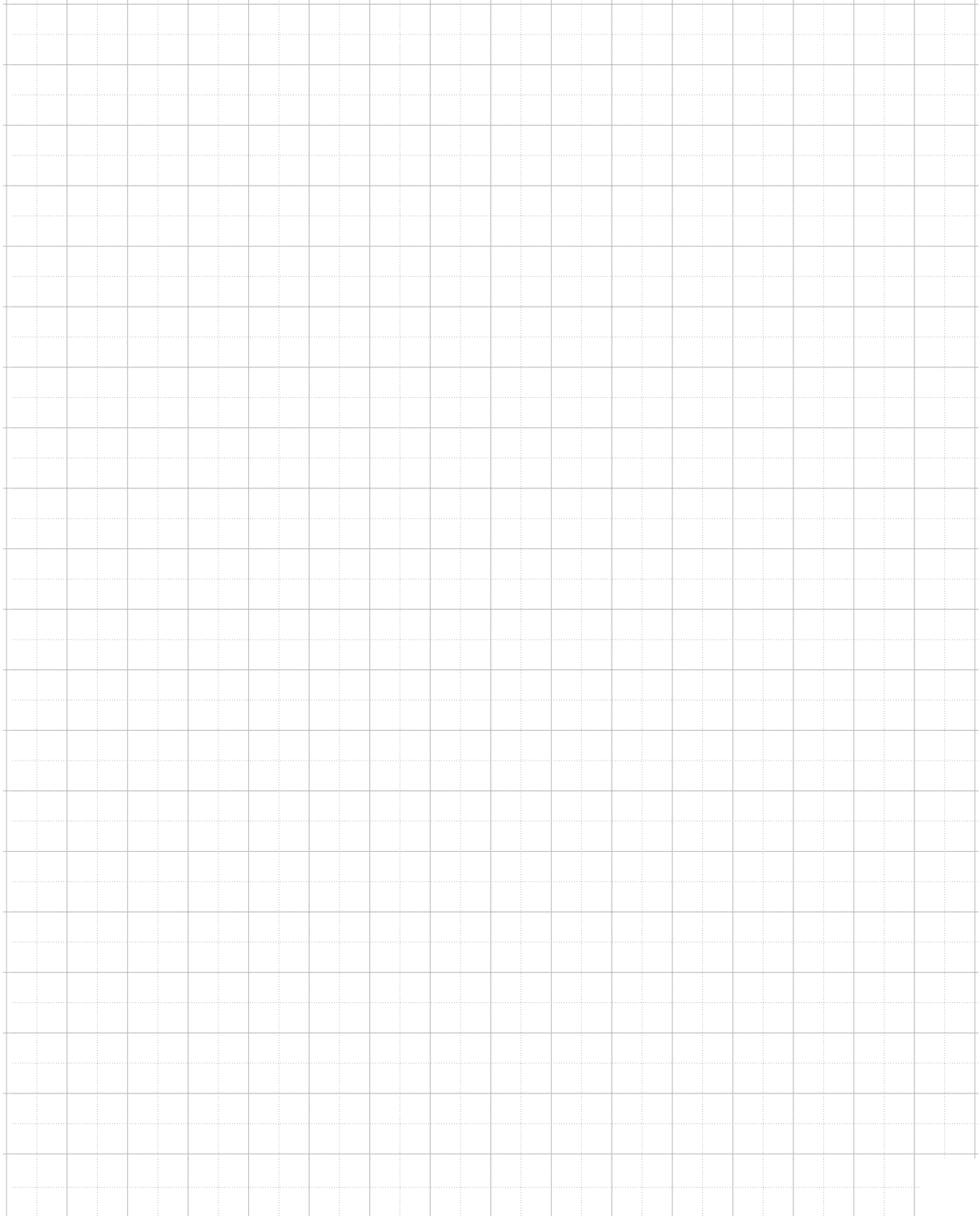
- *Paso 2:* Se aplican operaciones elementales (por renglones, es decir, el método de Gauss-Jordan) con el objetivo de llevar la matriz A a su FER. Si esto resulta posible, entonces

$$[A | I] \rightarrow [I | A^{-1}]$$

- Si el paso 2 falla, entonces A tiene renglones linealmente dependientes y su inversa no existe.

Ejemplo: Encuentre (si existe) la inversa de la matriz

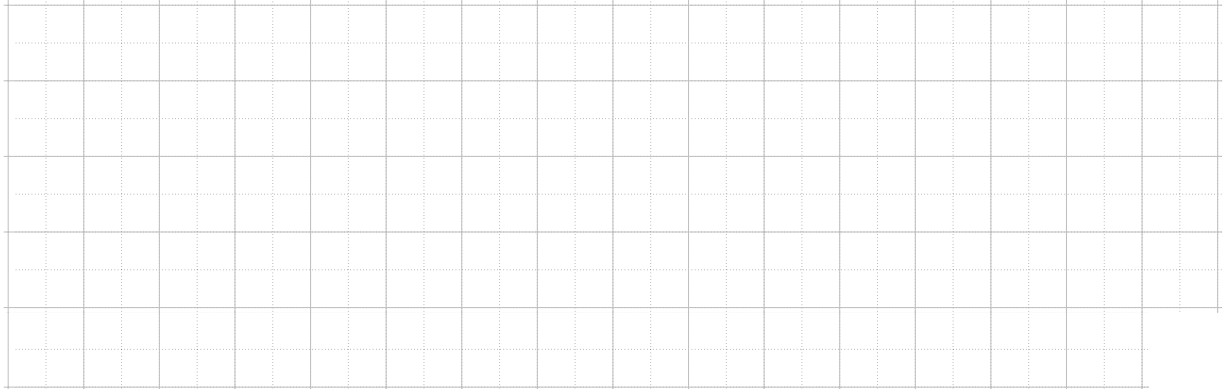
$$A = \begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{bmatrix} .$$



5. Matrices inversas

Ejemplo: Resuelve el sistema $A\vec{x} = \vec{b}$ con $\vec{b} = (1, 2, 3)^\top$ y la inversa de A dada por:

$$A^{-1} = \frac{1}{8} \begin{bmatrix} -5 & 7 & -4 \\ 1 & -3 & 4 \\ 7 & -5 & 4 \end{bmatrix}.$$



Ejemplo: Decide si la siguiente matriz tiene inversa:

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



5.1. Existencia de la inversa (I)

El siguiente teorema da criterios “intuitivos” para cuando existe la inversa de una matriz cuadrada. El nombre contiene “primer” ya que más adelante veremos más criterios.

 **Teorema 5.2** (Primer teorema de la inversa). Sea $A \in \mathbb{R}^{n \times n}$ una matriz cuadrada.

Los siguientes enunciados son equivalentes:

1. A es invertible / A tiene inversa.
2. $A\vec{x} = \vec{b}$ tiene una única solución para todo $\vec{b} \in \mathbb{R}^n$.
3. $A\vec{x} = \vec{0}$ tiene solución única (la trivial).
4. La FER de A es la matriz identidad.
5. Las columnas de A son linealmente independientes.
6. Los renglones de A son linealmente independientes.

 Recuerde que si tenemos un enunciado de la forma $p \implies q$, la contra-positiva es $\neg q \implies \neg p$.

Ejemplo abstracto: $p = “A \text{ es invertible}” \implies q = “A \text{ tiene columnas lin. independientes}”$.

Contra-positiva: $\neg q = “A \text{ tiene columnas lin. dependientes}” \implies \neg p = “A \text{ no tiene inversa}”$.

En palabras: Si “ A tiene columnas lin. dependientes”, entonces “ A no tiene inversa”.



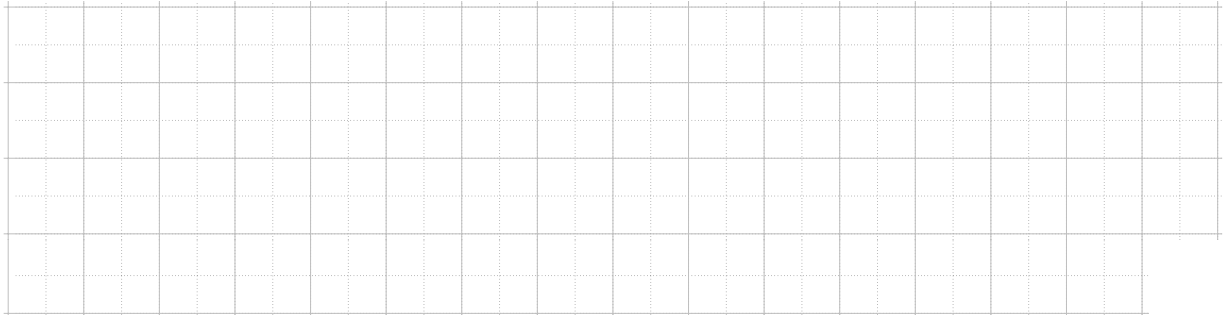
Ejemplos prácticos: En página 121 vimos que la matriz $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ no tiene inversa.



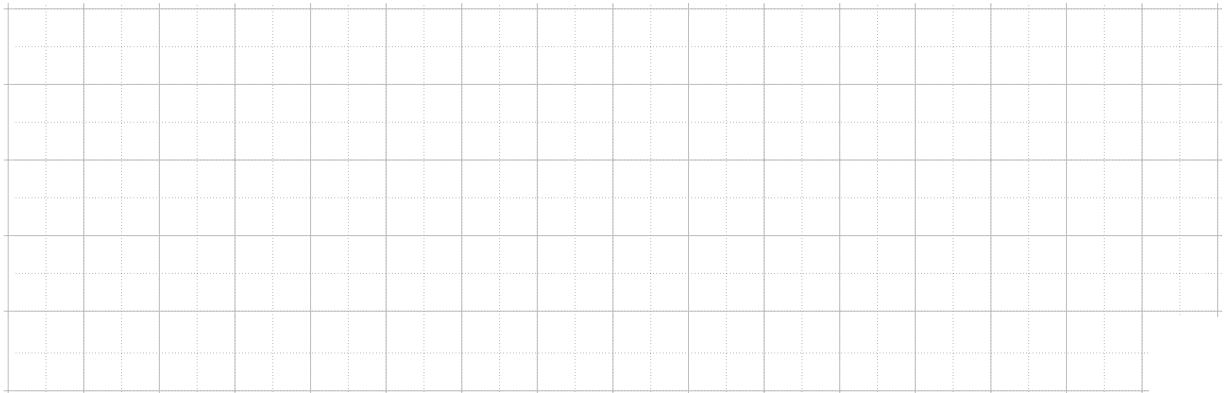
 Si conocemos A^{-1} entonces la solución del sistema $A\vec{x} = \vec{b}$ es $\vec{x} = A^{-1}\vec{b}$.

5.2. Propiedades de la inversa

Ejemplo zoom-zoom:



Ejemplo rotaciones:



Ejemplo rotación y zoom:



Si recuerdas las reglas $(A^\top)^\top = A$, $(AB)^\top = B^\top A^\top$ & $(A^\top)^k = (A^k)^\top$ del Teorema 2.4, entonces las igualdades de las inversas se te facilitarán.

Teorema 5.3. Sean A una matriz invertible, entonces se cumple que:

a) A^{-1} también es invertible y $(A^{-1})^{-1} = A$.

b) Sea $\gamma \in \mathbb{R}$ distinto de cero, entonces γA es una matriz invertible y $(\gamma A)^{-1} = \frac{1}{\gamma} A^{-1}$.

c) $A^\top$ también es una matriz invertible y $(A^\top)^{-1} = (A^{-1})^\top$.

Debido a que no importa ese orden, también se escribe $A^{-\top}$.

d) Si B es otra matriz invertible del tamaño de A , entonces AB es invertible y su inversa es

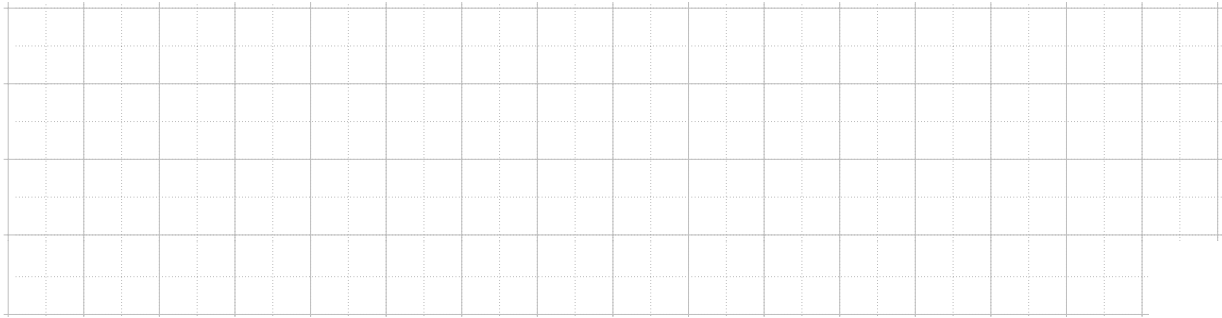
$$(AB)^{-1} = B^{-1}A^{-1}.$$

e) Sea $k \in \mathbb{N}_+$ (un natural positivo) entonces A^k es una matriz invertible y

$$(A^k)^{-1} = (A^{-1})^k.$$

Demostración. Para realizar la demostración de los incisos haremos uso de la ecuación (5.1), es decir, A es invertible si existe una matriz B tal que $AB = I = BA$ y se dice que $B^{-1} = A$.

Para mostrar a) sea $B = A^{-1}$, entonces como A es invertible A tiene inversa y se cumple



b) Sean $A^* = (\frac{1}{\gamma} A^{-1})$ & $B^* = (\gamma A)$. Por demostrar que $A^* B^* = I$ & $B^* A^* = I$.



c) **Ejercicio:**

d) Sean $A^* = B^{-1}A^{-1}$ & $B^* = AB$, entonces debemos mostrar que $A^*B^* = I$ & $B^*A^* = I$.

e) Usamos (de nuevo) inducción sobre k empezando con $k = 2$.

Paso Base:

Sea $k = 2$, entonces $(A^2)^{-1} = (A * A)^{-1} \stackrel{d)}{=} A^{-1} * A^{-1} = (A^{-1})^2$, lo cual confirma la igualdad. (Notar que el producto $A * A$ requiere (por la regla de oro) que A es cuadrada.)

Hipótesis inductiva:

Supongamos que el resultado se cumple para un $k \geq 2$, es decir, $(A^k)^{-1} = (A^{-1})^k$.

Paso inducción:

Tomando la hipótesis de inducción aumentamos k a $k + 1$:

Empecemos con el lado izquierdo:

$$(A^{k+1})^{-1} = (A^k * A)^{-1} \stackrel{d)}{=} A^{-1}(A^k)^{-1} \stackrel{B/H.d.I.}{=} A^{-1}(A^{-1})^k = (A^{-1})^{k+1}.$$

Con esto se termina la demostración.



5.3. Matrices elementales y sus inversas

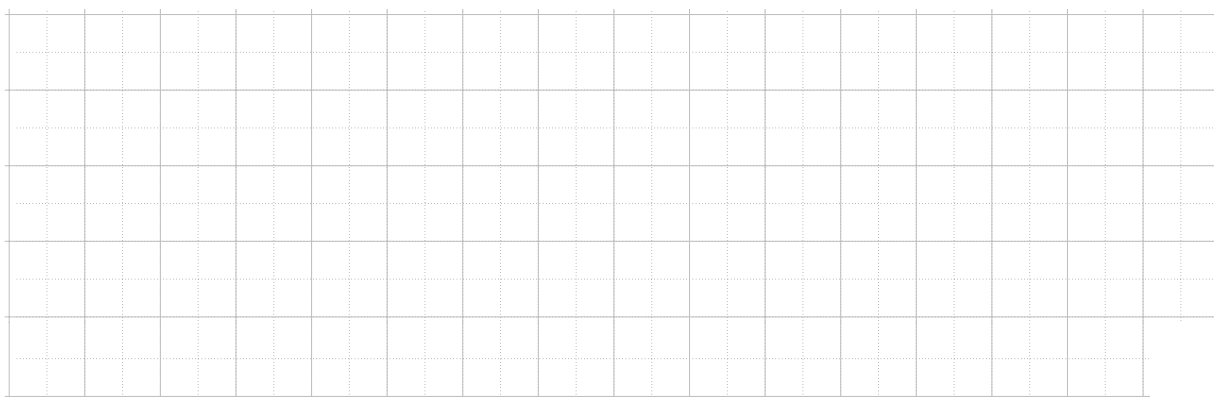


Definición. Una matriz cuadrada de $n \times n$ es una **matriz elemental** si es el resultado de aplicar **una sola** operación elemental (página 19) a la matriz identidad I . $\therefore$ Existen 3 tipos.



El producto de una “matriz elemental” y otra matriz A efectúa la operación elemental en A .

Ejemplo tipo 1: Intercambio de filas.



Ejemplo tipo 2: multiplicar por un número distinto de cero.



Ejemplo tipo 3: multiplicar una fila por un número y sumar el resultado a otra fila.



5.3.1. Eliminación Gaussiana con matrices elementales

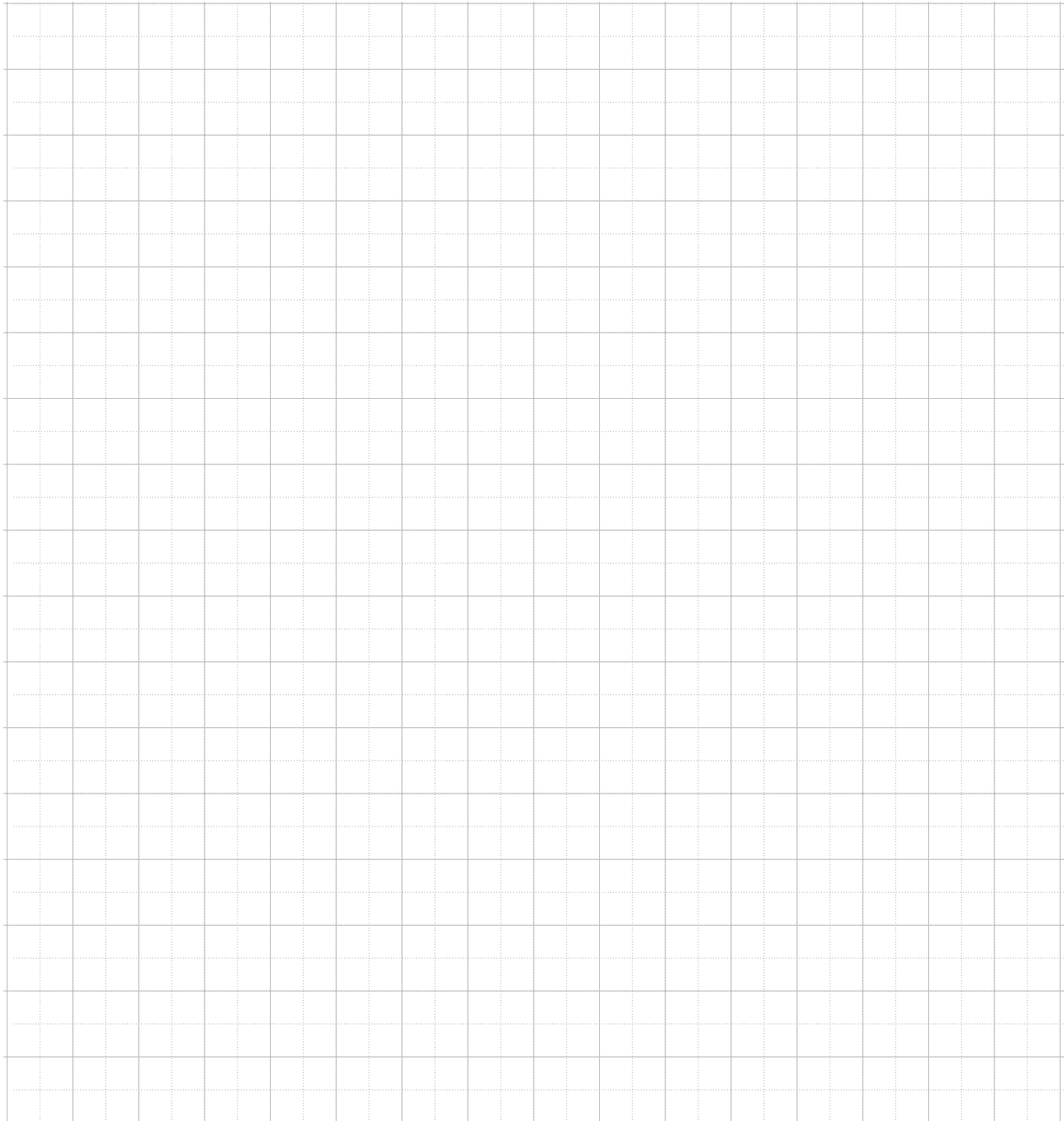
La eliminación Gaussiana se puede ver como muchos productos matriciales. Si “acumulamos” las matrices elementales, entonces podemos escribir A & A^{-1} como producto de ellas.

Ejemplo: Resuelve el siguiente sistemas de ecuaciones lineales y compare su respuesta con la multiplicación de matrices elementales:

$$x + 2y + 2z = 1$$

$$4x + 8y + 9z = 3$$

$$3y + 2z = 1$$





 Una forma escalonada de A se puede encontrar multiplicando la matriz A (siempre por la izquierda) por matrices elementales. Si esto se hace **sin** matrices elementales del tipo 2, entonces los unos principales no se pueden generar, pero A se pone en forma triangular superior U .

5. Matrices inversas

En el ejemplo anterior vimos un caso especial del siguiente resultado:



Lema 5.4. Se A es una matriz de $n \times n$ invertible, entonces existen k matrices elementales $E_1, \dots, E_k$ tales que $E_k \cdots E_2 E_1 A = I$. De esto se sigue que

$$A^{-1} = E_k \cdots E_2 E_1 \quad \& \quad A = E_1^{-1} E_2^{-1} \cdots E_k^{-1}.$$

Demostración. Por teorema 5.2 sabemos que la FER de A es la matriz identidad. Entonces, existen $k \in \mathbb{N}$ operaciones elementales que convierten A en I . Para cada operación elemental existe una matriz elemental que realiza la operación elemental como producto de matrices, por lo tanto, tenemos $E_k \cdots E_2 E_1 A = I$.

Por el teorema 2.3 sabemos que la multiplicación es asociativa por lo cual el el primer producto, es decir, la matriz $B = (E_k \cdots E_2 E_1)$ satisface $BA = I$. Entonces, por Lemma 5.1 concluimos que $AB = I$ y $B = A^{-1}$.

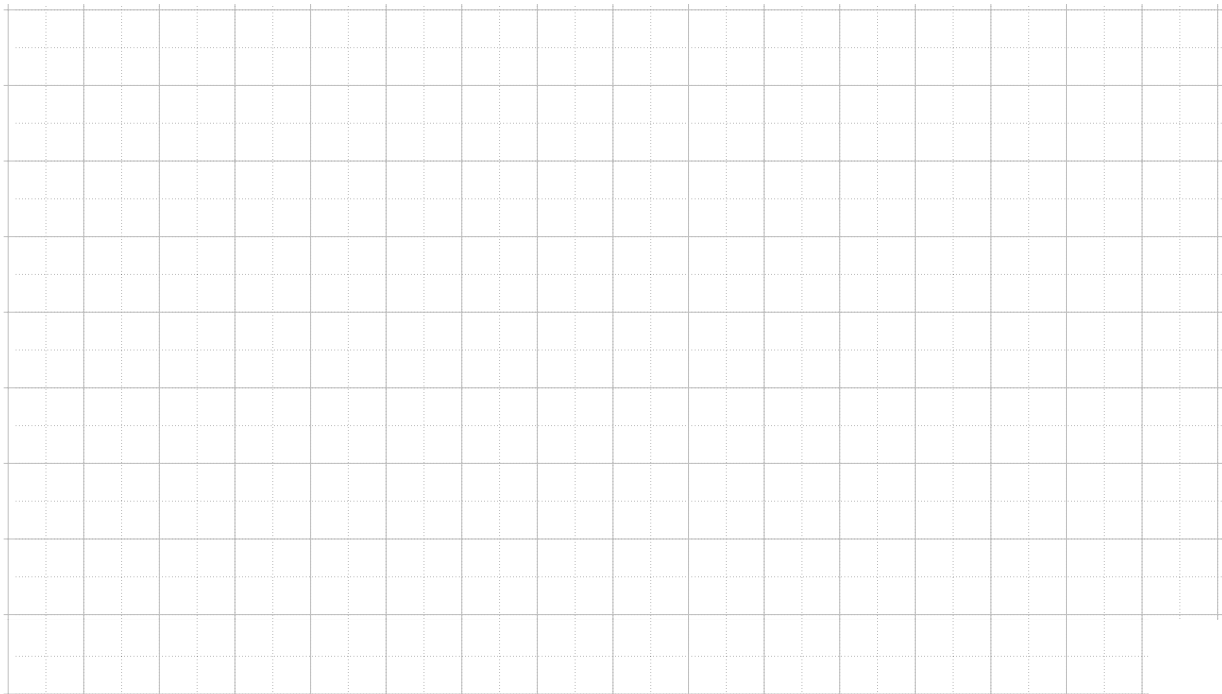
Por el teorema 5.3d) sabemos que la inversa de $B = E_k \cdots E_2 E_1$ es $B^{-1} = E_1^{-1} E_2^{-1} \cdots E_k^{-1}$ y por propiedad a) concluimos $B^{-1} = (A^{-1})^{-1} = A$. $\square$



Lema 5.5. Si $E \in \mathbb{R}^{n \times n}$ es una matriz elemental, entonces E es invertible y E^{-1} es elemental del mismo tipo que E . Más aún

1. Si E es de tipo 1, entonces $E^{-1} = E$.
2. Si E es de tipo 2 e $I \xrightarrow{\gamma R_i} E$, entonces E^{-1} se obtiene de $I \xrightarrow{\gamma^{-1} R_i} E^{-1}$.
3. Si E es de tipo 3 e $I \xrightarrow{\gamma R_i + R_j} E$, entonces E^{-1} se obtiene de $I \xrightarrow{-\gamma R_i + R_j} E^{-1}$.

Ejemplos:



5.4. Una formula rápida para la inversa de 2×2

Para cada matriz $A \in \mathbb{R}^{n \times n}$ existe un número (escalar) que determina si A es invertible.

Ese número será llamado “determinante de A ” $\det(A)$ en la siguiente sección.

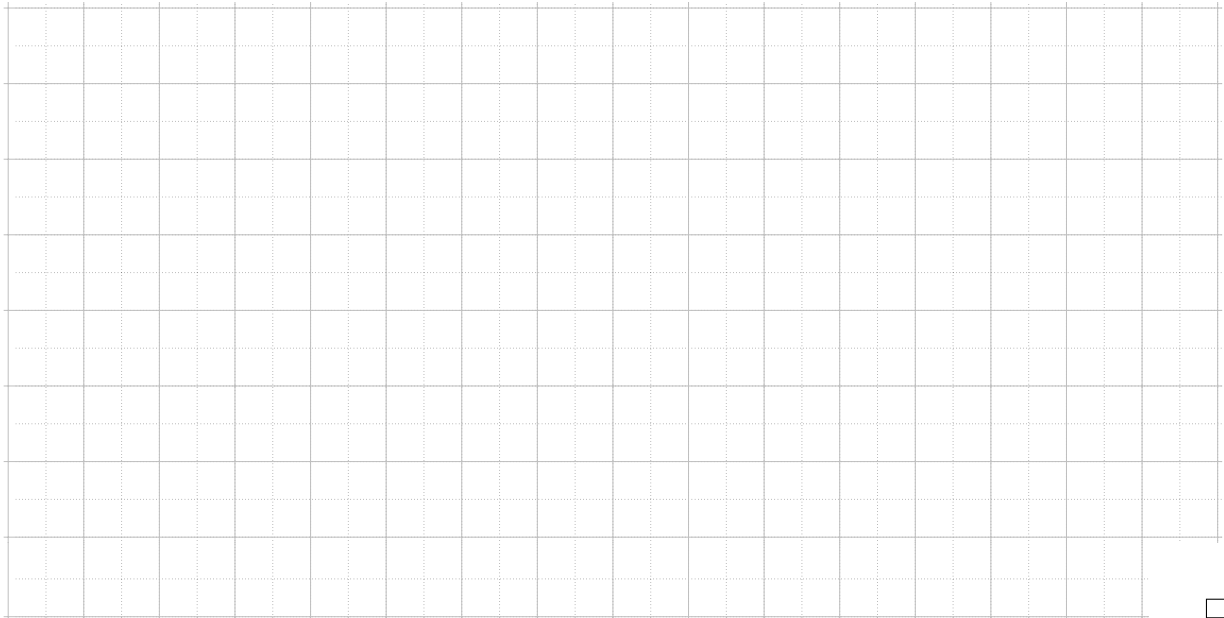
Para matrices 2×2 el número es sencillo.

 **Lema 5.6.** Si $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es invertible, entonces $ad - bc \neq 0$.

Demostración. Por hipótesis A es invertible y por ser cuadrada el teorema 5.2 implica que sus renglones y columnas son linealmente independientes. Como un renglón y una columna de ceros es lin. dependiente, podemos hacer dos casos.

Caso 1: Si $a = 0$, entonces c & b son distintos de cero y por lo tanto $ad - bc = -cb \neq 0$.

Caso 2: Si $a \neq 0$, entonces

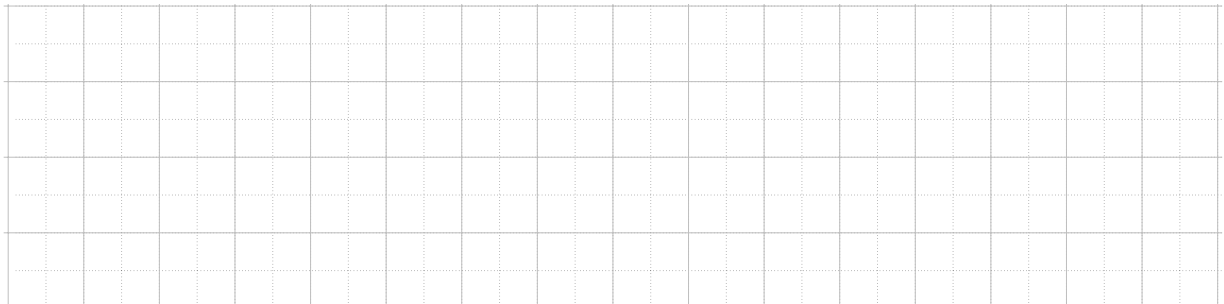


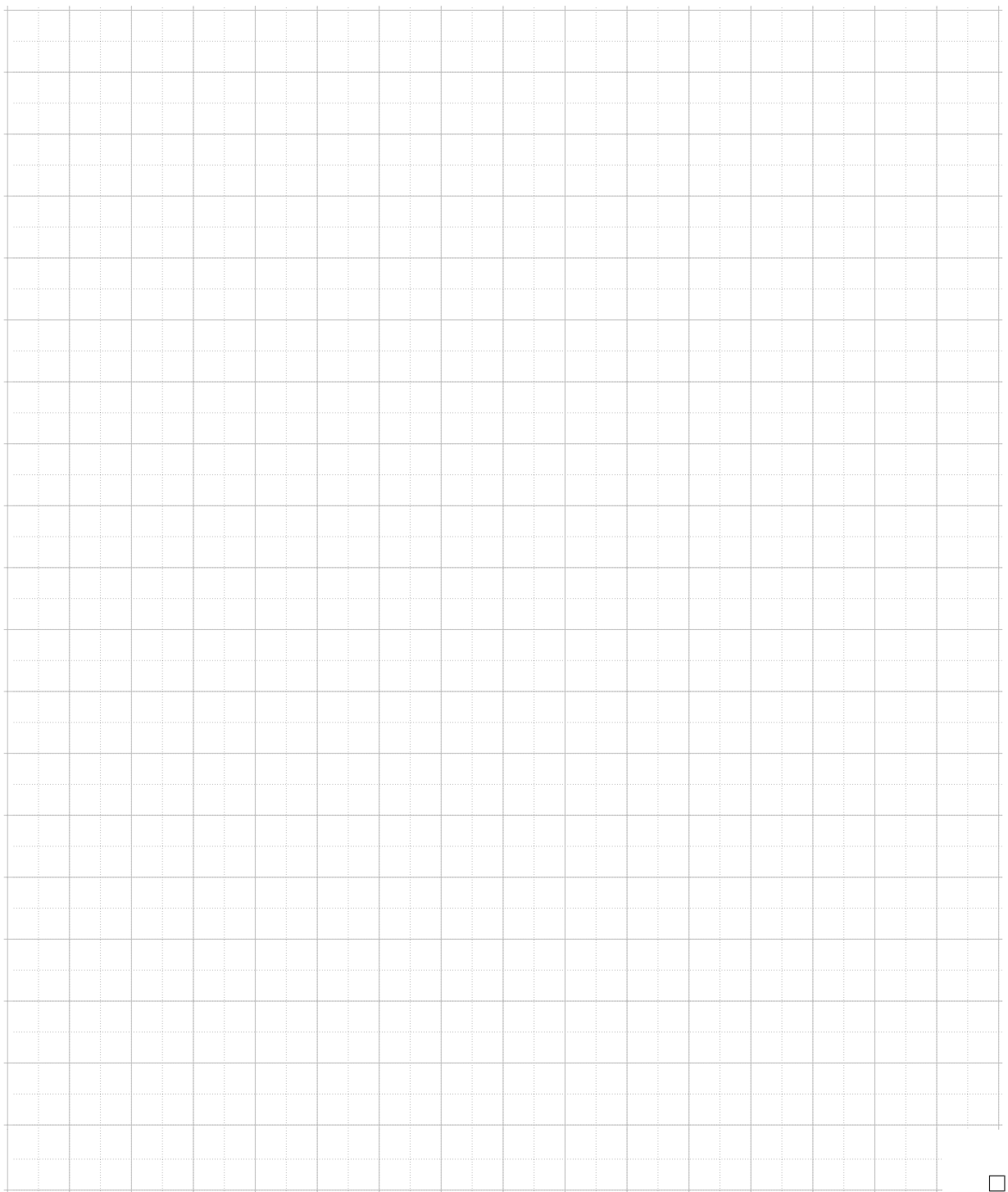
□

 **Lema 5.7.** Si la matriz $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ satisface $ad - bc \neq 0$, entonces A es invertible y

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (5.2)$$

Demostración.





6. Determinantes

Históricamente los determinantes fueron introducidos (siglo XVII) antes que las matrices (siglo XIX). El concepto de determinante apareció por primera vez en la publicaciones del matemático japonés de Seki en 1683. [en.wikipedia.org, visto el 16.10.2024]

La teoría en esta sección será tomada del libro de los profesores Mota & Rumbos.



Definición 6.1. El **determinante** de una matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ de 2×2 se define como

$$\det(A) = ad - bc.$$



Teorema 6.1. La matriz $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ es invertible si y solamente si $\det(A) = ad - bc \neq 0$. Además:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \& \quad A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Demostración. Si A es invertible, Lemma 5.6 muestra $\det(A) = ad - bc \neq 0$. Por otro lado, si $\det(A) = ad - bc \neq 0$, Lemma 5.7 muestra que A es invertible y da la fórmula. □



Recuerde que si tenemos un enunciado de la forma $p \implies q$, la contra-positiva es $\neg q \implies \neg p$.

Ejemplo abstracto: $p = "A \text{ es invertible}" \implies q = "\det(A) \neq 0"$.

Contra-positiva: $\neg q = "\det(A) = 0" \implies \neg p = "A \text{ no tiene inversa}"$

Por teo. 5.2 ser cuadrada y no tener inversa equivale a que A tiene columnas y renglones LD.

Ejemplo: ¿Las matrices $A = \begin{pmatrix} 4 & 5 \\ 2 & 3 \end{pmatrix}$ & $B = \begin{pmatrix} 4 & 6 \\ 2 & 3 \end{pmatrix}$ tienen inversa?

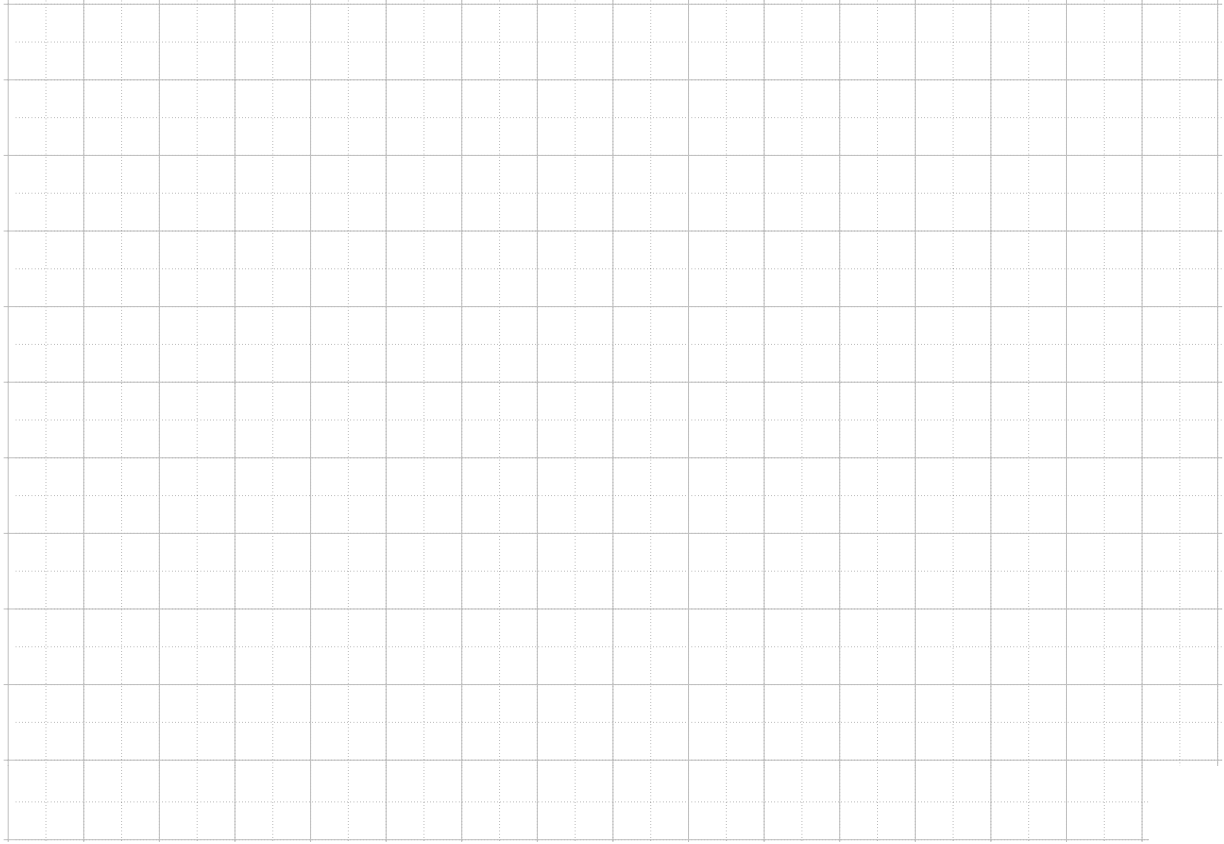


Interpretación geométrica (para Cálculo III).

 | El determinante es el factor con el cual se aumenta / reduce el volumen.

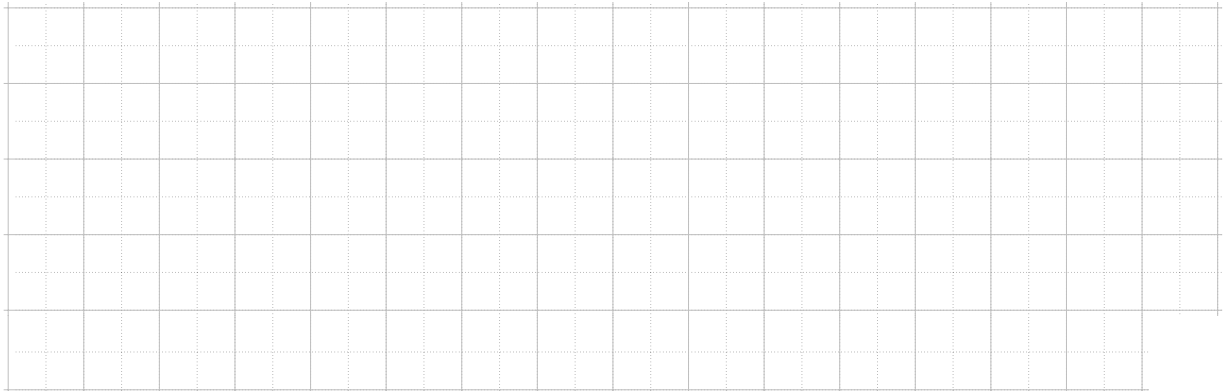
Ejemplo: Sea $A = \begin{pmatrix} 1 & 2 \\ 0 & -3 \end{pmatrix}$, entonces $\det(A) = 3$.

Calcule el área encerrado por $\vec{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{x}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$, $\vec{x}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ & $\vec{x}_4 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ y el área encerrado por $\vec{y}_1 = A\vec{x}_1$, $\vec{y}_2 = A\vec{x}_2$, $\vec{y}_3 = A\vec{x}_3$ & $\vec{y}_4 = A\vec{x}_4$.



Ejemplo: Sea $A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$, entonces $\det(A) = 0$.

Calcule el área encerrado por $\vec{x}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\vec{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\vec{x}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ & $\vec{x}_4 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ y el área encerrado por $\vec{y}_1 = A\vec{x}_1$, $\vec{y}_2 = A\vec{x}_2$, $\vec{y}_3 = A\vec{x}_3$ & $\vec{y}_4 = A\vec{x}_4$.



Importante: El volumen se mantiene positivo si y solo si $\det(A) \neq 0 \iff A^{-1}$ existe.

6.1. El determinante es una función multilinear.



Definición 6.2. Una función $F: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ que toma una matriz y asigna un valor real es **multilinear en los renglones** si satisface las siguientes dos propiedades para cada $i \in \{1, 2, \dots, n\}$:

- i) Sea $A \in \mathbb{R}^{n \times n}$ una matriz y $A_{i\gamma}$ la matriz que se obtiene de A multiplicamos la i -ésima fila por el número $\gamma \in \mathbb{R}$, es decir

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix} \quad \& \quad A_{i\gamma} = \begin{pmatrix} A_1 \\ \vdots \\ \gamma A_i \\ \vdots \\ A_n \end{pmatrix},$$

entonces la función F debe cumplir

$$F(A_{i\gamma}) = \gamma F(A).$$

- ii) Sean $A, A', B \in \mathbb{R}^{n \times n}$ matrices tales que A, A', B solo son diferentes en el renglón i y B contiene en su renglón i la suma de los renglones i de A, A' , es decir

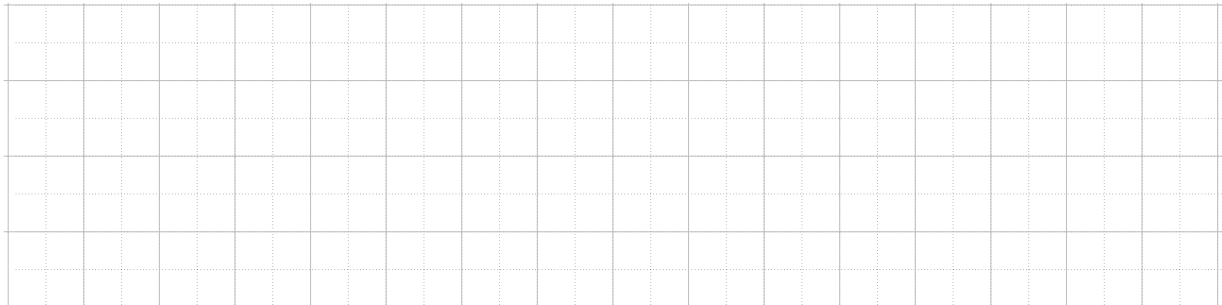
$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_n \end{pmatrix}, \quad A' = \begin{pmatrix} A_1 \\ \vdots \\ A'_i \\ \vdots \\ A_n \end{pmatrix} \quad \& \quad B = \begin{pmatrix} A_1 \\ \vdots \\ A_i + A'_i \\ \vdots \\ A_n \end{pmatrix},$$

entonces la función F debe cumplir

$$F(B) = F(A) + F(A').$$

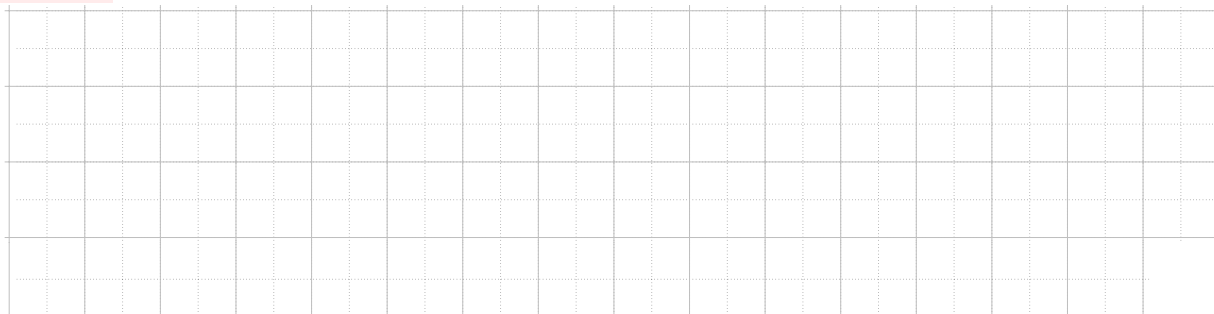
Ejemplo: Demuestre que el determinante de una matriz 2×2 es una función multilinear (por renglones). Verifique las propiedades de definición 6.2 para las matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A' = \begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}, \quad A_{1\gamma} = \begin{pmatrix} \gamma a & \gamma b \\ c & d \end{pmatrix} \quad \& \quad B = \begin{pmatrix} a + a_2 & b + b_2 \\ c & d \end{pmatrix}.$$





Ejercicio: Sea $A, I \in \mathbb{R}^{2 \times 2}$. Verifique que $\det(A) = 0$ si A tiene 2 renglones iguales y $\det(I) = 1$.



En general, el determinante de una matriz de $n \times n$ es una función multilineal que adicionalmente cumple lo del ejercicio anterior.



Definición 6.3. La **función determinante** ó el **determinante**

$$\det: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$$

satisface las siguientes propiedades

1. $\det$ es una función multilineal en los renglones.
2. $\det(A) = 0$ cuando A tiene dos renglones consecutivos iguales.
3. $\det(I) = 1$.

Nota: Algunos autores (*e.g.* en Cálculo) ocupan la notación de $|A|$ en lugar de $\det(A)$.

De la definición 6.3 salen muchas propiedades generales que permiten construir una nueva forma de resolver sistemas lineales (la “Regla de Cramer”), calcular el determinante para matrices $n \times n$ y verificar que el resultado cuando $n = 2$ coincide con el valor dado por la definición 6.1.

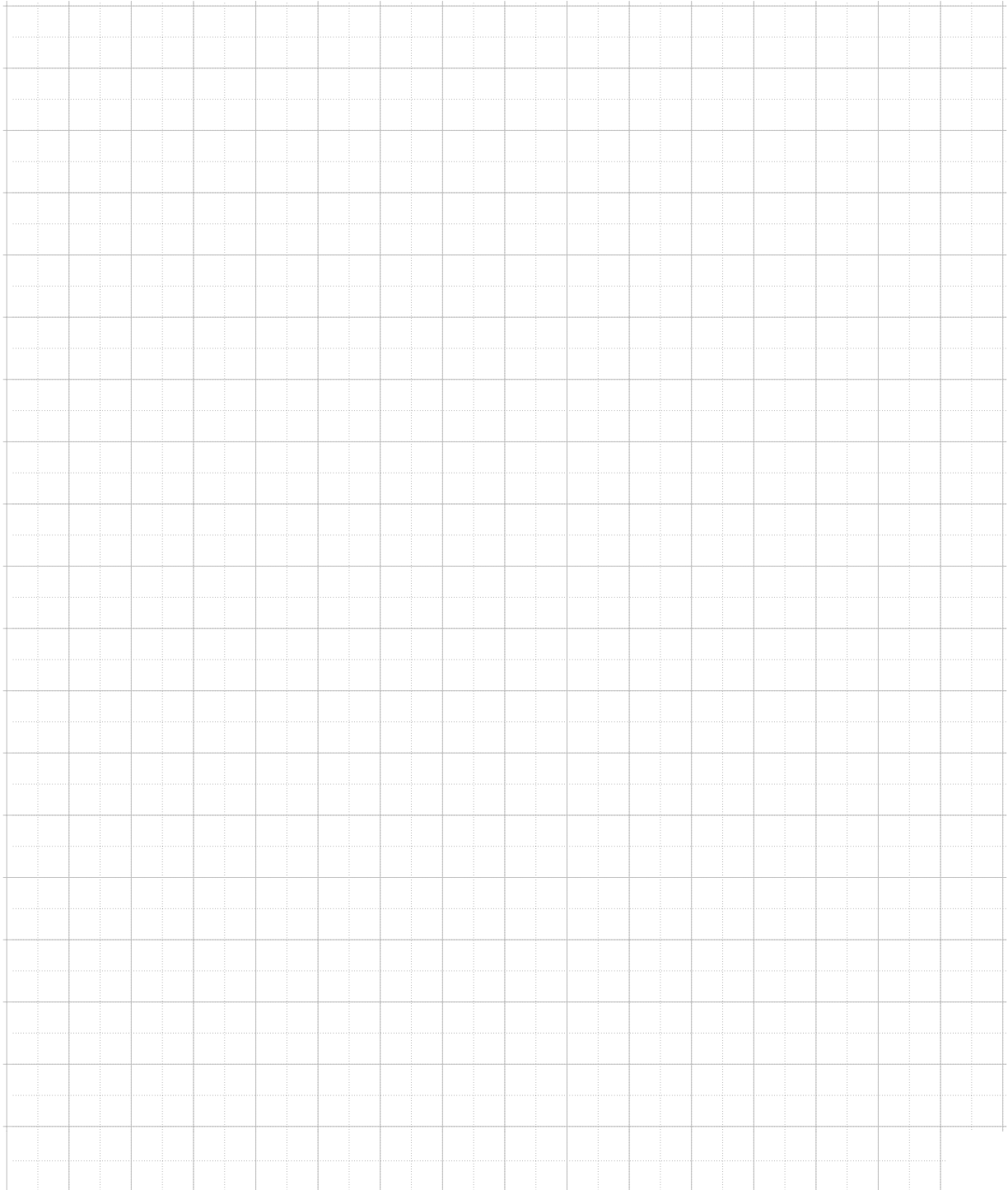
Ejemplo: Cambiamos renglones de una matriz y vemos como cambia el determinante.



Ejemplo: “Multilinealidad” Si $\det(A) = 42$ y

$$A = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & \ell \\ m & n & o & p \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \frac{2}{3}a & \frac{2}{3}b & \frac{2}{3}c & \frac{2}{3}d \\ e & f & g & h \\ 66i & 66j & 66k & 66\ell \\ \pi m & \pi n & \pi o & \pi p \end{bmatrix}$$

¿Cuál es el determinante de B ?



6.2. Propiedades de determinantes

El objetivo de esta sección es demostrar 4 propiedades principales de determinantes. Estas servirán más adelante en esta materia y luego en otras materias avanzadas. También sirven para entender la formula del determinante que veremos más adelante.

Las 4 propiedades:

1. Si A tiene dos renglones iguales, entonces $\det(A) = 0$.
(Esa da determinantes de matrices elementales)
2. Una matriz cuadrada A es invertible si y solo si $\det(A) \neq 0$.
3. Si A, B son matrices cuadradas, entonces $\det(AB) = \det(A)\det(B)$.
Interpretación: Dos matrices A, B son invertibles si y solo si su producto AB es invertible.
4. Sea A una matriz cuadrada, entonces $\det(A^\top) = \det(A)$.
Importante para valores propios (más tarde).

6.2.1. La primera propiedad

Esa propiedad permite calcular determinantes de matrices elementales (el corazón del resto).

Paso 1. Si A' se obtiene a partir de A intercambiando dos renglones consecutivos, digamos renglón i e $i + 1$, entonces $\det(A') = -\det(A)$.

Demostración. Eso se demuestra usando definición 6.3 2) y definición 6.2 ii):

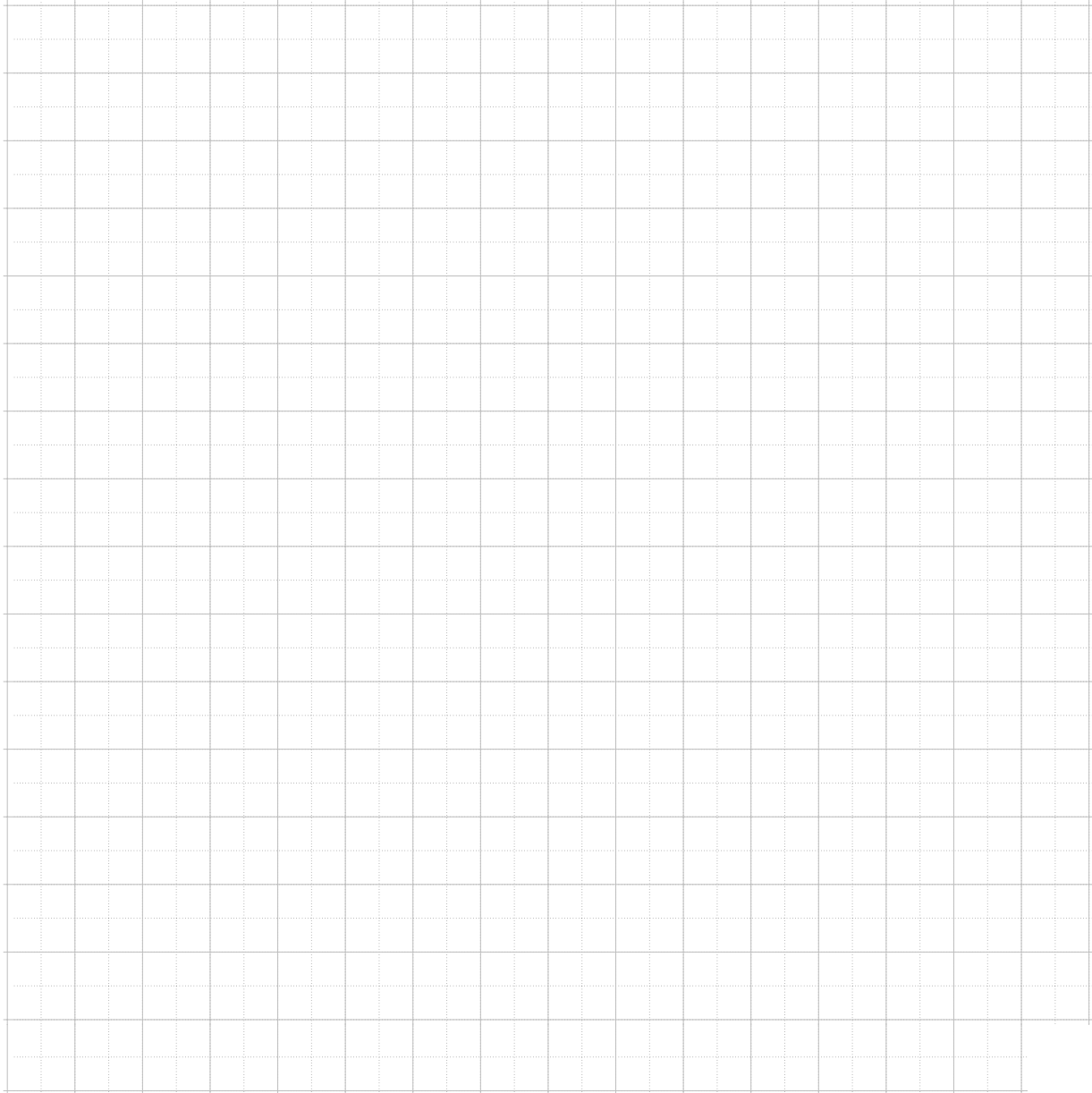
$$\begin{aligned}
 \text{Por def. 6.3 2):} \quad 0 &= \det \begin{pmatrix} \vdots \\ A_i + A_{i+1} \\ A_i + A_{i+1} \\ \vdots \end{pmatrix} \\
 \text{Por def. 6.2 ii):} \quad &= \det \begin{pmatrix} \vdots \\ A_i \\ A_i + A_{i+1} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_{i+1} \\ A_i + A_{i+1} \\ \vdots \end{pmatrix} \\
 \text{Por def. 6.2 ii):} \quad &= \det \begin{pmatrix} \vdots \\ A_i \\ A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_i \\ A_{i+1} \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_{i+1} \\ A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_{i+1} \\ A_{i+1} \\ \vdots \end{pmatrix} \\
 \text{Por def. 6.3 2):} \quad &= 0 + \det(A) + \det(A') + 0.
 \end{aligned}$$



Es decir, $0 = \det(A) + \det(A')$ o bien $\det(A') = -\det(A)$.

□

Ejemplo: La definición 6.3 generaliza y no contradice definición 6.1. Para verificar esto, calculamos el determinante de una matriz 2×2 usando definición 6.3.



 | **Lema 6.2** (Ayuda). Si A tiene dos renglones iguales, entonces $\det(A) = 0$.

Demostración. Sea A igual en los renglones i & k con $i < k$, entonces $k = i + 1 + d$ para $d = k - (i + 1)$.

Si $d = 0$, entonces $k = i + 1$ y A tiene dos renglones **consecutivos** iguales, con lo cual la definición 6.3(2) da $\det(A) = 0$.

Si $d > 0$, entonces se cambia el renglón k con el renglón anterior d veces. Después, de d cambios tendremos una matriz A_d tal que el renglón k de A se encuentre en el renglón $i + 1$ de A_d , por lo cual los renglones $i, i + 1$ de A_d son iguales y por definición 6.3(2) se tiene $\det(A_d) = 0$. Además, por paso 1 tenemos $\det(A_d) = (-1)^d \det(A)$ y se concluye $\det(A) = (-1)^d \det(A_d) = 0$. $\square$

6.2.2. Determinantes y operaciones elementales

Para demostrar la equivalencia vemos primero como operaciones elementales afectan determinantes.



Lema 6.3. Sea $A \in \mathbb{R}^{n \times n}$ es una matriz cuadrada, las operaciones elementales afectan el determinante como sigue:

1. Si $A \xrightarrow{R_i \leftrightarrow R_k} A'$, entonces $\det(A') = -\det(A)$.
2. Si $A \xrightarrow{\gamma R_i} A'$, entonces $\det(A') = \gamma \det(A)$.
3. Si $A \xrightarrow{\gamma R_i + R_j} A'$, entonces $\det(A') = \det(A)$.

Demostración. Consideramos la primera operación $A \xrightarrow{R_i \leftrightarrow R_k} A'$. Usando los renglones de A definimos una matriz de ayuda que tiene en el renglón i y renglón k la suma de los renglones A_i, A_k de A . Los puntos verticales representan otros renglones de A , por lo que se tiene

$$\begin{aligned}
 \text{Por lema 6.2:} \quad 0 &= \det \begin{pmatrix} \vdots \\ A_i + A_k \\ \vdots \\ A_i + A_k \\ \vdots \end{pmatrix} \\
 \text{Por def. 6.2 ii):} \quad &= \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_k \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_k \\ \vdots \\ A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_k \\ \vdots \\ A_k \\ \vdots \end{pmatrix} \\
 \text{Por lema. 6.2:} \quad &= 0 + \det(A) + \det(A') + 0.
 \end{aligned}$$

Se concluye que $0 = \det(A) + \det(A')$ o bien $\det(A') = -\det(A)$.

Ahora, la segunda operación $A \xrightarrow{\gamma R_i} A'$ elemental solo para $\gamma \neq 0$. Por definición 6.3 $\det$ es una función multilineal y se cumple la propiedad i) de la definición 6.2. Por lo tanto, dado que $A' = A_{i\gamma}$ concluimos $\det(A') = \det(A_{i\gamma}) = \gamma \det(A)$.

Finalmente, consideramos $A \xrightarrow{\gamma R_i + R_j} A'$. Aquí, la definición 6.2 ii), i) y lema 6.2 dan:

$$\begin{aligned}
 \det(A') &= \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ \gamma A_i + A_j \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ \gamma A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \end{pmatrix} = \gamma \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ A_i \\ \vdots \\ A_j \\ \vdots \end{pmatrix} \\
 &= \gamma 0 + \det(A) = \det(A).
 \end{aligned}$$

Para la tercera operación elemental concluimos que $\det(A') = \det(A)$. □

Consecuencias del lema 6.3:

Paso 2. Si en lema 6.3 $A = I$, entonces A' es una matriz elemental y su determinante es:

1. Si $I \xrightarrow{R_i \leftrightarrow R_k} E_{t1}$, entonces $\det(E_{t1}) = -\det(I) = -1$.
2. Si $I \xrightarrow{\gamma R_i} E_{t2}$, entonces $\det(E_{t2}) = \gamma \det(I) = \gamma$.
3. Si $I \xrightarrow{\gamma R_i + R_j} E_{t3}$, entonces $\det(E_{t3}) = \det(I) = 1$.

Paso 3. Las operaciones elementales son productos con matrices elementales y se cumple:

5. Dado que $A \xrightarrow{R_i \leftrightarrow R_k} A' \iff A' = E_{t1}A$, se cumple

$$\det(E_{t1}A) = \det(A') = -\det(A) = \det(E_{t1}) \det(A).$$

6. Dado que $A \xrightarrow{\gamma R_i} A' \iff A' = E_{t2}A$, se cumple

$$\det(E_{t2}A) = \det(A') = \gamma \det(A) = \det(E_{t2}) \det(A).$$

7. Dado que $A \xrightarrow{\gamma R_i + R_j} A' \iff A' = E_{t3}A$, se cumple

$$\det(E_{t3}A) = \det(A') = \det(A) = \det(E_{t3}) \det(A).$$

Paso 4. Sean $E(1), E(2), \dots, E(k)$ matrices elementales de cualquier tipo, entonces

$$\begin{aligned} \det(E(k) \cdots E(2)E(1)A) &= \det(E(k)) \cdot \dots \cdot \det(E(2)) \cdot \det(E(1)) \cdot \det(A) \\ &= \det(A) \prod_{i=1}^k \det(E(i)). \end{aligned} \tag{6.1}$$

Demostración. La asociatividad del producto de matrices y las propiedades anteriores permiten cambiar una-por-una la matriz elemental por su determinante. $\square$

Paso 5. Si una matriz A de $n \times n$ tiene un renglón de ceros, entonces $\det(A) = 0$.

Demostración. Suponga que el renglón i de A es cero. Multiplicamos ese renglón con cero para usar la multi-linealidad (definición 6.2 propiedad i), es decir $A \xrightarrow{0R_i} A_{i0}$ implica $\det(A_{i0}) = 0 \det(A) = 0$. Pero, dado que el renglón ya era cero tenemos $A = A_{i0}$ y $\det(A) = \det(A_{i0}) = 0$. $\square$

Contra-positiva: Si $\det(A) \neq 0$, entonces A no tiene un renglón de ceros.

 **Teorema 6.4.** Sea A una matriz cuadrada, entonces A es invertible si y solo si $\det(A) \neq 0$.

Demostración. Sea A' la FER de A que se obtiene aplicando operaciones elementales a la matriz A . Por lema 5.4 existen matrices elementales tales que $E(k) \cdots E(1)A = A'$.

Si A es invertible, entonces la FER de A es la identidad y $A' = I$ y ecuación (6.1) nos permite escribir $1 = \det(I) = \det(A') = \det(A) \prod_{i=1}^k \det(E(i))$. Por lo tanto, $\det(A) \neq 0$.

Por otro lado, si $0 \neq \det(A)$, entonces $0 \neq \det(A) \prod_{i=1}^k \det(E(i)) = \det(A')$, ya que matrices elementales tienen $\det(E(i)) = \gamma_i \neq 0$ (algunos ± 1), ver paso 2. Entonces $\det(A') \neq 0$, es decir, A' no tiene un renglón de ceros (contra positiva de paso 5). Es decir, por ser cuadrada la FER de A tiene rango $r = n$ y A es invertible. $\square$

 El teorema 6.4 negado:

$$\det(A) = 0 \iff \text{“} A \text{ no es invertible”} \quad (6.2)$$

Ya que individualmente podemos usar contra-positivas:

Positiva	Contra-positiva
$A \text{ invertible} \implies \det(A) \neq 0$	$\det(A) = 0 \implies A \text{ no invertible}$
$\det(A) \neq 0 \implies A \text{ invertible}$	$A \text{ no invertible} \implies \det(A) = 0$

6.2.3. Determinantes de productos

El siguiente lema y teorema podrían estar en las páginas 120 y 126.

 **Lema 6.5.** Si $AC = I$, entonces $CA = I$ y A es invertible con $C = A^{-1}$.

Demostración. Dado $AC = I$ transponemos $(AC)^\top = I \iff C^\top A^\top = I$, por teorema 2.4 d). Luego, por lema 5.1 existe D tal que $A^\top D = I$ y $D = C^\top$. Transponemos de nuevo y usamos la ley $(A^\top)^\top = A$ para concluir $I = (A^\top D)^\top = D^\top A = CA$. Con la definición sabemos que C es inversa de A , la unicidad sigue de lema 5.1. $\square$

 **Teorema 6.6.** Sean A y B matrices cuadradas, entonces

$$A \text{ \& B son invertibles} \iff \text{su producto } AB \text{ es invertible.}$$

Demostración. Si A y B son invertibles, entonces por teorema 5.3 AB es invertible.

Por otro lado, sea AB invertible. Entonces $(AB)^{-1}$ existe y por asociatividad del producto de matrices tenemos $(AB)(AB)^{-1} = I \iff A(B(AB)^{-1}) = I$. Es decir, A tiene inversa derecha $C = B(AB)^{-1}$ que por lema 6.5 es la inversa de A y A es invertible.

El mismo argumento, da $(AB)^{-1}AB = I \iff ((AB)^{-1}A)B = I$. Es decir, B tiene inversa izquierda $(AB)^{-1}A$ que por lema 5.1) es la inversa de B y B es invertible. $\square$

 **Teorema 6.7.** Sean $A, B \in \mathbb{R}^{n \times n}$ matrices, entonces $\det(AB) = \det(A) \det(B)$.

Demostración. Si AB no es invertible, entonces A o B no lo son (contra-positiva de teorema 6.6). Por lo tanto, con (6.2) tenemos

$$\det(A) \det(B) = 0 = \det(AB) \quad \therefore \quad \det(AB) = \det(A) \det(B).$$

Por otro lado, si AB es invertible, entonces teorema 6.6 justifica que A y B son invertibles. Por lema 5.4 podemos escribir a las matrices A y B como producto de matrices elementales $A = E_1^{-1} \cdots E_k^{-1}$ y $B = F_1^{-1} \cdots F_k^{-1}$. Más aún lema 5.5 justifica que inversas de matrices elementales son matrices elementales, es decir, $A = \tilde{E}_1 \cdots \tilde{E}_k$ y $B = \tilde{F}_1 \cdots \tilde{F}_k$. En conclusión, usando la igualdad (6.1) el determinante cumple:

$$\begin{aligned} \det(AB) &= \det(\tilde{E}_1 \cdots \tilde{E}_k B) = \det(\tilde{E}_1) \cdots \det(\tilde{E}_k) \cdot \det(B) \\ &= \det(\tilde{E}_1 \cdots \tilde{E}_k) \cdot \det(B) = \det(A) \cdot \det(B). \end{aligned} \quad \square$$

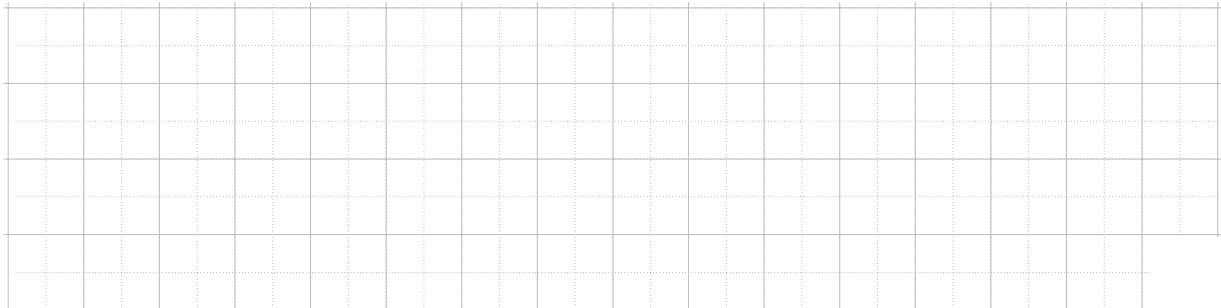
 **Corolario 6.8.** Si A es invertible, entonces $\det(A^{-1}) = \frac{1}{\det(A)}$.

Ejemplo: Calcula los determinantes de $A = \begin{pmatrix} -1 & 1 \\ 3 & 2 \end{pmatrix}$ y de su inversa A^{-1} .



6.2.4. Determinantes de transpuestas

La matriz elemental de tipo uno $I \xrightarrow{R_i \leftrightarrow R_j} E_{t1}$ es simétrica, es decir $E_{t1} = E_{t1}^\top$. Ejemplo:



6. Determinantes

La matriz elemental de tipo dos $I \xrightarrow{\gamma R_i} E_{t2}$ es simétrica, es decir $E_{t2} = E_{t2}^\top$. Ejemplo:

$$E_{t2} = \begin{pmatrix} \ddots & & & \\ & 1 & & \\ & & \gamma & \\ & & & 1 \\ & & & & \ddots \end{pmatrix} \leftarrow i$$

La transpuesta de matriz elemental de tipo tres $I \xrightarrow{\gamma R_i + R_j} E_{t3}$ es también una matriz elemental de tipo tres. Ejemplo:

$$E_{t3} = \begin{pmatrix} \ddots & & & \\ & 1 & & \\ & & \ddots & \\ & \gamma & & 1 \\ & & & & \ddots \end{pmatrix} \begin{matrix} \leftarrow i \\ \\ \leftarrow j \end{matrix} \quad \& \quad E_{t3}^\top = \begin{pmatrix} \ddots & & & \\ & 1 & & \gamma \\ & & \ddots & \\ & & & 1 \\ & & & & \ddots \end{pmatrix} \begin{matrix} \leftarrow i \\ \\ \leftarrow j \end{matrix}$$

Observe que $I \xrightarrow{\gamma R_j + R_i} E_{t3}^\top$, por lo cual $E_{t3}^\top$ es una matriz elemental de tipo tres.



Conclusión: En cada caso, la transpuesta de una matriz elemental es una matriz elemental del mismo tipo. Por lo tanto, paso 2 nos da sus determinantes:

$$\det(E_{t1}^\top) = -1 = \det(E_{t1}), \quad \det(E_{t2}^\top) = \gamma = \det(E_{t2}) \quad \& \quad \det(E_{t3}^\top) = 1 = \det(E_{t3}). \quad (6.3)$$



Teorema 6.9. Sea A una matriz cuadrada, entonces $\det(A) = \det(A^\top)$.

Demostración. Primero, A tiene renglones y columnas LI syss $A^\top$ tiene columnas y renglones LI. Entonces, por el teorema 5.2 A es invertible si y solo si $A^\top$ es invertible.

Caso 1: Si A no es invertible, entonces $A^\top$ no lo es y equivalencia (6.2) da $\det(A) = 0$ y $\det(A^\top) = 0$, por lo cual $\det(A) = \det(A^\top)$.

Caso 2: Si A es invertible, entonces por lema 5.4 podemos escribir a la matrices A como producto de matrices elementales $A = E_1^{-1} \cdots E_k^{-1}$. Más aún lema 5.5 justifica que inversas de matrices elementales son matrices elementales, es decir, $A = \tilde{E}_1 \cdots \tilde{E}_k$. Ahora, por el teorema 2.4(d) la transpuesta tiene la forma $A^\top = \tilde{E}_k^\top \cdots \tilde{E}_1^\top$. Finalmente, el paso 4 y las igualdades (6.3) nos permiten concluir

$$\det(A) = \prod_{i=1}^k \det(\tilde{E}_i) = \prod_{i=1}^k \det(\tilde{E}_i^\top) = \det(\tilde{E}_k^\top \cdots \tilde{E}_1^\top) = \det(A^\top).$$

□



Por $\det(A) = \det(A^\top)$ se puede concluir que la función $\det$ es multilineal en columnas también.

6.3. La función determinante

Hasta ahora solo vimos propiedades de funciones determinantes que satisfacen la definición 6.3.

- Falta ver que solo existe una única función que satisface la definición.
- Falta encontrarla forma explícita de esta función (para calcular).

a) La unicidad:



Teorema 6.10 (Unicidad). Si A es una matriz cuadrada de $n \times n$ y la función $\det$ satisface las tres propiedades de la definición 6.3:

- $\det$ es una función multilineal en los renglones,
- $\det(A) = 0$ cuando A tiene dos renglones consecutivos iguales,
- $\det(I) = 1$,

entonces la función $\det$ es única.

Demostración. Suponga que existen dos funciones $\det$ y $\widetilde{\det}$ que satisfacen las tres condiciones anteriores. Todas las propiedades en la sección 6.2 son válidas para ambas funciones, ya que fueron demostrados solo usando las 3 condiciones. Hacemos dos casos.

Caso 1: Si A no es invertible, entonces por equivalencia (6.2) $\det(A) = 0$ y $\widetilde{\det}(A) = 0$. Concluimos que $\det(A) = \widetilde{\det}(A)$ para cualquier matriz A que no es invertible.

Caso 2: Si A es invertible, entonces por lemas 5.4 y 5.5 podemos escribir A como producto de matrices elementales $A = \tilde{E}_1 \cdots \tilde{E}_k$. Por paso 2 (página 143) tenemos $\det(E_{ti}) = \widetilde{\det}(E_{ti})$ para los 3 tipos de matrices elementales. Luego paso 4 permite igualar $\det(A)$ con $\widetilde{\det}(A)$ como sigue:

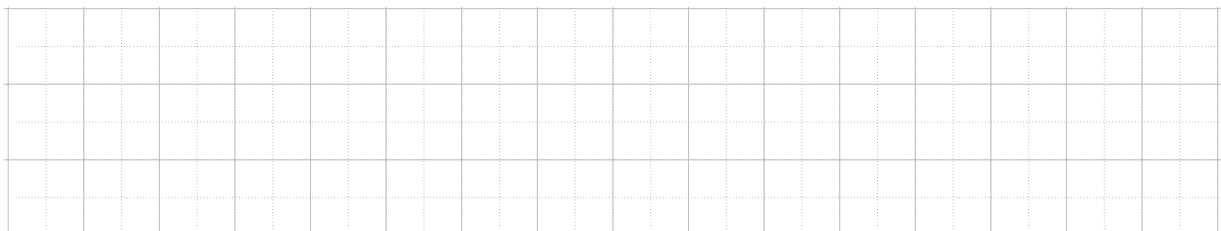
$$\begin{aligned} \det(A) &= \det(\tilde{E}_1 \cdots \tilde{E}_k) = \det(\tilde{E}_1) \cdot \det(\tilde{E}_2) \cdots \det(\tilde{E}_k) \\ &= \widetilde{\det}(\tilde{E}_1) \cdot \widetilde{\det}(\tilde{E}_2) \cdots \widetilde{\det}(\tilde{E}_k) = \widetilde{\det}(\tilde{E}_1 \cdots \tilde{E}_k) = \widetilde{\det}(A). \end{aligned}$$

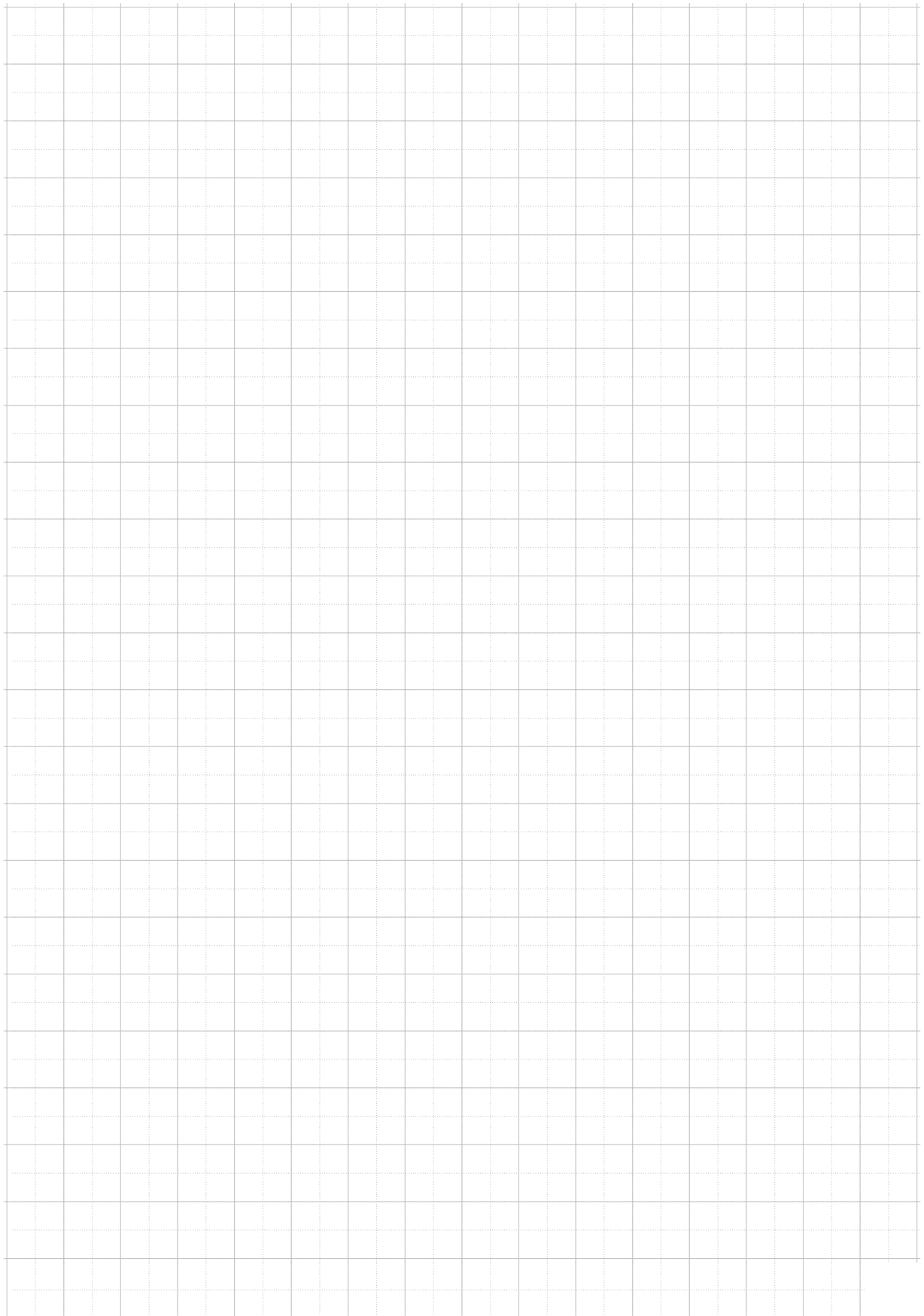
Esto ocurre para cada A invertible (con matrices elementales que dependen de A). $\square$

b) La fórmula explícita:

Intuición: Por teorema 6.9 tenemos $\det(A) = \det(A^\top)$ y se concluye que $\det$ también es multilineal en columnas. Es decir, si cambiamos columnas con el producto AE_{t1} entonces por $E_{t1} = E_{t1}^\top$ y paso 3 (página 143) se tiene

$$\det(AE_{t1}) = \det((AE_{t1})^\top) = \det(E_{t1}^\top A^\top) = \det(E_{t1} A^\top) = \det(E_{t1}) \det(A^\top) = -\det(A)$$



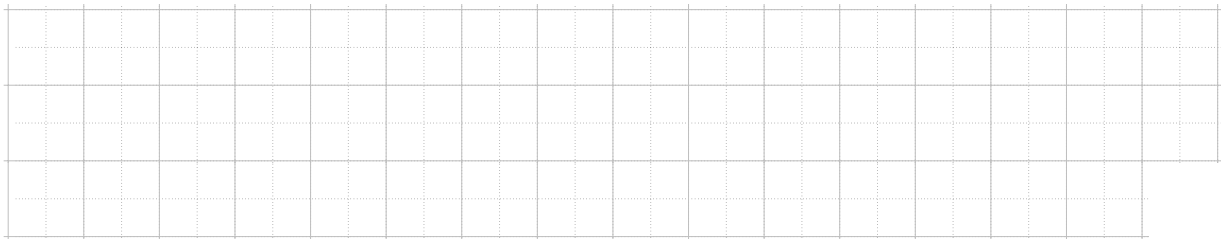


Con los argumentos del ejemplo en la página 141 se puede justificar que

$$\det \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & c & d \end{bmatrix} = ad - bc = \det \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

La igualdad anterior permite evaluar el resultado de la página anterior:

$$\begin{aligned} \det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} &= a \cdot \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{bmatrix} + b \cdot \det \begin{bmatrix} 0 & 1 & 0 \\ d & 0 & f \\ g & 0 & i \end{bmatrix} + c \cdot \det \begin{bmatrix} 0 & 0 & 1 \\ d & e & 0 \\ g & h & 0 \end{bmatrix} \\ &= a \cdot \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & e & f \\ 0 & h & i \end{bmatrix} - b \cdot \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & f \\ 0 & g & i \end{bmatrix} + c \cdot \det \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & e \\ 0 & g & h \end{bmatrix} \\ &= a \cdot \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \cdot \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \cdot \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}. \end{aligned} \quad (6.4)$$



6.3.1. La fórmula explícita para tamaño 3×3 :

El siguiente lema simplifica el cálculo visualizando el resultado:



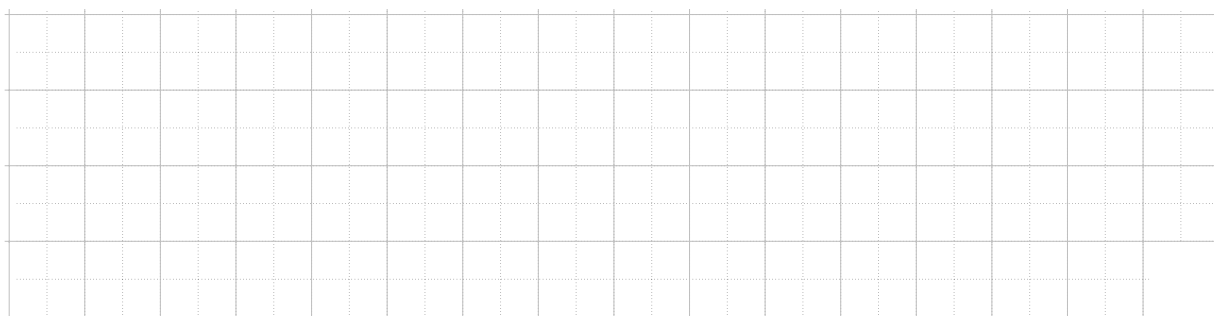
Lema 6.11 (Regla de Sarrus). Sea A una matriz 3×3 . Si $\tilde{A}$ es $[A | A_{:,1} | A_{:,2}]$, entonces el determinante de A es la suma de productos de las diagonales (de $\tilde{A}$) que bajan menos los productos de las diagonales (de $\tilde{A}$) que suben. Es decir, para

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \quad \tilde{A} = \begin{bmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{bmatrix}$$

tenemos

$$\det(A) = +aei + bfg + cdh - ceg - afh - bdi.$$

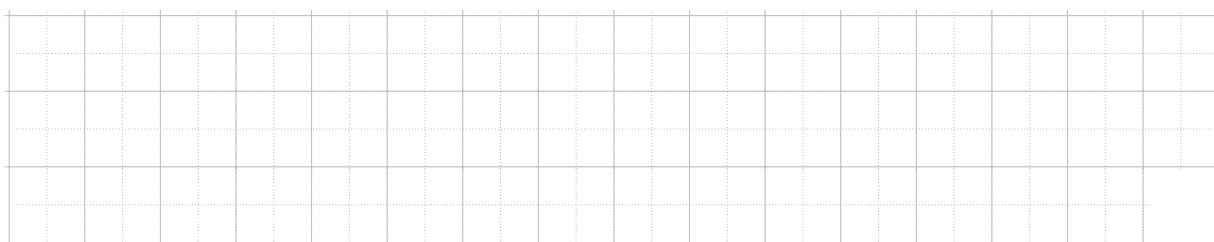
Ejemplo 1: Sea $A = \begin{pmatrix} 5 & 4 & 3 \\ 10 & 8 & 6 \\ 0 & 1 & 3 \end{pmatrix}$. Calcule $\det(A)$.



Ejemplo 2: Sea $A = \begin{pmatrix} -5 & 2 & 4 \\ 0 & 3 & -5 \\ 8 & -9 & 7 \end{pmatrix}$. Calcule $\det(A)$.



Ejemplo 3: Sea $A = \begin{pmatrix} 1 & 3 & 5 \\ 4 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Calcule $\det(A)$.



La regla de Sarrus se usa mucho en Cálculo.



La regla de Sarrus **solo** funciona para matrices de tamaño 3×3 .

6.3.2. La fórmula explícita para tamaño $n \times n$:

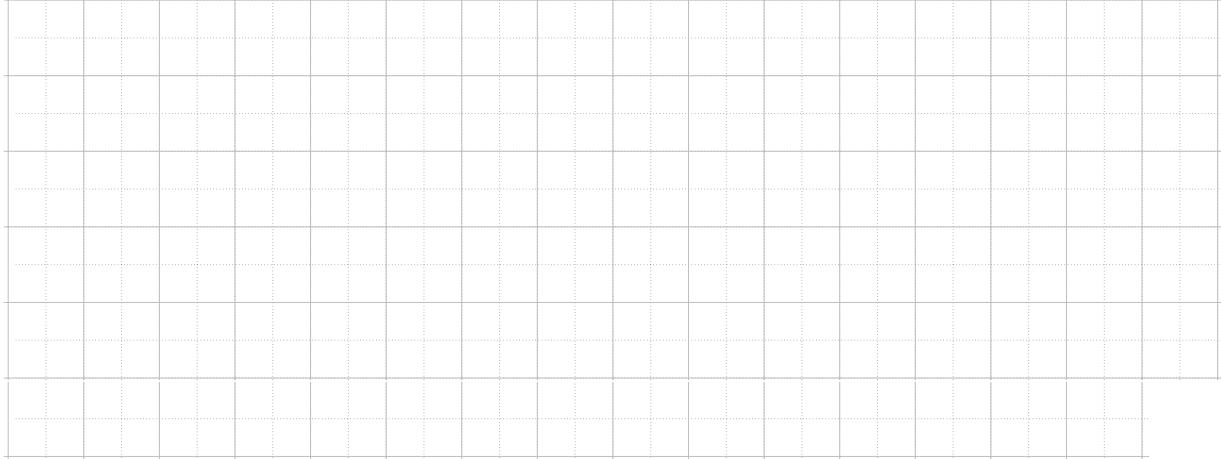
La ecuación (6.4) da la intuición para la evaluación recursiva. Es decir, la multi-linealidad permite escribir el determinante como suma de determinantes al separar un renglón en “renglones canónicos” con factores. Esos factores se pueden sacar (por multi-linealidad) y operaciones elementales de tipo 3 (que no cambian el determinante) muestran que cada determinante solo depende de elementos en la matriz que NO pertenecen al renglón ni a la columna del factor.

Eso se puede hacer empezando en cualquier renglón (o columna).

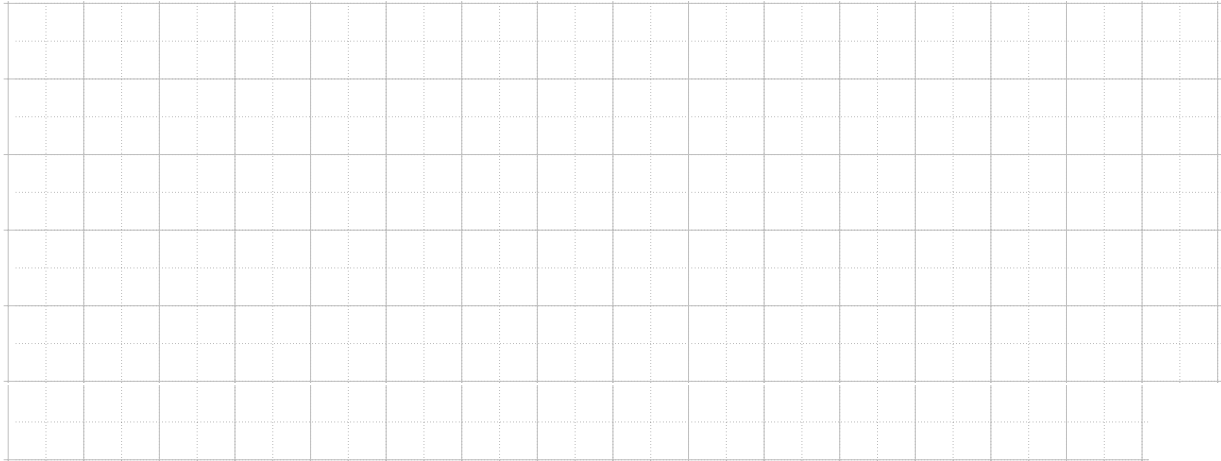


Definición 6.4. Sea A una matriz de $(n \times n)$ y sean $i, j \in \{1, \dots, n\}$ dos números naturales. Definimos $\tilde{A}_{ij}$ como la matriz de tamaño $(n-1) \times (n-1)$ que se obtiene de la matriz A eliminando el renglón i y la columna j .

Ejemplo I: Si $A = \begin{pmatrix} 2 & -3 & 1 \\ -3 & 5 & 7 \\ 1 & 7 & -1 \end{pmatrix}$. Calcule $\tilde{A}_{13}$, $\tilde{A}_{31}$ y $\tilde{A}_{22}$.



Ejemplo II: Si $B = \begin{pmatrix} 2 & 5 & 7 & \sqrt{3} \\ 3 & 2 & 0 & \pi \\ 1 & 8 & 6 & 4 \\ 0 & 0 & \sqrt{3} & \frac{1}{2} \end{pmatrix}$. Calcule $\tilde{B}_{34}$ y $\tilde{B}_{43}$.



En la ecuación (6.4) se ve que la matriz $\tilde{A}_{ij}$ (para el factor b) aparece “distribuida”. Por eso, se cambia el 1 del factor a_{ij} (b es este caso) a la esquina izquierda superior, es decir, se hacen $(i-1)$ cambios de renglones y $(j-1)$ cambios de columnas, todo con el fin de mantener los renglones y columnas de $\tilde{A}_{ij}$ en el orden como aparecen en A . Esos cambios cambian el signo del determinante por el factor $(-1)^{i-1+j-1} = (-1)^{i+j}$, por lo cual el determinante de $\tilde{A}_{ij}$ multiplicado con el factor a_{ij} recibe el factor $(-1)^{i+j}$.



Teorema 6.12 (Expansiones de P. S. Laplace). Sea $n \geq 3$ y sea A de tamaño $n \times n$. Entonces, para cualesquiera $i, j \in \{1, \dots, n\}$ el determinante tiene las siguientes expansiones:

$$\det(A) \stackrel{j \text{ fijo}}{=} \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}) \quad \text{“por columna } j\text{”,} \quad (6.5a)$$

$$\stackrel{i \text{ fijo}}{=} \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\tilde{A}_{ij}) \quad \text{“por renglón } i\text{”.} \quad (6.5b)$$

En palabras: Fijar i ó j y sumar determinantes más chicos.



Las expansiones (6.5) se usan para calcular determinantes de matrices $n \times n$ con $n \geq 3$.

Demostración. (Bosquejo)

Se define $D: M_{n \times n} \rightarrow \mathbb{R}$ por $D(A) := \det(A)$ cuando $n = 2$. Esto dará el inicio de inducción. Para $n \geq 3$ se define $D(A) := \sum_{i=1}^n (-1)^{i+j} a_{ij} D(\tilde{A}_{ij})$. Luego, con un paso de inducción se verifica que D satisface la definición 6.3, esto, por ejemplo está hecho en Proposición 8.33 del libro de [MR20].

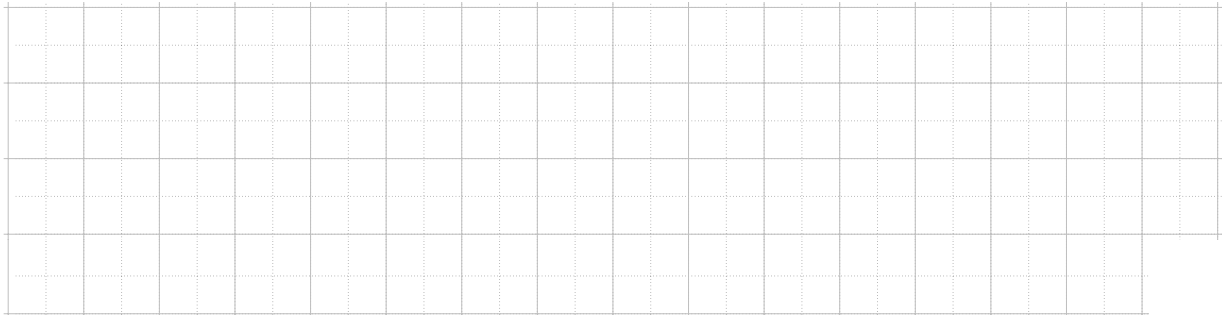
Después, por el teorema 6.10 de unicidad tenemos que $D(A) = \det(A)$.

Finalmente, la propiedad $\det(A) = \det(A^\top)$ da la expansión por renglón i . □

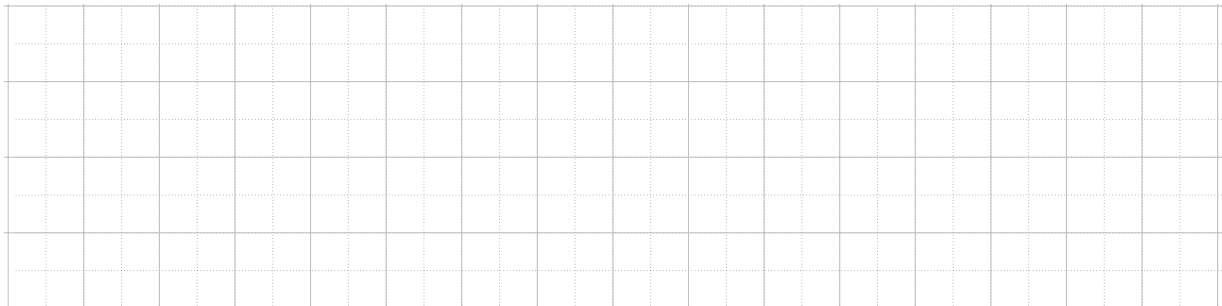
Ejemplo 1: Calcule el determinante de la siguiente matriz por el método de Sarrus & de Laplace:

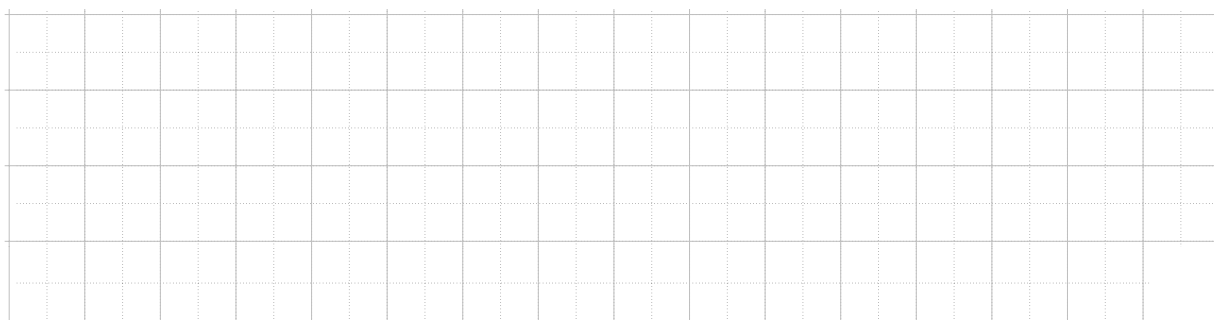
$$A = \begin{pmatrix} -2 & 4 & 3 \\ 0 & \sqrt{3} & 1/2 \\ 0 & -4/3 & 5 \end{pmatrix}$$

Por Sarrus:



Por Laplace: escogemos expandir la columna $j = 1$ ya que tiene muchos ceros:





Ejemplo 2: Determine si siguiente matriz es invertible: $A = \begin{pmatrix} 1 & 0 & 2 & -3 \\ 2 & -3 & -2 & 3 \\ -1 & 2 & 1 & 2 \\ 3 & 2 & 5 & 0 \end{pmatrix}$.



6. Determinantes

Ejemplo 3: Determine si siguiente matriz tiene columnas L.I.: $A = \begin{pmatrix} 8 & 4 & 0 \\ 8 & 4 & 1 \\ 1 & 3 & 0 \end{pmatrix}$.



Ejemplo 4: Calcule el determinante de la siguiente matriz: $A = \begin{pmatrix} 0 & 0 & \pi & 0 \\ 1 & 2 & 7 & 3 \\ 1 & 1 & -5 & 1 \\ 2 & 2 & 6 & 4 \end{pmatrix}$.



6.4. Resultados generales

Lema 6.13. El determinante de una matriz diagonal es el producto de sus entradas en la diagonal, i.e.

$$\det \begin{pmatrix} a_{11} & 0 & 0 & 0 \\ 0 & a_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & a_{nn} \end{pmatrix} = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn} = \prod_{i=1}^n a_{ii}.$$

Demostración. Ejercicio. Idea: usar que la función $\det$ es multi-lineal. $\square$

Lema 6.14. Sea A una matriz de $n \times n$. Si A es triangular superior o inferior, entonces

$$\det(A) = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.$$

Demostración. Hacemos una inducción sobre n .

Paso Base: Sabemos que para $n = 2$ el determinante es $\det \begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} = a_{11} \cdot a_{22} = \det \begin{pmatrix} a_{11} & a_{12} \\ 0 & a_{22} \end{pmatrix}$, lo cual confirma la igualdad.

Hipótesis inductiva: Supongamos que el resultado se cumple para un $n - 1 \geq 2$, es decir, si $A \in \mathbb{R}^{(n-1) \times (n-1)}$ es triangular superior o inferior, entonces $\det(A) = a_{11} \cdot a_{22} \cdot \dots \cdot a_{n-1}$ es cierto.

Paso inducción: Aumentamos $n - 1$ a n . Entonces, A tiene una de las formas

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 & 0 \\ a_{21} & a_{22} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n-1} & 0 \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n-1} & a_{nn} \end{pmatrix} \quad \text{o} \quad A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n-1} & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & \cdots & 0 & 0 & a_{n,n} \end{pmatrix}.$$

La expansión de Laplace (teorema 6.12) permite expandir el $\det(A)$ a lo largo de la última columna o último renglón. Dado que solamente la entrada a_{nn} puede ser distinto de cero, tenemos

$$\det(A) = (-1)^{n+n} a_{nn} \det(\tilde{A}_{nn}) \stackrel{\text{H. ind.}}{=} a_{nn} (a_{11} \cdot a_{22} \cdot \dots \cdot a_{n-1,n-1}) = \prod_{i=1}^n a_{ii}.$$

Lo cual termina la prueba. $\square$

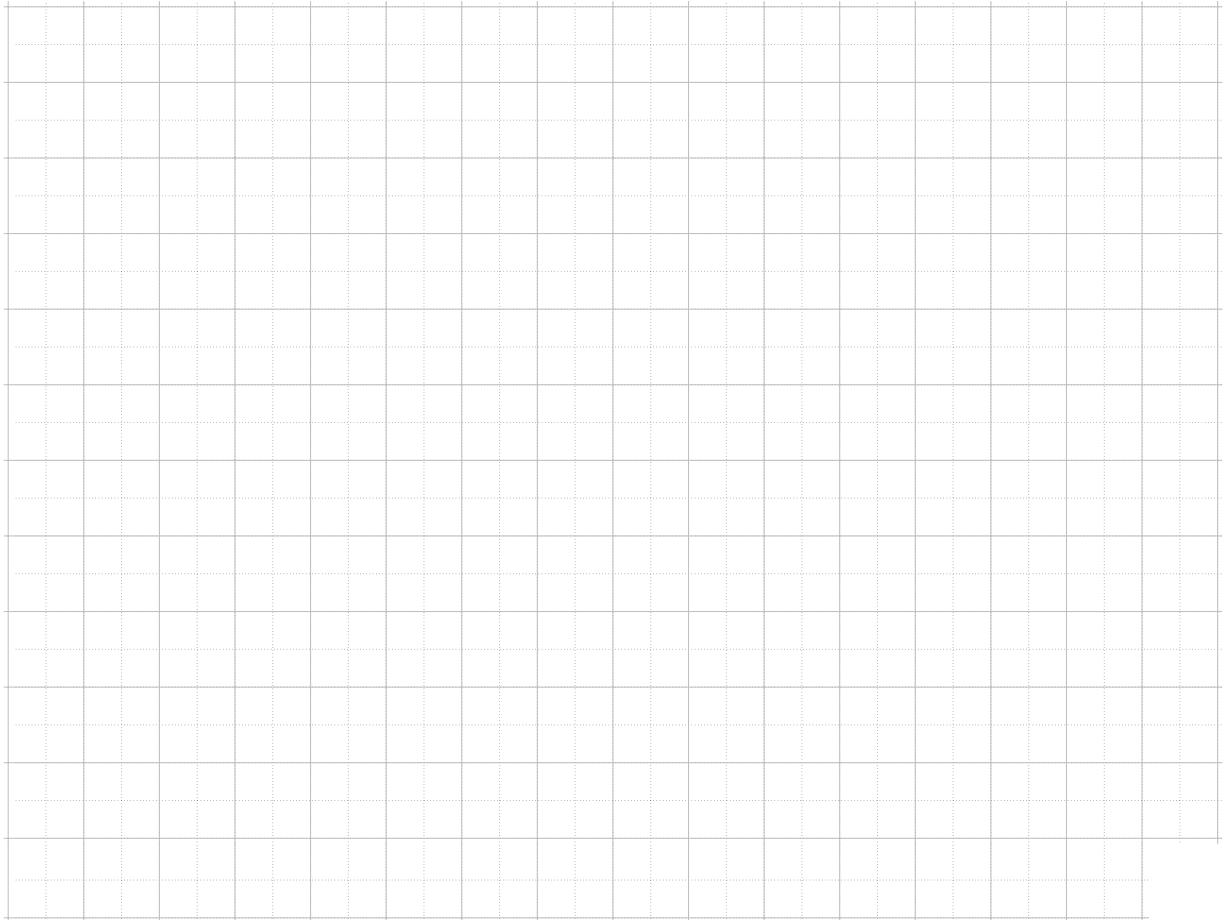
Ejemplo: Calcule el determinante de la matriz $A = \begin{pmatrix} 7 & 8 & 1 \\ 0 & \pi & -666 \\ 0 & 0 & 9 \end{pmatrix}$.



6. Determinantes

Ejemplo: Sean los vectores $\vec{v}_1 = \begin{pmatrix} 0 \\ a-3 \\ a^2-9 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} a-2 \\ a \\ \pi \end{pmatrix}$ & $\vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}$. Determine los valores de a para que los vectores sean linealmente independientes.

Solución: Los vectores son L.I., si son las columnas de una matriz invertible, es decir, si la matriz tiene inversa lo cual equivale a $\det(A) \neq 0$.



Un resumen más: El siguiente teorema reemplaza al teorema 5.2.



Teorema 6.15 (Segundo “teorema” de la inversa). Sea $A \in \mathbb{R}^{n \times n}$ una matriz cuadrada.

Los siguientes enunciados son equivalentes:

1. A es invertible / A tiene inversa.
2. $A\vec{x} = \vec{b}$ tiene una única solución para todo $\vec{b} \in \mathbb{R}^n$.
3. $A\vec{x} = \vec{0}$ tiene solución única (la trivial).
4. La FER de A es la matriz identidad.
5. Las columnas de A son linealmente independientes.
6. Los renglones de A son linealmente independientes.
7. $A^\top$ es invertible.
8. $\det(A) \neq 0$.

El determinante de matrices triangulares se puede generalizar.



Lema 6.16. Suponga que tenemos la siguiente matriz $n \times n$ “triangular en bloques”, i.e.

$$\mathcal{M} := \begin{bmatrix} A & B \\ \mathcal{O} & C \end{bmatrix}$$

con $A \in M_{p \times p}$, $B \in M_{p \times (n-p)}$, $C \in M_{(n-p) \times (n-p)}$ & $\mathcal{O} \in M_{(n-p) \times p}$ la matriz de ceros. Entonces,

$$\det(\mathcal{M}) = \det(A) \det(C).$$

Demostración. Notar que

$$\mathcal{M} = \begin{bmatrix} I_{p \times p} & \mathcal{O}_{p \times (n-p)} \\ \mathcal{O}_{(n-p) \times p} & C \end{bmatrix} \begin{bmatrix} A & B \\ \mathcal{O} & I_{(n-p) \times (n-p)} \end{bmatrix}$$

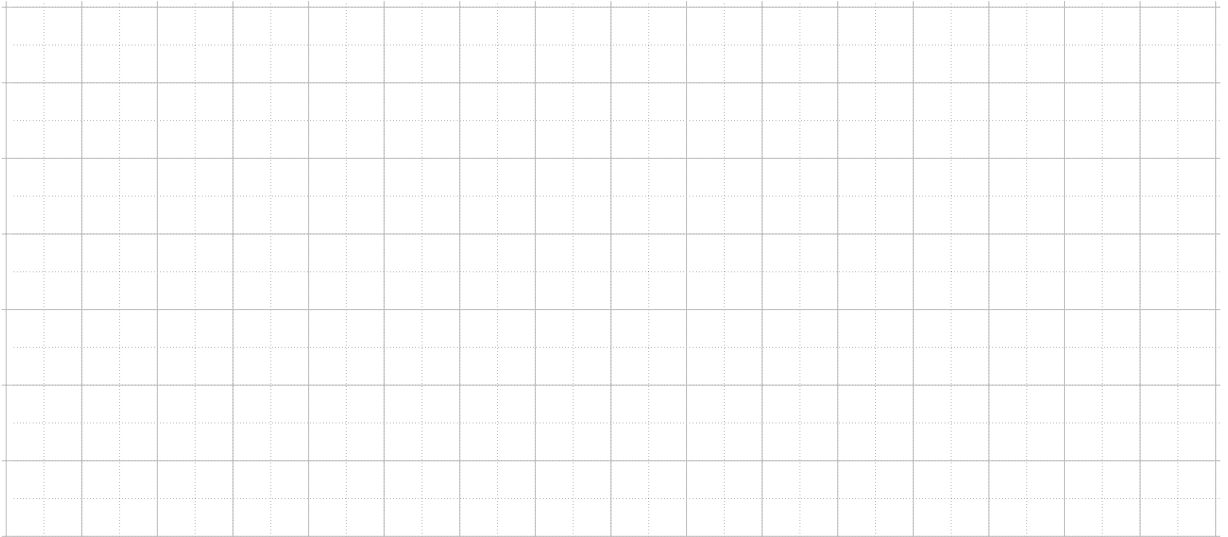
Por teorema 6.7 $\det(\mathcal{M})$ es producto de los determinantes de los dos matrices. Más aún los determinantes de los dos matrices se simplifican a $\det(C)$ y $\det(A)$, cuando se usa la expansión de Laplace en las filas de las identidades. En resumen llegamos a:

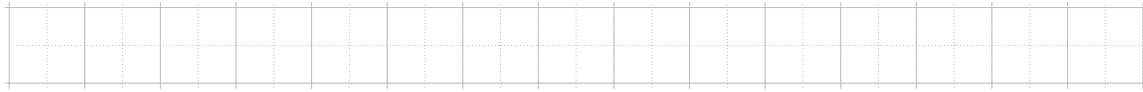
$$\det(\mathcal{M}) = \det \begin{pmatrix} I_{p \times p} & \mathcal{O}_{p \times (n-p)} \\ \mathcal{O}_{(n-p) \times p} & C \end{pmatrix} \det \begin{pmatrix} A & B \\ \mathcal{O} & I_{(n-p) \times (n-p)} \end{pmatrix} = \det(C) \det(A).$$

□

Ejemplo: Calcular los determinantes de las siguientes matrices:

$$M = \begin{pmatrix} 1 & -1 & \pi & 2 \\ -3 & 5 & -7 & 66 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 5 & -7 & 66 \\ 1 & 0 & -2 & \pi \\ 0 & 0 & 0 & 2 \end{pmatrix}$$





6.5. La regla de Cramer

Esa regla es para resolver sistemas lineales con matrices cuadradas e invertibles, es decir, con solución única. Es popular para resolver sistemas chicos, pero no es eficiente para matrices grandes. La regla de Cramer se usa en Cálculo diferencial para demostrar el *teorema de la función implícita*.



Teorema 6.17 (Regla de Cramer). Sean A una matriz invertible de $n \times n$ y $\vec{b} \in \mathbb{R}^n$ un vector (lado derecho) del sistema lineal $A\vec{x} = \vec{b}$. Entonces, cada entrada $x_1, \dots, x_k, \dots, x_n$ de la solución $\vec{x}^*$ es cociente de dos determinantes

$$x_k = \frac{\det(A^{k \leftrightarrow \vec{b}})}{\det(A)}$$

donde $A^{k \leftrightarrow \vec{b}}$ es la matriz $\begin{bmatrix} A_{:1} & \cdots & A_{:k-1} & \vec{b} & A_{:k+1} & \cdots & A_{:n} \end{bmatrix}$ depende de k .

Demostración. Sea $\{\vec{e}_1, \dots, \vec{e}_n\}$ la base estándar (canónica) de $\mathbb{R}^n$. Suponga que $\vec{x}^*$ es solución de $A\vec{x} = \vec{b}$. Formamos el siguiente producto

$$\begin{aligned} A \begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_{k-1} & \vec{x}^* & \vec{e}_{k+1} & \cdots & \vec{e}_n \end{bmatrix} &= \begin{bmatrix} A\vec{e}_1 & \cdots & A\vec{e}_{k-1} & A\vec{x}^* & A\vec{e}_{k+1} & \cdots & A\vec{e}_n \end{bmatrix} \\ &= \begin{bmatrix} A_{:1} & \cdots & A_{:k-1} & \vec{b} & A_{:k+1} & \cdots & A_{:n} \end{bmatrix} = A^{k \leftrightarrow \vec{b}}. \end{aligned}$$

Entonces, por el resultado $\det(AB) = \det(A)\det(B)$ tenemos

$$\det(A) \det \begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_{k-1} & \vec{x}^* & \vec{e}_{k+1} & \cdots & \vec{e}_n \end{bmatrix} = \det(A^{k \leftrightarrow \vec{b}}).$$

En determinante “en medio” se evalúa la lo largo del k -ésimo renglón con la expansión de Laplace lo cual nos lleva a

$$\det \begin{bmatrix} \vec{e}_1 & \cdots & \vec{e}_{k-1} & \vec{x}^* & \vec{e}_{k+1} & \cdots & \vec{e}_n \end{bmatrix} = x_k \det(I_{n-1}) = x_k.$$

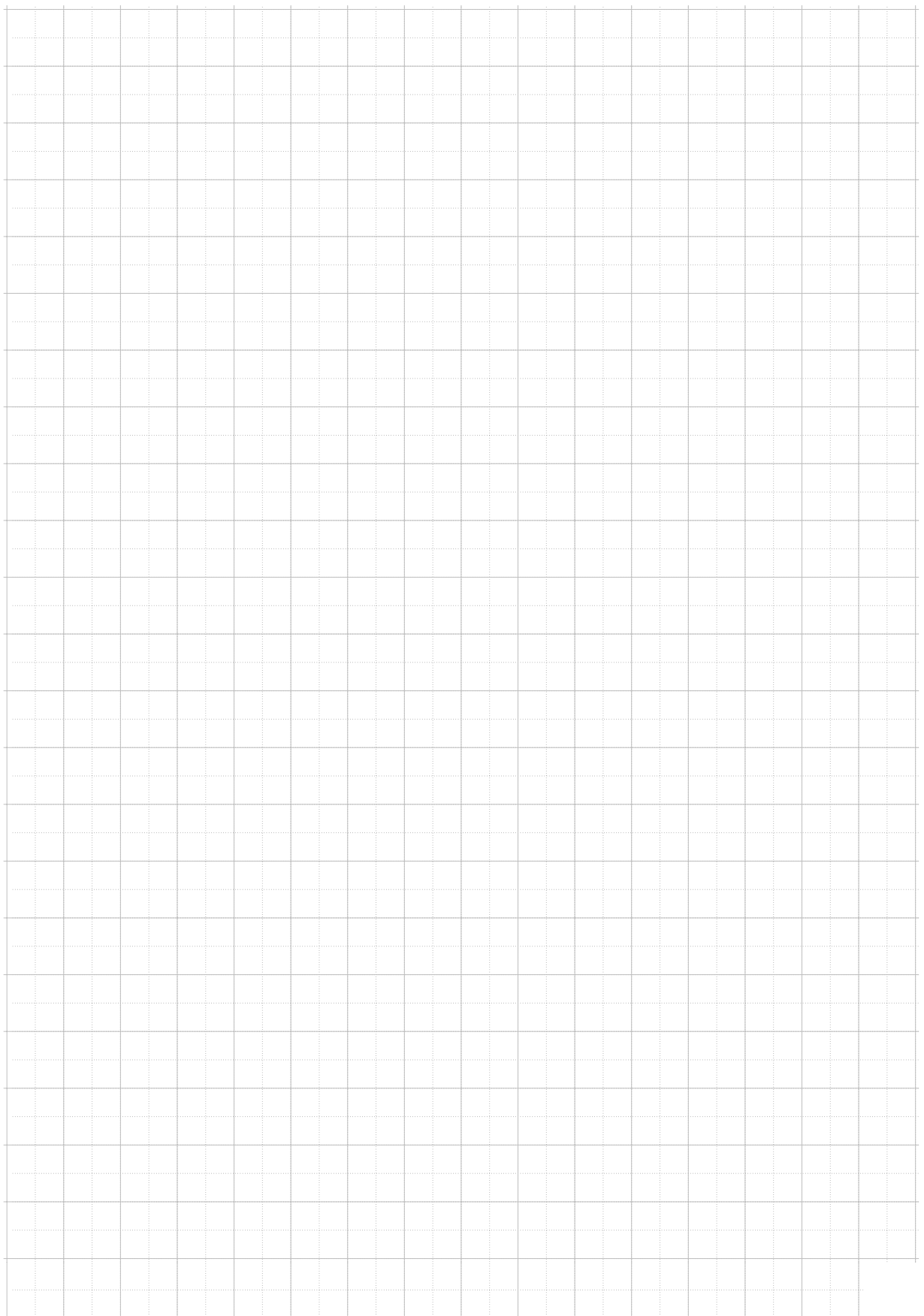
Entonces, la igualdad previa se reduce a

$$\det(A)x_k = \det(A^{k \leftrightarrow \vec{b}}) \quad \text{o bien} \quad x_k = \frac{\det(A^{k \leftrightarrow \vec{b}})}{\det(A)}$$

ya que $\det(A) \neq 0$ puesto que A es invertible. □

Ejemplo: Resuelve el siguiente sistema con la regla de Cramer.

$$\begin{pmatrix} 2 & 4 & 6 \\ 2 & 5 & 6 \\ 3 & 1 & -2 \end{pmatrix} \vec{x} = \begin{pmatrix} 9 \\ 12 \\ 2 \end{pmatrix}$$



6.6. La matriz de cofactores (de adjuntos)

La regla de Cramer (teorema 6.17) permite encontrar la inversa de A . De hecho, la columna j de A^{-1} se obtiene tomando $\vec{b} = \vec{e}_j$ ya que $\vec{e}_j = I\vec{e}_j = AA^{-1}\vec{e}_j = A(A_{:,j}^{-1})$.



Corolario 6.18. La entrada en el renglón i y la columna j de A^{-1} es

$$[A^{-1}]_{ij} = \frac{\det(A^{i \leftrightarrow \vec{e}_j})}{\det(A)} = \frac{1}{\det(A)} (-1)^{i+j} \det(\tilde{A}_{ji}).$$

Demostración. Sea $\vec{x}_j$ la solución de $A\vec{x} = \vec{e}_j$, es decir, $\vec{x}_j$ es la j -ésima columna de A^{-1} . Por la regla de Cramer tenemos que la entrada i del vector $\vec{x}_j$ es:

$$[A^{-1}]_{ij} = x_{ij} = \frac{\det(A^{i \leftrightarrow \vec{e}_j})}{\det(A)}.$$

Solo falta usar la expansión de Laplace a lo largo de la i -ésima columna la cual tiene un solo uno en el renglón j . Por lo tanto:

$$\det(A^{i \leftrightarrow \vec{e}_j}) = \det \begin{bmatrix} A_{:,1} & \cdots & A_{:,i-1} & \vec{e}_j & A_{:,i+1} & \cdots & A_{:,n} \end{bmatrix} = (-1)^{i+j} \det(\tilde{A}_{ji}).$$

La última igualdad se debe a que las columnas de $A^{i \leftrightarrow \vec{e}_j}$ & A son iguales excepto en la columna i . Por lo cual, al quitar la columna i y el renglón j de ambas matrices, obtenemos $\tilde{A}_{ji}$. $\square$



Definición. Al determinante $\det(\tilde{A}_{ij})$ se le conoce como **menor ij de A** .

Al número $C_{ij} := (-1)^{i+j} \det(\tilde{A}_{ij})$ se le conoce como el **cofactor ij de A** .

La **expansión por menores o cofactores** a lo largo de la columna j de A es $\sum_{i=1}^n a_{ij} C_{ij}$.

La **expansión por menores o cofactores** a lo largo del renglón i de A es $\sum_{j=1}^n a_{ij} C_{ij}$.

Por el teorema 6.12 ambas expansiones dan el valor $\det(A)$.

La siguiente definición permite generalizar la formula de la inversa vista en teorema 6.1.



Definición. Sea A una matriz $n \times n$. La **matriz de adjuntos** de A la denotamos como $\text{adj}(A)$ se define por

$$[\text{adj}(A)]_{ij} := C_{ji} \quad \text{donde} \quad C_{ji} = (-1)^{i+j} \det(\tilde{A}_{ji}),$$

y es una matriz de cofactores.

Por el corolario 6.18, cuando $\det(A) \neq 0$ tenemos $\det(A)A^{-1} = \text{adj}(A)$ lo cual implica:

$$(\text{adj}(A)) \cdot A = A \cdot (\text{adj}(A)) = \det(A) \cdot I. \quad (6.6)$$



7. Valores & vectores propios y Diagonalización de Matrices

Recordamos el ejemplo de dos fabricantes que compiten por clientes (sección 1.6).

Ejemplo: Suponga que únicamente dos empresas rivales R, S fabrican cierto producto. Cada año, la empresa R conserva $1/4$ de sus clientes mientras que $3/4$ cambian a S . En el mismo lapso, S conserva $2/3$ de sus clientes mientras $1/3$ cambia a R .

Al comenzar por primera vez la fabricación del producto, R tiene $x_r = 3/5$ del mercado (el mercado es la cantidad total de los clientes), mientras S tiene los otros $x_s = 2/5$. Notar que eso son 60% y 40%.

Habíamos encontrado la siguiente relación: Si $\vec{x}_k$ representa la distribución de los clientes en el lapso de tiempo $k \in \mathbb{N}$, entonces $\vec{x}_1 = A\vec{x}_0$. Además, hemos encontrado un vector $\vec{x}_0$ tal que

$$A\vec{x}_0 = \vec{x}_0 \quad \text{donde} \quad A = \begin{pmatrix} 1/4 & 1/3 \\ 3/4 & 2/3 \end{pmatrix}.$$

Entonces, **a veces** multiplicar una matriz con un vector no cambia la dirección del vector.

Ejemplo: Sean $A = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ & $\vec{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Dibujar los imágenes de $\vec{u}$ & $\vec{v}$ bajo la multiplicación por A .



En esta sección estudiamos ecuaciones del tipo $A\vec{x} = \lambda\vec{x}$ con incógnitas $\vec{x} \in \mathbb{R}^n$ y $\lambda \in \mathbb{R}$.

Por ejemplo, para $\lambda = 1$ los vectores $\vec{x}$ son estados estacionarios de una cadena de Markov.

Si te interesan las aplicaciones en la naturaleza, entonces recomiendo ver el ejemplo que trata de la supervivencia de los búhos en el libro de David C. Lay, véase [Lay07, pag. 301].



Definición. Un **eigenvector (vector propio)** de una matriz A de $n \times n$ es un vector $\vec{x}$ **distinto del vector cero** tal que $A\vec{x} = \lambda\vec{x}$ para algún escalar λ .

Un escalar λ se llama **eigenvalor (valor propio)** de A si existe una solución **no-trivial** $\vec{x}$ (denominada *vector propio correspondiente a λ*) de

$$A\vec{x} = \lambda\vec{x}.$$

Ejemplo: (de Lay). Sean $A = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 6 \\ -5 \end{pmatrix}$ & $\vec{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$.

¿Son $\vec{u}, \vec{v}$ vectores propios de A ?



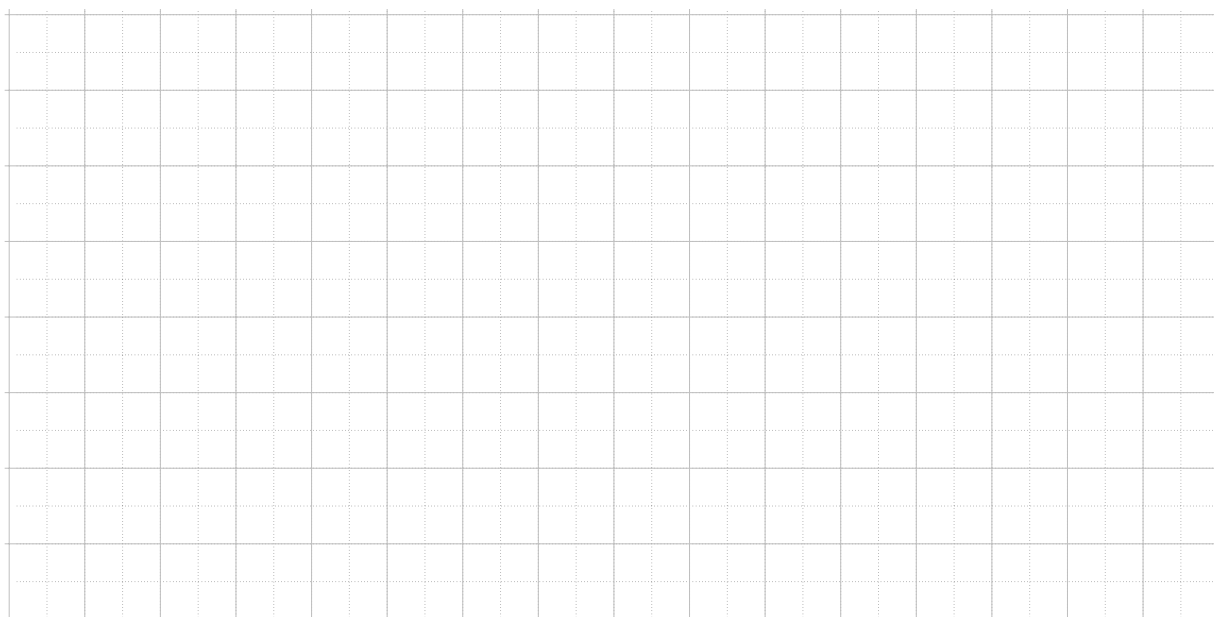
Ejemplo: (de Poole). Demuestre que 5 es un eigenvalor de $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ y determine todos los eigenvectores correspondientes a este eigenvalor.





 **Definición.** Sea A una matriz de $n \times n$ y sea λ un eigenvalor de A . El **Eigenespacio** de (λ, A) es el conjunto $\mathcal{E}_\lambda(A) = \{\vec{x}: A\vec{x} = \lambda\vec{x}\}$, es decir el conjunto que contiene al vector cero y todos los eigenvectores correspondientes a λ .

Ejemplo: Demuestre que 6 es un eigenvalor de $A = \begin{pmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{pmatrix}$ y encuentre una base para su eigenespacio $\mathcal{E}_6(A)$.





🍃 | *Observación:* El eigenspacio es un espacio nulo: $\mathcal{E}_\lambda(A) = \text{Null}(A - \lambda I)$.

La siguiente igualdad se cumple independiente de que si λ es o no es eigenvalor de A .



Lema 7.1. Sea A una matriz de tamaño $n \times n$ y λ un escalar. Entonces

$$\mathcal{E}_\lambda(A) := \{\vec{x} \in \mathbb{R}^n : A\vec{x} = \lambda\vec{x}\} = \text{Null}(A - \lambda I).$$

Demostración.



□

¿Cómo encontramos los eigenvalores?

Respuesta: Sea A una matriz de tamaño $n \times n$, entonces

λ es eigenvalor de $A \iff A\vec{x} = \lambda\vec{x}$ tiene una solución $\vec{x}$ no-trivial.

$\iff \text{Null}(A - \lambda I) \neq \{\vec{0}\}$, i.e. es no-trivial.

$\iff A - \lambda I$ no es invertible.

$\iff \det(A - \lambda I) = 0$.

$\iff A - \lambda I$ tiene columnas/renglones linealmente dependientes.

El criterio principal:



Definición. Dada una matriz A de tamaño $n \times n$ y una variable z se define el **polinomio característico** por

$$p_A(z) := \det(A - zI).$$



Corolario 7.2. Un escalar λ es eigenvalor de una matriz A ($n \times n$) si y solo si λ es **raíz** (o un “cero”) del polinomio característico, es decir

$$p_A(\lambda) = \det(A - \lambda I) = 0.$$

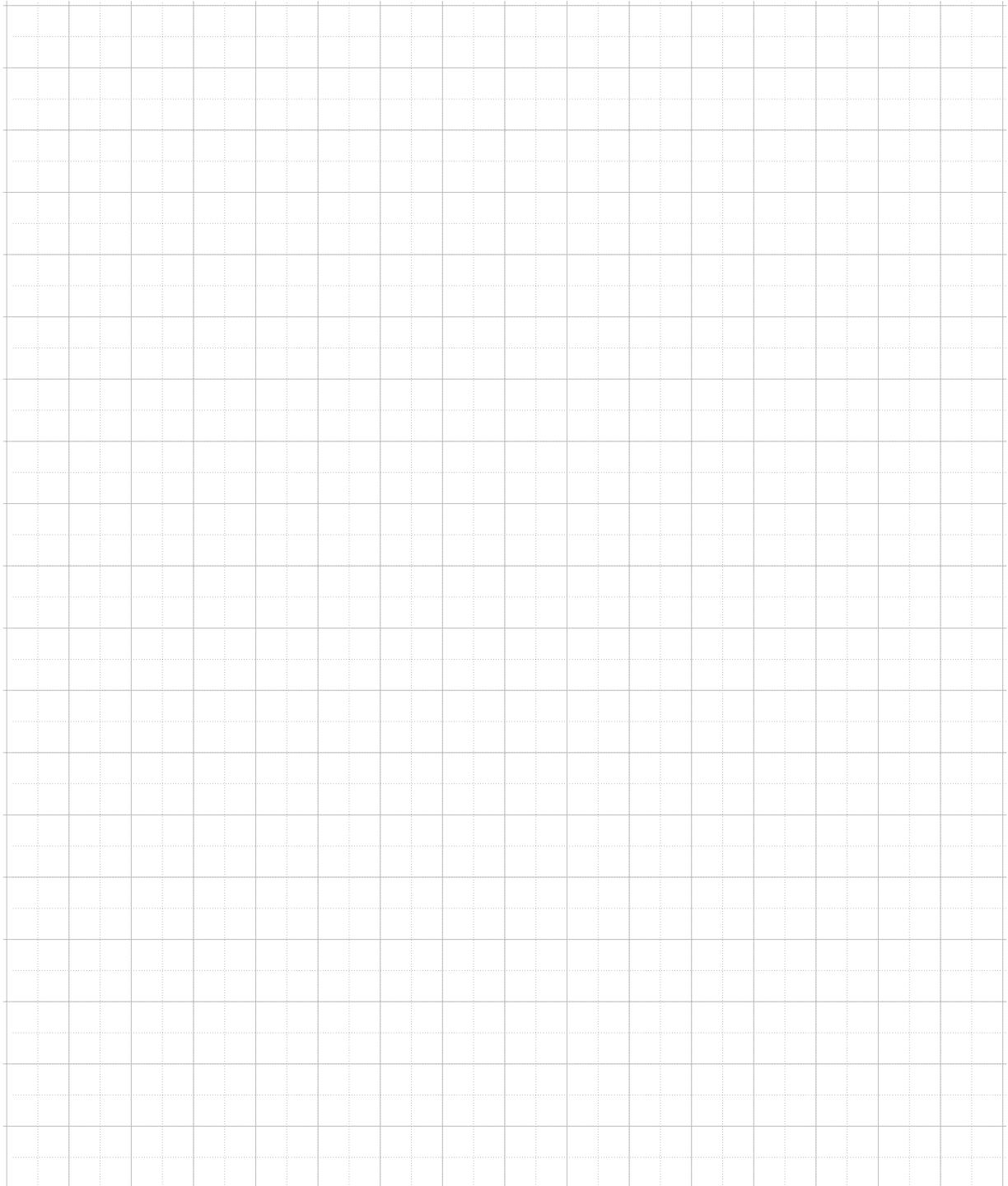
7. Valores y vectores propios y Diagonalización de Matrices



Hecho: En general se cumple que $p_A(z) = (-z)^n + [\dots]z + \det(A)$ donde $[\dots]$ representa un polinomio de grado $n - 2$ en z . El primer termino $(-z)^n$ se muestra con inducción (pero no lo hacemos), el último termino es evidente ya que $p_A(0) = \det(A)$.

Ejemplo:

Encuentre todos los eigenvalores y correspondientes eigenvectores de la matriz $A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$.





En el siguiente ejemplo usaremos el siguiente hecho.

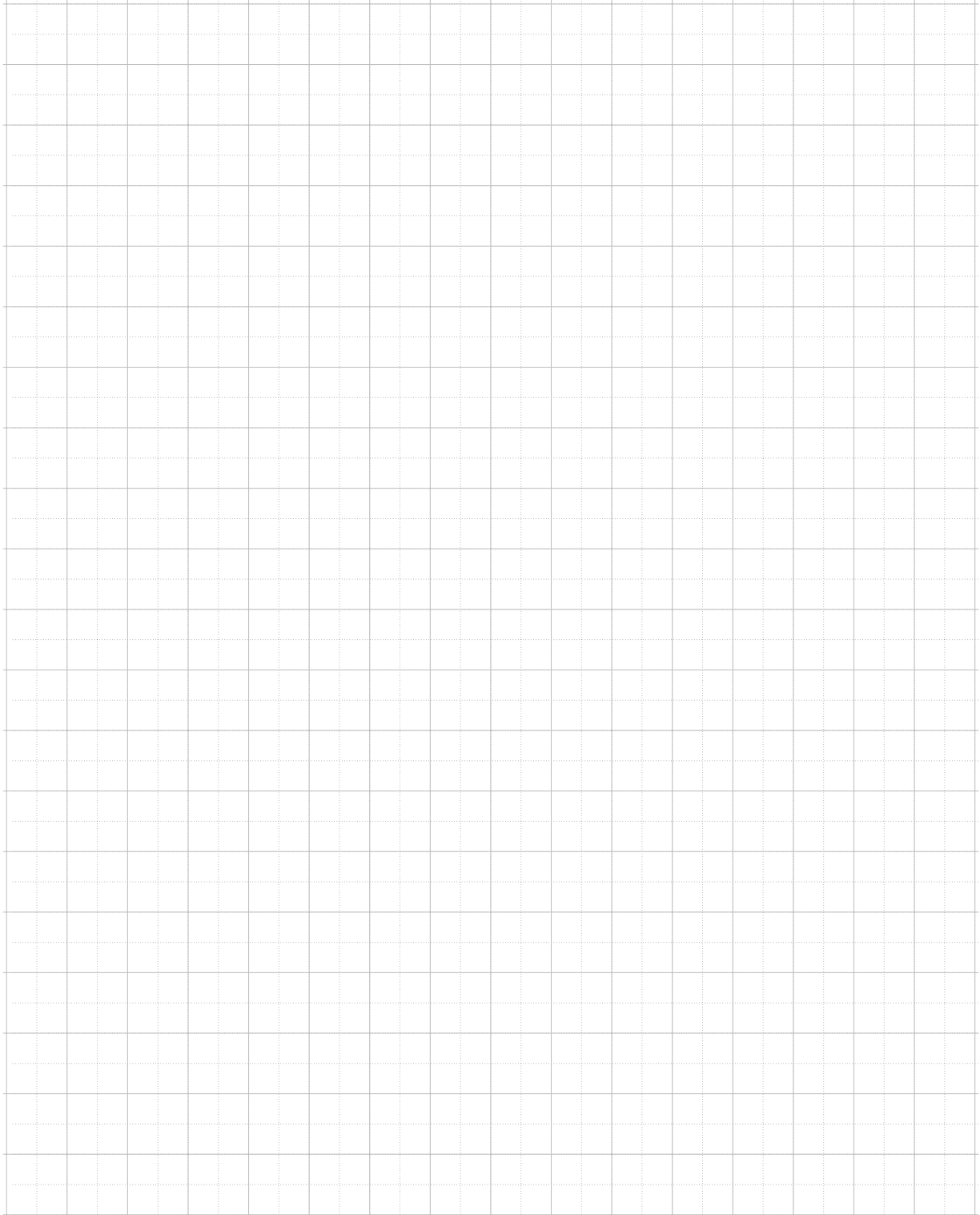


Si $p_A(z)$ tiene coeficientes enteros, entonces la forma de $p_A(z)$ y el *teorema de las raíces racionales* implican que cada raíz racional es entera y es divisor de $\det(A)$. Esto nos da candidatos para posibles raíces, pero no implica que todos los raíces son racionales.

7. Valores y vectores propios y Diagonalización de Matrices

Ejemplo: (importante) Sea A la matriz dada por $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{pmatrix}$.

1. Calcule los eigenvalores (resultado: $\lambda_1 = \lambda_2 = 1, \lambda_3 = 2$).
2. Calcule los eigenvectores y eigespacios.

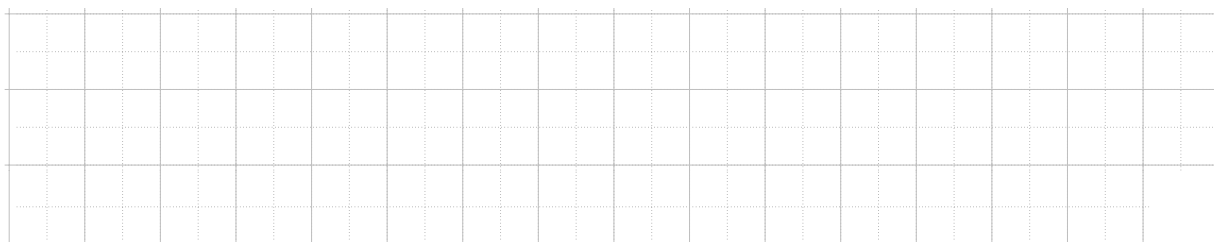




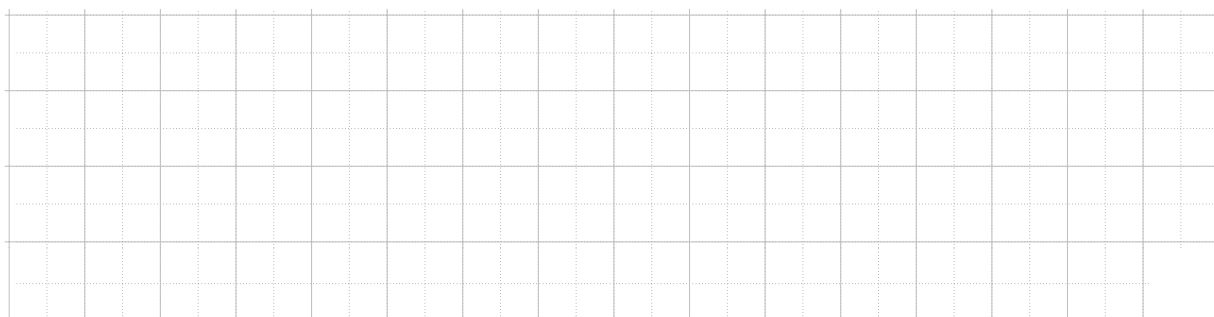



Definición 7.1. Si $p(x)$ es un polinomio y r es una raíz de $p(x)$, entonces definimos la **multiplicidad (algebraica) de r** como el mayor número natural m tal que $(x - r)^m$ divide al polinomio $p(x)$.

Ejemplo: En el ejemplo anterior el polinomio característico es $p_A(\lambda) = -(\lambda - 1)^2(\lambda - 2)$.
 Sus raíces y multiplicidades son:



Ejemplo: Sea $p(z) = (z - 7)^5(z + 3)(z - 42)^3$ entonces sus raíces y multiplicidades de ellas son:



7.1. Propiedades y resultados importantes



Definición 7.2. El **espectro** de A es el conjunto de todos los eigenvalores de A . Lo denotamos por

$$\text{spec}(A) = \{\lambda: \lambda \text{ es eigenvalor de } A\}.$$



Lema 7.3. Sea A una matriz $n \times n$ triangular inferior o superior. Entonces sus eigenvalores son las entradas de su diagonal, es decir $\text{spec}(A) = \{a_{11}, a_{22}, \dots, a_{nn}\}$.

Demostración. Sabemos que A es inferior o superior y restar una matriz identidad no lo cambia:

$$A = \begin{pmatrix} a_{11} & 0 & 0 \\ \vdots & \ddots & 0 \\ a_{n,1} & \cdots & a_{nn} \end{pmatrix} \quad \text{y} \quad A - \lambda I = \begin{pmatrix} a_{11} - \lambda & 0 & 0 \\ \vdots & \ddots & 0 \\ a_{n,1} & \cdots & a_{nn} - \lambda \end{pmatrix}$$

Por lo tanto, lema 6.14 dice que $\det(A - \lambda I)$ es el producto de sus entradas diagonales, es decir

$$p_A(\lambda) = \det(A - \lambda I) = \prod_{i=1}^n (a_{ii} - \lambda).$$

Este polinomio es de grado n y tiene a lo más n raíces.

Por corolario 7.2 sus raíces $\{a_{11}, a_{22}, \dots, a_{nn}\}$ son los eigenvalores de A . □

Ejemplo: Encuentre los eigenvalores de $A = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & \pi \end{pmatrix}$.



Ejemplo OjO: Si A es triangular pero $A - \lambda I$ no es triangular, entonces los eigenvalores no se pueden “leer” directamente. Encuentre los eigenvalores de $B = \begin{pmatrix} 0 & 0 & 9 \\ 0 & 4 & 0 \\ \pi & 0 & 0 \end{pmatrix}$.

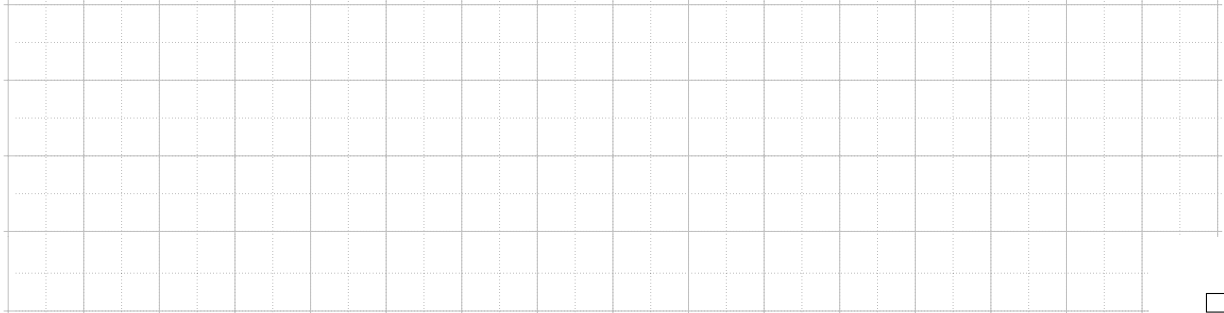


7. Valores y vectores propios y Diagonalización de Matrices

El siguiente resultado es importante en estadística, probabilidad, ... y dice que los eigenvalores de una matriz y su transpuesta son iguales.

 **Lema 7.4.** Sea A una matriz de tamaño $n \times n$. Entonces $\text{spec}(A) = \text{spec}(A^\top)$.

Demostración. Idea: Usar el teorema 6.9.



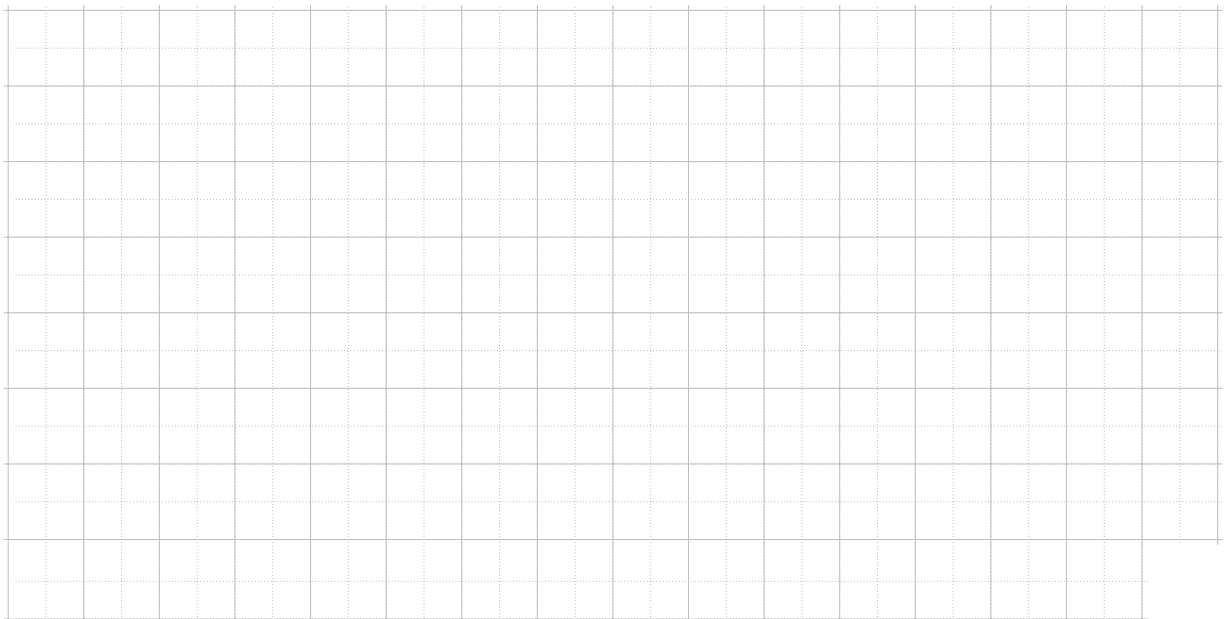
□

El siguiente resultado se usa para definir *valores singulares* (otra materia).

 **Lema 7.5.** Sean A una matriz $m \times n$, B otra matriz $n \times m$ y $\lambda \neq 0$ un escalar. Entonces,

$$\lambda \in \text{spec}(AB) \iff \lambda \in \text{spec}(BA).$$

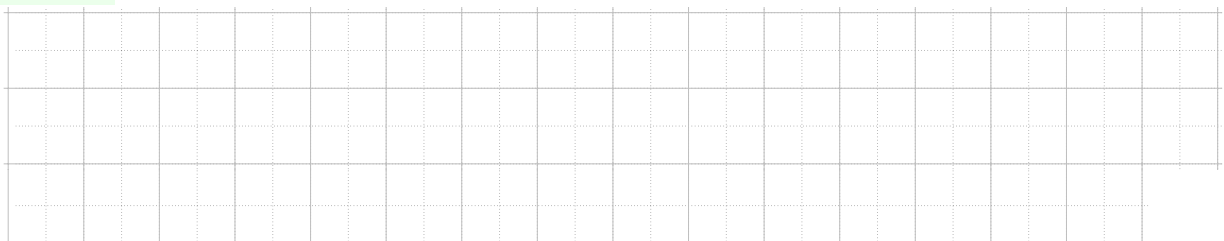
Demostración. Idea: Usar la definición de eigenvector y -valor.



La otra dirección queda como ejercicio.

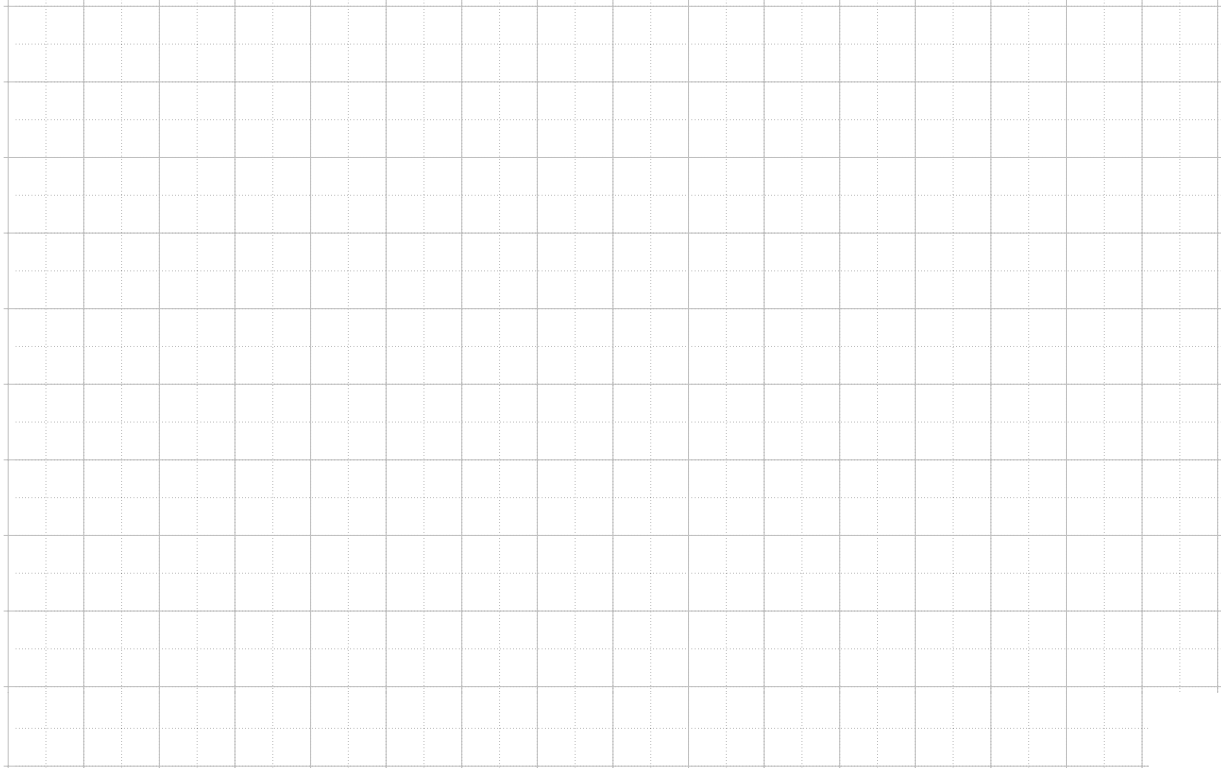
□

Ejemplos:



Ejemplo: Lema 7.4 dice que los eigenvalores de A y A^T son iguales.

Verifique que los eigenespacios son distintos (en general) tomando el ejemplo $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.



Definición. Sea A una matriz de $n \times n$ y λ uno de sus eigenvalores. Un vector $\vec{x} \in \mathbb{R}^{n \times 1} \setminus \{\vec{0}\}$ es **eigenvector por la izquierda** cuando $A^T \vec{x} = \lambda \vec{x}$.



Para su futuro: En estadística y probabilidad los eigenvectores por la izquierda se escriben en una de las 3 formas: $A^T \vec{x} = \lambda \vec{x} \iff \vec{x}^T A = \lambda \vec{x}^T \iff \vec{y} A = \lambda \vec{y}$ donde $\vec{y} = \vec{x}^T \in \mathbb{R}^{1 \times n}$, es decir $\vec{y}$ es $\vec{x}$ pero escrito como renglón.

7.1.1. Eigenvalores y matrices inversas



Lema 7.6. Una matriz A de $n \times n$ es invertible si y solo si 0 **no** es un eigenvalor de A .

Demostración. Mostramos la contra-positiva: A no es invertible si y solo si 0 es eigenvalor de A . Sabemos que A no es invertible si y solo si $0 = \det(A) \iff 0 = \det(A - 0I)$. Por el corolario 7.2 lo último equivale a que 0 es eigenvalor de A . $\square$



Lema 7.7. Sea A es una matriz invertible de tamaño $n \times n$ con eigenvalor λ y eigenvector correspondiente $\vec{x}$. Entonces $A\vec{x} = \lambda\vec{x} \iff A^{-1}\vec{x} = \lambda^{-1}\vec{x}$.

Es decir, λ^{-1} es eigenvalor de A^{-1} y el vector $\vec{x}$ es eigenvector de A y de A^{-1} .

Ejercicio: Demuestre lema 7.7

7.1.2. Eigenvalores de matrices triangulares en bloques



Lema 7.8. Suponga que tenemos la siguiente matriz $n \times n$ “triangular en bloques”, i.e.

$$\mathcal{M} := \begin{bmatrix} A & B \\ \mathcal{O} & C \end{bmatrix}$$

con $A \in M_{p \times p}$, $B \in M_{p \times (n-p)}$, $C \in M_{(n-p) \times (n-p)}$ & $\mathcal{O} \in M_{(n-p) \times p}$ la matriz de ceros. Entonces,

$$\det(\mathcal{M} - \lambda I) = \det(A - \lambda I) \det(C - \lambda I),$$

es decir $\text{spec}(\mathcal{M}) = \text{spec}(A) \cup \text{spec}(C)$.

Demostración. Solo tenemos que aplicar lema 6.16 ya que

$$\mathcal{M} - \lambda I = \begin{bmatrix} A - \lambda I & B \\ \mathcal{O} & C - \lambda I \end{bmatrix},$$

es decir, $\mathcal{M} - \lambda I$ tiene la misma estructura triangular por bloques que $\mathcal{M}$. □

Ejemplo (tipo examen): Calcular los eigenvalores de las siguientes matrices:

$$M = \begin{pmatrix} 1 & -1 & \pi & 2 \\ -3 & 5 & -7 & 66 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 5 & -7 & 66 \\ 1 & 0 & -2 & \pi \\ 0 & 0 & 0 & 2 \end{pmatrix}$$



7.1.3. Sistemas dinámicos discretos y cadenas de Markov

En el ejemplo de los búhos o de las cadenas de Markov, los vectores de cada momento en el tiempo son formados de la forma $\vec{x}_{\ell+1} = A\vec{x}_\ell$ para $\ell \in \{0, 1, 2, \dots\}$. Haciendo los primeros tres pasos

$$\vec{x}_0 \mapsto \vec{x}_1 = A\vec{x}_0 \mapsto \vec{x}_2 = A\vec{x}_1 = AA\vec{x}_0 = A^2\vec{x}_0 \mapsto \vec{x}_3 = A\vec{x}_2 = A^3\vec{x}_0.$$

concluimos que

$$\vec{x}_\ell = A^\ell \vec{x}_0 \quad \text{para} \quad \ell \in \mathbb{N}_+ = \{1, 2, 3, \dots\}. \quad (7.1)$$

El siguiente resultado describe como se comporte $A^\ell \vec{x}$ cuando $\vec{x}$ es eigenvector de A . Este resultado permite (bajo otras condiciones) determinar de antemano el límite de la sucesión $\{\vec{x}_\ell\}_\ell$.



Lema 7.9. Sea A una matriz $(n \times n)$ con eigenvalor λ y un eigenvector correspondiente $\vec{x}$. Entonces, para $\ell \in \mathbb{N}_+$ el valor λ^ℓ es eigenvalor de A^ℓ con eigenvector $\vec{x}$, i.e. $A^\ell \vec{x} = \lambda^\ell \vec{x}$.

Demostración. Por inducción.

La base: Sea $\ell = 1$ entonces $A^1 \vec{x} = A\vec{x} = \lambda \vec{x} = \lambda^1 \vec{x}$.

La hipótesis de inducción (H.d.I.): Supongamos que para $\ell \geq 1$ se tiene $A^\ell \vec{x} = \lambda^\ell \vec{x}$.

El paso de inducción: Aumentando $\ell \mapsto \ell + 1$ y usando la hipótesis y reglas concluimos:

$$A^{\ell+1} \vec{x} = AA^\ell \vec{x} = A(A^\ell \vec{x}) \stackrel{H.d.I.}{=} A(\lambda^\ell \vec{x}) = \lambda^\ell (A\vec{x}) = \lambda^\ell (\lambda \vec{x}) = \lambda^{\ell+1} \vec{x}. \quad \square$$

Ejemplo: Si $\vec{x}$ es eigenvector de A cuyo eigenvalor es λ . ¿Quién es $A^3 \vec{x}$?

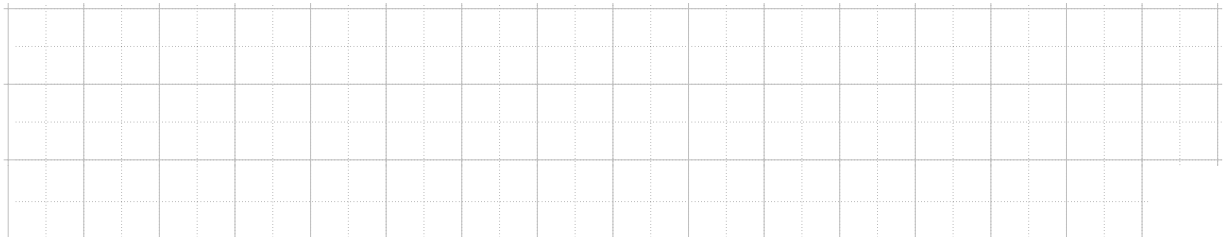


Ejemplo (importante): Recordemos el ejemplo de dos fabricantes que compiten por clientes. Sea

$$A = \begin{pmatrix} 1/4 & 1/3 \\ 3/4 & 2/3 \end{pmatrix} \quad \& \quad \vec{x}_0 = \begin{pmatrix} 3/5 \\ 2/5 \end{pmatrix}.$$

Si “la situación no cambia”, es decir, el número de fabricantes y la opinión publica, entonces podemos predecir como están distribuidos los clientes en 30 meses, es decir, deseamos conocer $\vec{x}_{30} = A^{30} \vec{x}_0$.

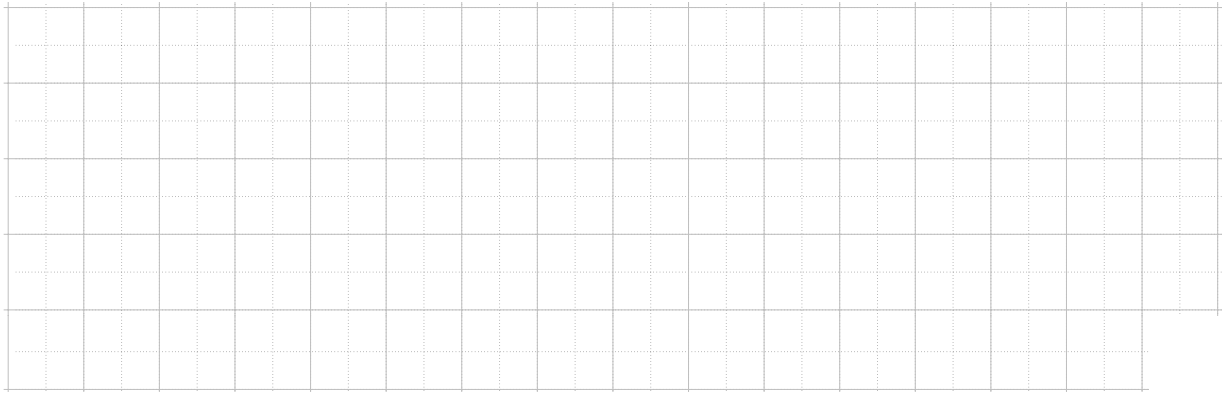
Caso 1: Si $\vec{x}_0$ **fuese** eigenvector de A , entonces



7. Valores y vectores propios y Diagonalización de Matrices

Caso 2: Podemos observar que $\vec{x}_0$ no es eigenvector de A . Pero yo soy terco y quiero aplicar la misma idea. Para poder hacer esto debo escribir $\vec{x}_0$ como combinación lineal de eigenvectores de A , ya que entonces





La idea anterior funciona cuando $\vec{x}_0$ se puede escribir como combinación lineal de eigenvectores. Esto requiere una base de eigenvectores de $\mathbb{R}^n$ y nos lleva al próximo objetivo.

7.2. Bases de eigenvectores

El ejemplo anterior genera el deseo de encontrar una base de $\mathbb{R}^n$ de eigenvectores de A .

Como $\mathbb{R}^n$ es el espacio completo solo tenemos que encontrar n vectores linealmente independientes (los cuales automáticamente generan a $\mathbb{R}^n$ por corolario 4.8).

Teorema 7.10. Sea A una matriz de $n \times n$ con $r \leq n$ eigenvalores **distintos** $\lambda_1, \lambda_2, \dots, \lambda_r$. Entonces, A tiene (por lo menos) r eigenvectores correspondientes $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_r$ que son L.I.

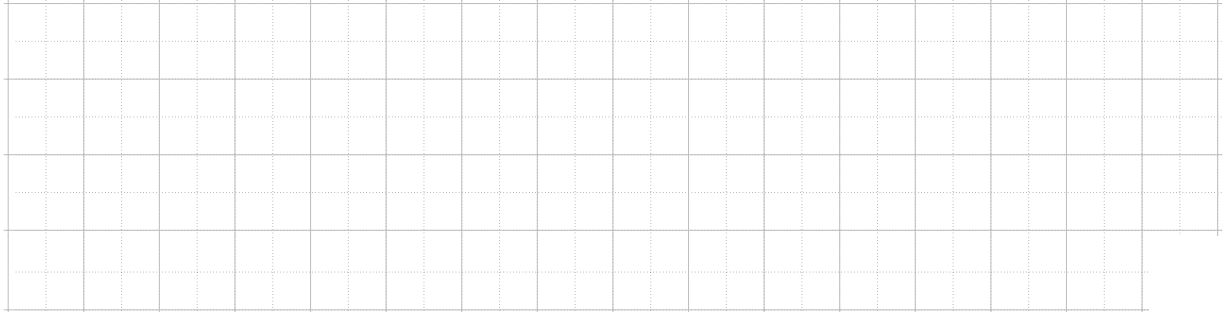
Demostración. Procedemos por inducción.

La base: Si $r = 1$ y λ_1 es eigenvalor entonces por definición existe un eigenvector $\vec{v}_1 \neq \vec{0}$, es decir $\vec{v}_1 \in \mathcal{E}_{\lambda_1}(A) \setminus \{\vec{0}\}$. Entonces $\{\vec{v}_1\}$ es lin. indep., ya que $\alpha_1 \vec{v}_1 = \vec{0}$ y $\vec{v}_1 \neq \vec{0}$ implica $\alpha_1 = 0$ por el hecho 3.2.4 en página 77 y el teorema 4.2.

Idea del paso para $r = 2$. Sean $\vec{v}_1 \in \mathcal{E}_{\lambda_1}(A) \setminus \{\vec{0}\}$ y $\vec{v}_2 \in \mathcal{E}_{\lambda_2}(A) \setminus \{\vec{0}\}$.

Por el teorema 4.2 tenemos que probar la implicación $\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 = \vec{0} \implies \alpha_1 = \alpha_2 = 0$.





Hipótesis de inducción: Suponga que el conjunto $\{\vec{v}_1, \dots, \vec{v}_{r-1}\}$ para $r \geq 2$ es L.I.

Paso de inducción: $r - 1 \mapsto r$

Sean $\vec{v}_i \in \mathcal{E}_{\lambda_i}(A) \setminus \{\vec{0}\}$ para $i \in \{1, \dots, r\}$ eigenvectores.

Por el teorema 4.2 tenemos que probar la implicación

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_r \vec{v}_r = \vec{0} \implies \alpha_1 = \alpha_2 = \dots = \alpha_r = 0.$$

Multiplicar la igualdad $\sum_i \alpha_i \vec{v}_i = \vec{0}$ por la izquierda con A da la igualdad:

$$\begin{aligned} A[\alpha_1 \vec{v}_1 + \dots + \alpha_{r-1} \vec{v}_{r-1} + \alpha_r \vec{v}_r] &= A\vec{0} = \vec{0} \\ \iff \alpha_1 A\vec{v}_1 + \dots + \alpha_{r-1} A\vec{v}_{r-1} + \alpha_r A\vec{v}_r &= \vec{0} \\ \iff \alpha_1 \lambda_1 \vec{v}_1 + \dots + \alpha_{r-1} \lambda_{r-1} \vec{v}_{r-1} + \alpha_r \lambda_r \vec{v}_r &= \vec{0} \end{aligned} \quad (\#)$$

Multiplicar la igualdad $\sum_i \alpha_i \vec{v}_i = \vec{0}$ con λ_r da la igualdad:

$$\alpha_1 \lambda_r \vec{v}_1 + \dots + \alpha_{r-1} \lambda_r \vec{v}_{r-1} + \alpha_r \lambda_r \vec{v}_r = \lambda_r \vec{0} = \vec{0}. \quad (\star)$$

Restando $(\star)$ de $(\#)$ se elimina $\alpha_r \lambda_r \vec{v}_r$ y se obtiene

$$\alpha_1 (\lambda_1 - \lambda_r) \vec{v}_1 + \alpha_2 (\lambda_2 - \lambda_r) \vec{v}_2 + \dots + \alpha_{r-1} (\lambda_{r-1} - \lambda_r) \vec{v}_{r-1} = \vec{0}$$

Por la hipótesis de inducción los vectores $\vec{v}_1, \dots, \vec{v}_{r-1}$ son lin. independientes y el teorema 4.2 implica que los coeficientes son cero, es decir, para $i \in \{1, 2, \dots, r-1\}$ tenemos $\alpha_i (\lambda_i - \lambda_r) = 0$. El teorema supone que los eigenvalores son diferentes, por lo cual

$$\underbrace{\alpha_i (\lambda_i - \lambda_r)}_{\neq 0} = 0 \iff \alpha_i = 0 \quad \text{para} \quad i \in \{1, 2, \dots, r-1\}.$$

Finalmente la combinación lineal se reduce, *i.e.* $\vec{0} = \sum_{i=1}^r \alpha_i \vec{v}_i = \alpha_r \vec{v}_r \iff \vec{0} = \alpha_r \vec{v}_r$ por lo cual $\alpha_r = 0$ ya que $\vec{v}_r \neq 0$ por ser eigenvector. Concluimos que $\alpha_1 = \alpha_2 = \dots = \alpha_r = 0$ y por el teorema 4.2 los vectores $\vec{v}_1, \dots, \vec{v}_r$ son lin. independientes. $\square$

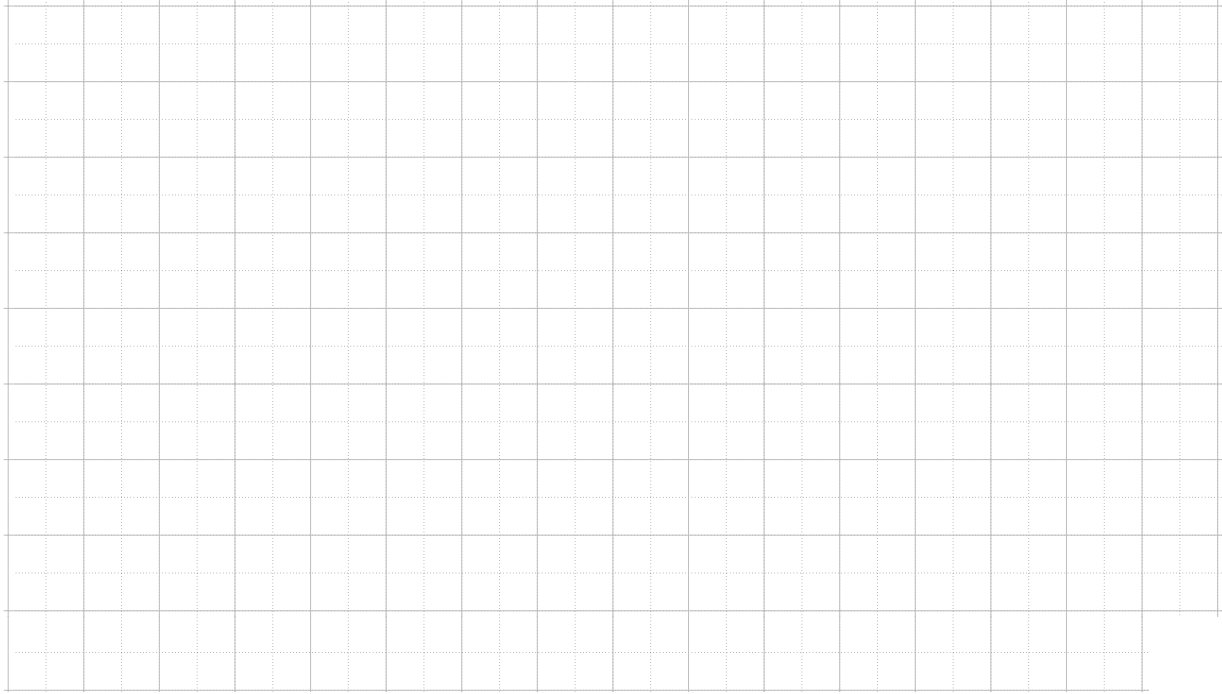


Interpretación importante: Sea A una matriz de $n \times n$. Si A tiene n eigenvalores distintos, entonces por el teorema 7.10 existe un conjunto de n eigenvectores L.I. que (por corolario 4.8) forma una base de $\mathbb{R}^n$ que consiste de eigenvectores de A .



Ejemplo 1: El teorema 7.10 dice lo mínimo que se puede (en general).

Calcule los eigenvalores y eigenespacios de la matriz $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.



La prueba del teorema 7.10 elige para cada λ_i solo un vector L.I. de $\mathcal{E}_{\lambda_i}(A) = \text{Null}(A - \lambda_i I)$. Cuando la dimensión de este espacio es más grande podemos escoger más vectores linealmente independientes. Esa dimensión tiene el siguiente nombre.

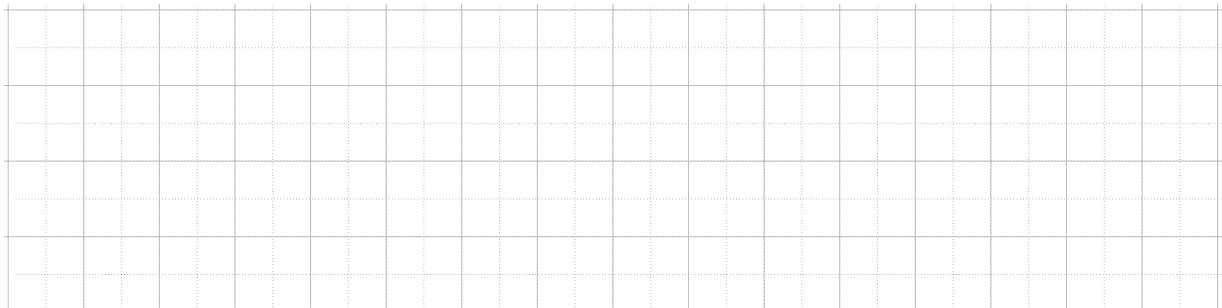


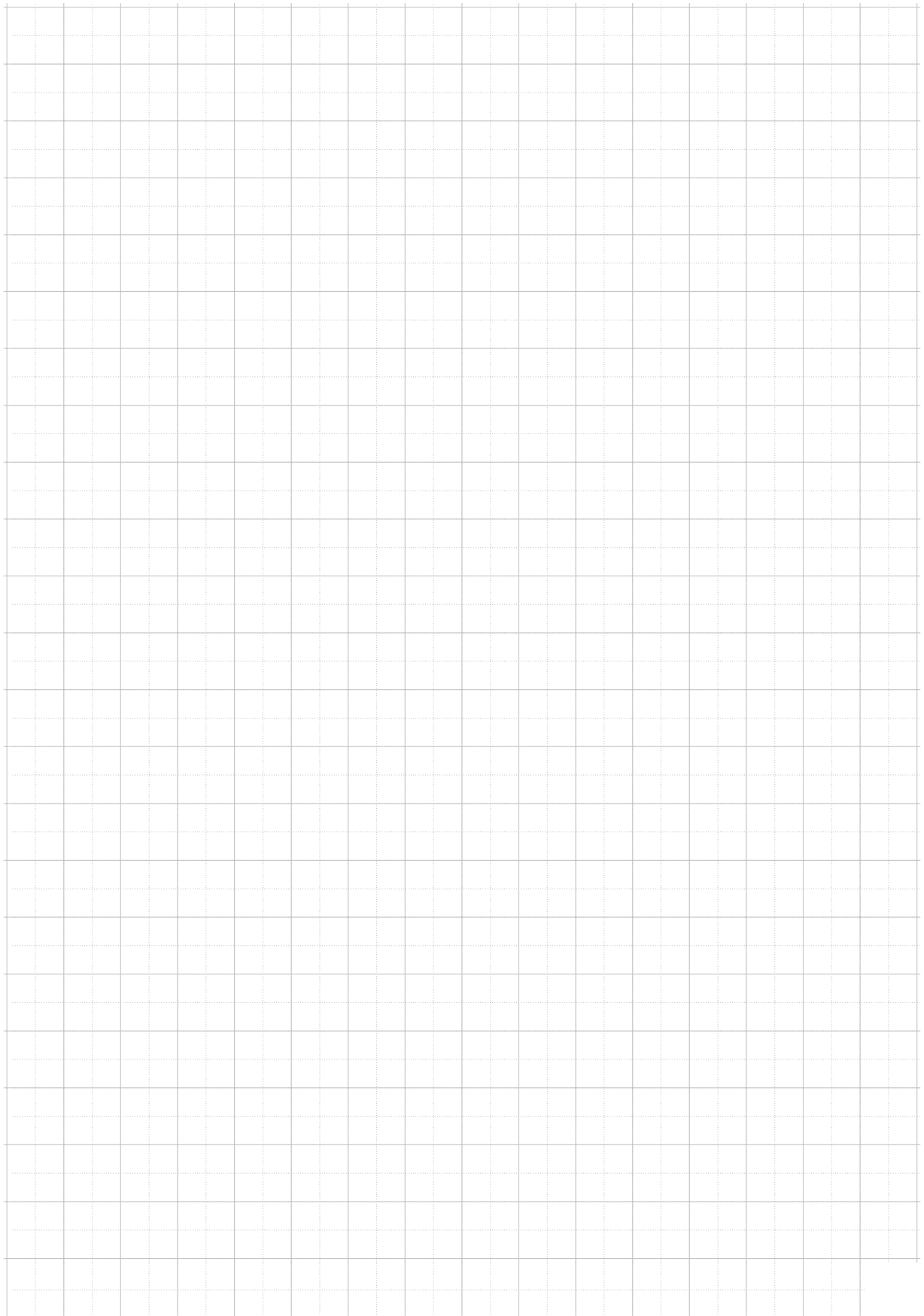
Definición. Sea A una matriz de $n \times n$ y λ un eigenvalor de A .

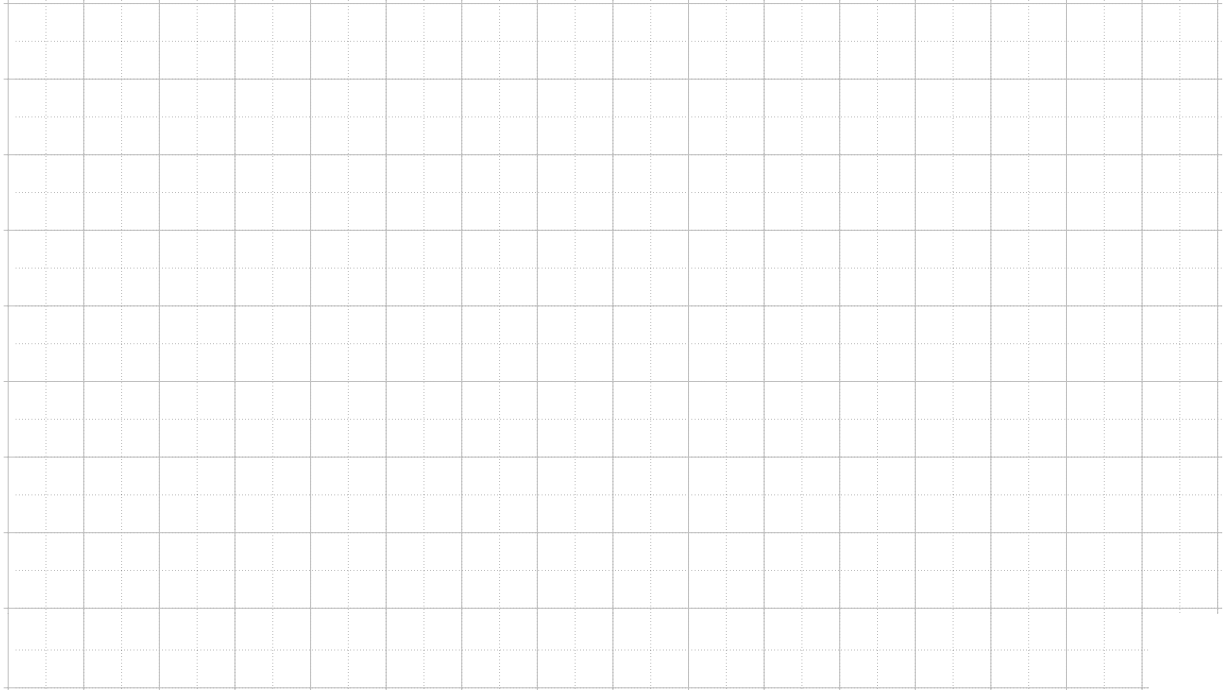
La **multiplicidad geométrica** de λ es la dimensión del eigenespacio $\mathcal{E}_{\lambda}(A) = \text{Null}(A - \lambda I)$.

Ejemplo 2: Existen matrices cuyos eigenvectores generan $\mathbb{R}^n$ y que tienen eigenvalores repetidos, es decir, con multiplicidad (algebraica) > 1 . Calcule las multiplicidades geométricas de los eigenvalores de las matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \& \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad \& \quad C = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & \pi \\ 1 & 0 & 1 \end{pmatrix}.$$







En los ejemplos anteriores vimos que las multiplicidades algebraicas dominan las geométricas:

$$\dim(\mathcal{E}_1(A)) = 3 \leq 3, \quad \dim(\mathcal{E}_0(B)) = 1 \leq 1, \quad \dim(\mathcal{E}_1(B)) = 2 \leq 2 \quad \& \quad \dim(\mathcal{E}_1(C)) = 1 < 2.$$



En general se cumple que:

$$1 \leq \dim(\mathcal{E}_\lambda(A)) = \text{“mult. geométrica”} \leq \text{“mult. algebraica”}.$$



Lema 7.11. Si λ_1, λ_2 son eigenvalores distintos de A , entonces $\mathcal{E}_{\lambda_1}(A) \cap \mathcal{E}_{\lambda_2}(A) = \{\vec{0}\}$.

Demostración. Sea $\vec{v} \in \mathcal{E}_{\lambda_1}(A) \cap \mathcal{E}_{\lambda_2}(A)$, entonces $A\vec{v} = \lambda_1\vec{v} = \lambda_2\vec{v} \implies (\lambda_1 - \lambda_2)\vec{v} = \vec{0}$. Usando $(\lambda_1 - \lambda_2) \neq 0$ y el hecho 3.2.4 (página 77) justifica $\vec{v} = \vec{0}$. $\square$

En los ejemplos tambien se pueden reconocer los siguientes resultados:



Corolario 7.12. Sea A una matriz de $n \times n$ y sean $\lambda_1, \lambda_2, \dots, \lambda_r$ eigenvalores distintos de A . Si $\vec{v}_i \in \mathcal{E}_{\lambda_i}(A)$ para $i \in \{1, 2, \dots, r\}$, entonces

$$\vec{v}_1 + \vec{v}_2 + \dots + \vec{v}_r = \vec{0} \implies \vec{v}_i = \vec{0} \quad \text{para } i \in \{1, 2, \dots, r\}.$$

Demostración. Usamos $\neg B$ y A para llegar a una contradicción.

Supongamos que algunos vectores $\vec{v}_i$ son distintos del vector cero. Entonces $\sum_{i=1}^r \vec{v}_i = \vec{0}$ es una combinación lineal no trivial de $\vec{0}$. Eso implica que los vectores $\{\vec{v}_1, \dots, \vec{v}_r\} \setminus \{\vec{0}\}$ son lin. dependientes, pero, los mismos vectores son eigenvectores para distintos eigenvalores y por teorema 7.10 son lin. independientes; una contradicción. Concluimos que $\vec{v}_i = \vec{0}$ para $i \in \{1, 2, \dots, r\}$. $\square$

Finalmente, el resultado general para matrices reales.

Proposición 7.13. Sea $A \in \mathbb{R}^{n \times n}$ una matriz cuadrada. Supongamos que A tiene $k \leq n$ eigenvalores **reales** distintos. Entonces $\mathbb{R}^n$ tiene una base que consisten de eigenvectores de A si y solo si

$$n = \dim(\mathcal{E}_{\lambda_1}(A)) + \dim(\mathcal{E}_{\lambda_2}(A)) + \dots + \dim(\mathcal{E}_{\lambda_k}(A)).$$

En tal caso, si $\mathcal{B}_i$ es base del eigenespacio $\mathcal{E}_{\lambda_i}(A)$, entonces la unión de las bases $\mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ es una base para $\mathbb{R}^n$.

 | La proposición 7.13 se mantiene valida si cambiamos el campo de los reales $\mathbb{R}$ por los complejos $\mathbb{C}$.

Demostración. Primero denotamos por $g_i := \dim(\mathcal{E}_{\lambda_i}(A))$ la multiplicidad geométrica del eigenvalor λ_i .

Ahora, mostramos que el conjunto $\mathcal{B} := \mathcal{B}_1 \cup \dots \cup \mathcal{B}_k$ es linealmente independiente. Por lema 7.11 las bases no tienen elementos en común, por lo cual el conjunto $\mathcal{B}$ posee $n_A := (g_1 + g_2 + \dots + g_k)$ elementos. Ahora, cada conjunto $\mathcal{B}_i$ consiste de g_i vectores LI, i.e. denotamos

$$\mathcal{B}_i := \{\vec{v}_1^i, \vec{v}_2^i, \dots, \vec{v}_{g_i}^i\}$$

y vemos si la combinación lineal del vector cero es unica. Sea

$$\sum_{\ell=1}^{g_1} \alpha_{1\ell} \vec{v}_\ell^1 + \sum_{\ell=1}^{g_2} \alpha_{2\ell} \vec{v}_\ell^2 + \dots + \sum_{\ell=1}^{g_k} \alpha_{k\ell} \vec{v}_\ell^k = \vec{0} \quad (\star)$$

una combinación lineal de vectores de $\mathcal{B}$ igual al vector $\vec{0}$. Esa siempre se puede escribir y agrupar de la manera anterior (por conmutatividad y asociatividad). Observamos que

$$\vec{w}_i := \sum_{\ell=1}^{g_i} \alpha_{i\ell} \vec{v}_\ell^i \in \mathcal{E}_{\lambda_i}(A) \quad \text{para} \quad i \in \{1, \dots, k\}$$

y el corolario 7.12 nos implica $\vec{w}_i = \vec{0}$ para $i \in \{1, 2, \dots, k\}$. Dado que $\mathcal{B}_i$ es LI, la igualdad $\vec{w}_i = \vec{0}$ implica $\alpha_{i\ell} = 0$ para $\ell \in \{1, \dots, g_i\}$. Esto aplica para cada suma en $(\star)$ y todos los coeficientes cumplen $\alpha_{i\ell} = 0$. Concluimos con teorema 4.2 que los vectores en $\mathcal{B}$ son L.I.

Ahora, si $n_A = (g_1 + g_2 + \dots + g_k) = n$, entonces lo anterior implica que el conjunto $\mathcal{B}$ es L.I. y por el corolario 4.8 es una base de $\mathbb{R}^n$.

Por otro lado si $\mathcal{B}$ es base de $\mathbb{R}^n$ entonces $\mathcal{B}$ tiene n vectores LI, es decir, las multiplicidades geométricas suman $n_A = (g_1 + g_2 + \dots + g_k) = n$. $\square$

7.3. Semejanza & Diagonalización

En demostraciones, algoritmos y materias superiores se usará el concepto de matrices semejantes y de la diagonalización.



Definición. Sean A, B matrices de $n \times n$. Se dice que A es **semejante** a B , si existe una matriz invertible P de $n \times n$ tal que

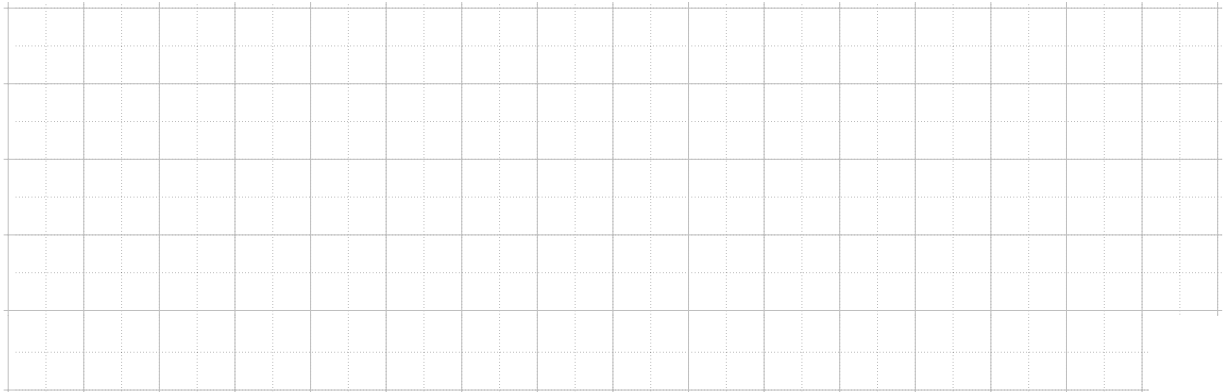
$$P^{-1}AP = B \quad \Longleftrightarrow \quad AP = PB.$$

Si A es semejante a B se escribe $A \sim B$.

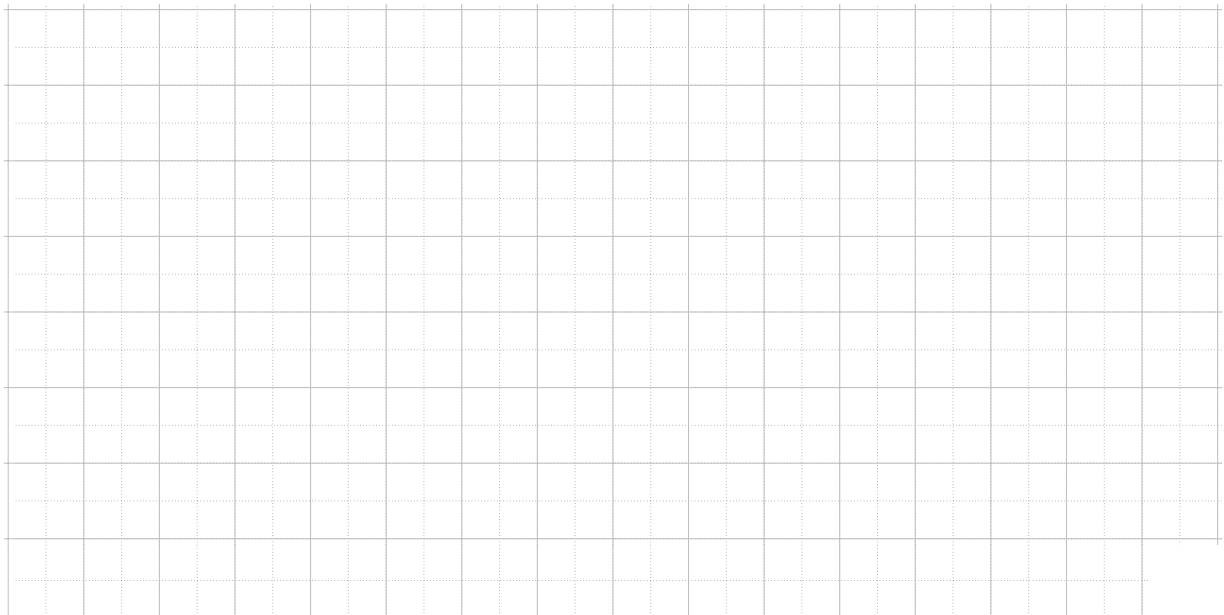
Ejemplo: Muestre que la matriz A es semejante con B a través de P usando las matrices dadas:

$$A = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ -2 & -1 \end{pmatrix} \quad \& \quad P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Sol. Forma I: Mostraremos que $AP = PB$:



Sol. Forma II: Mostraremos que $P^{-1}AP = B$:



Observación Algebraica:

La relación de semejanza $\sim$ es una *relación de equivalencia*, es decir, es

1. *reflexiva*, es decir, $A \sim A$
2. *simétrica*, es decir, $A \sim B$ y $B \sim A$
3. y *transitiva*, es decir, si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Demostración. La relación $\sim$ es reflexiva, ya que

La relación $\sim$ es simétrica, ya que

Ejercicio: Demuestre que la relación $\sim$ es transitiva. Consejo:

□



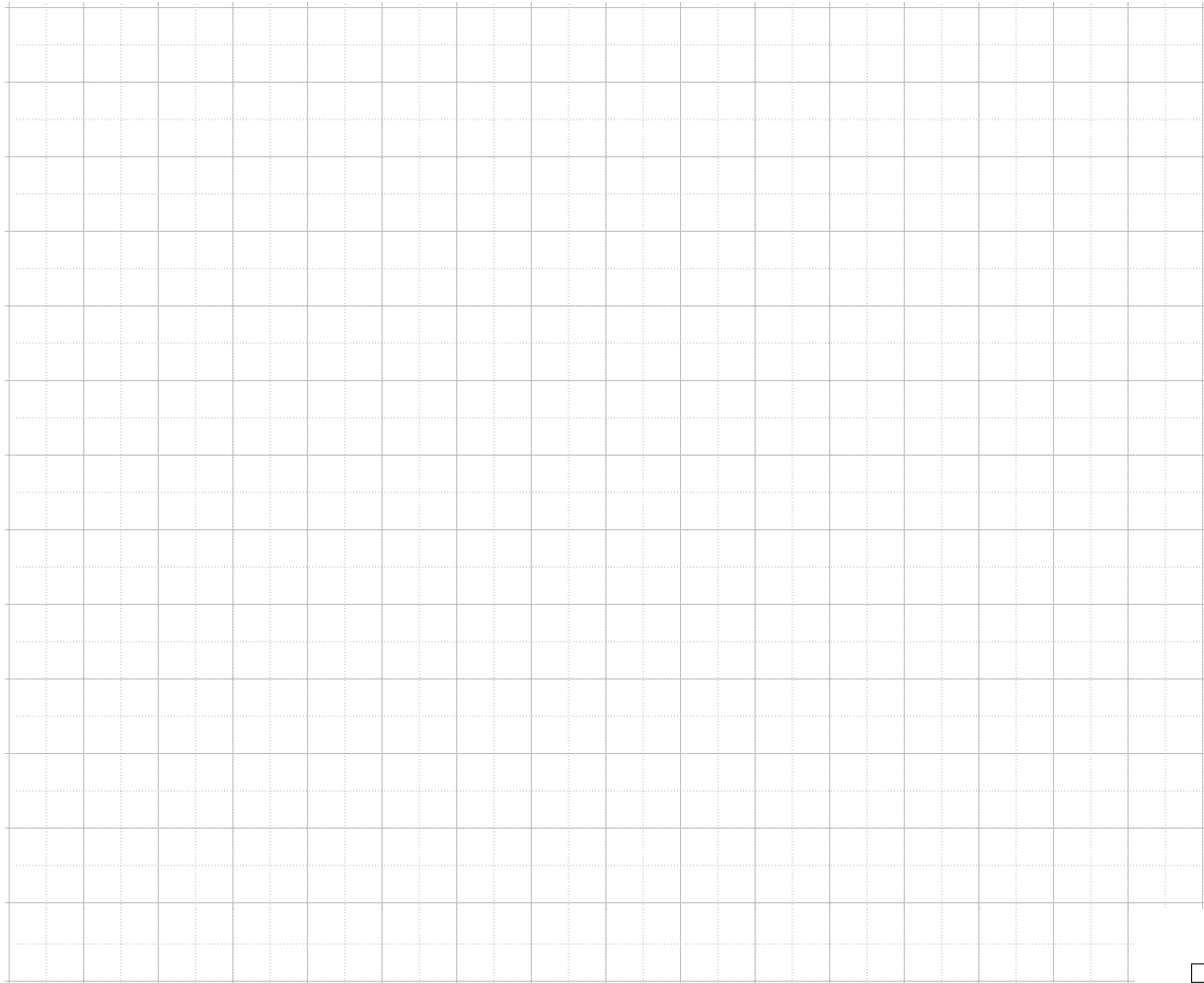
Teorema 7.14. Si las matrices A y B de $(n \times n)$ son semejantes, entonces tienen el mismo polinomio característico y los mismos eigenvalores con multiplicidades (algebraicas y geométricas) iguales.

Demostración.

Supongamos que $A \sim B$ por lo cual $P^{-1}AP = B$ para una matriz invertible P .



El polinomio característico se obtiene tomando el determinante:



□

Ejercicio: Suponga que P es invertible y que el conjunto $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k\}$ es L.I.

Demuestre que el conjunto $\{P\vec{v}_1, P\vec{v}_2, \dots, P\vec{v}_k\}$ es L.I.

7. Valores y vectores propios y Diagonalización de Matrices

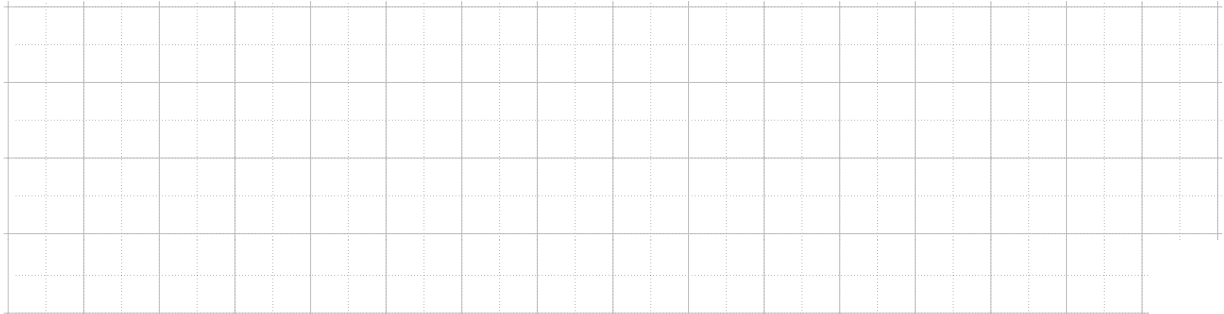
Un conjunto especial de matrices semejantes son las diagonalizables.



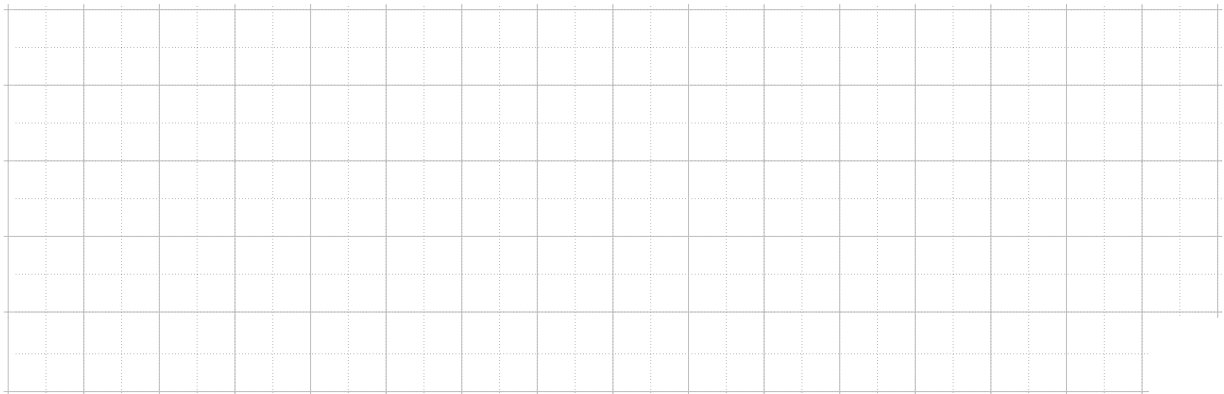
Definición. Una matriz A de $n \times n$ es **diagonalizable** si es semejante a una matriz diagonal D , es decir $P^{-1}AP = D$ donde P es una matriz invertible.

Determinantes y potencias de matrices diagonalizables se simplifican mucho.

Ejemplo (es bueno saber): Si $P^{-1}AP = D$, entonces $\det(A) = \det(D)$.



Ejemplo: Si $D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$, entonces



Ejemplo (es bueno saber): Si $A = PDP^{-1}$, entonces $A^k = PD^kP^{-1}$.





Teorema 7.15. Sea $A \in \mathbb{R}^{n \times n}$ una matriz. Suponga que existe una base de $\mathbb{R}^n$ formada por eigenvectores de A , entonces A es diagonalizable con D siendo la diagonal de eigenvalores y las columnas de P son eigenvectores de A . En símbolos: $AP = PD \iff P^{-1}AP = D$.



Nota: También, se cumple la otra dirección: Si A es diagonalizable ($AP = PD$), entonces la diagonal D contiene los eigenvalores de A y las columnas de P son eigenvectores de A . Pero los eigenvalores y -vectores para algunas matrices reales pueden ser complejos.

Demostración. Justificamos la existencia de una matriz P invertible y una diagonal D tal que A y D son semejantes.

Por hipótesis $\mathbb{R}^n$ tiene una base de n eigenvectores de A . Entonces el conjunto de ellos $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ es lin. independiente y los eigenvalores $\lambda_1, \dots, \lambda_n$ son reales (ya que $A \in \mathbb{R}^{n \times n}$, $\vec{v}_i \in \mathbb{R}^n$ y $A\vec{v}_i = \lambda_i\vec{v}_i$) pero no se pide que sean distintos. Definimos la matriz P como sigue: En cada columna P tiene un eigenvector de A , es decir

$$P = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \end{bmatrix}.$$

Esta matriz es invertible ya que tiene columnas L.I. (por teorema 5.3). Por la definición del producto de matrices deducimos que

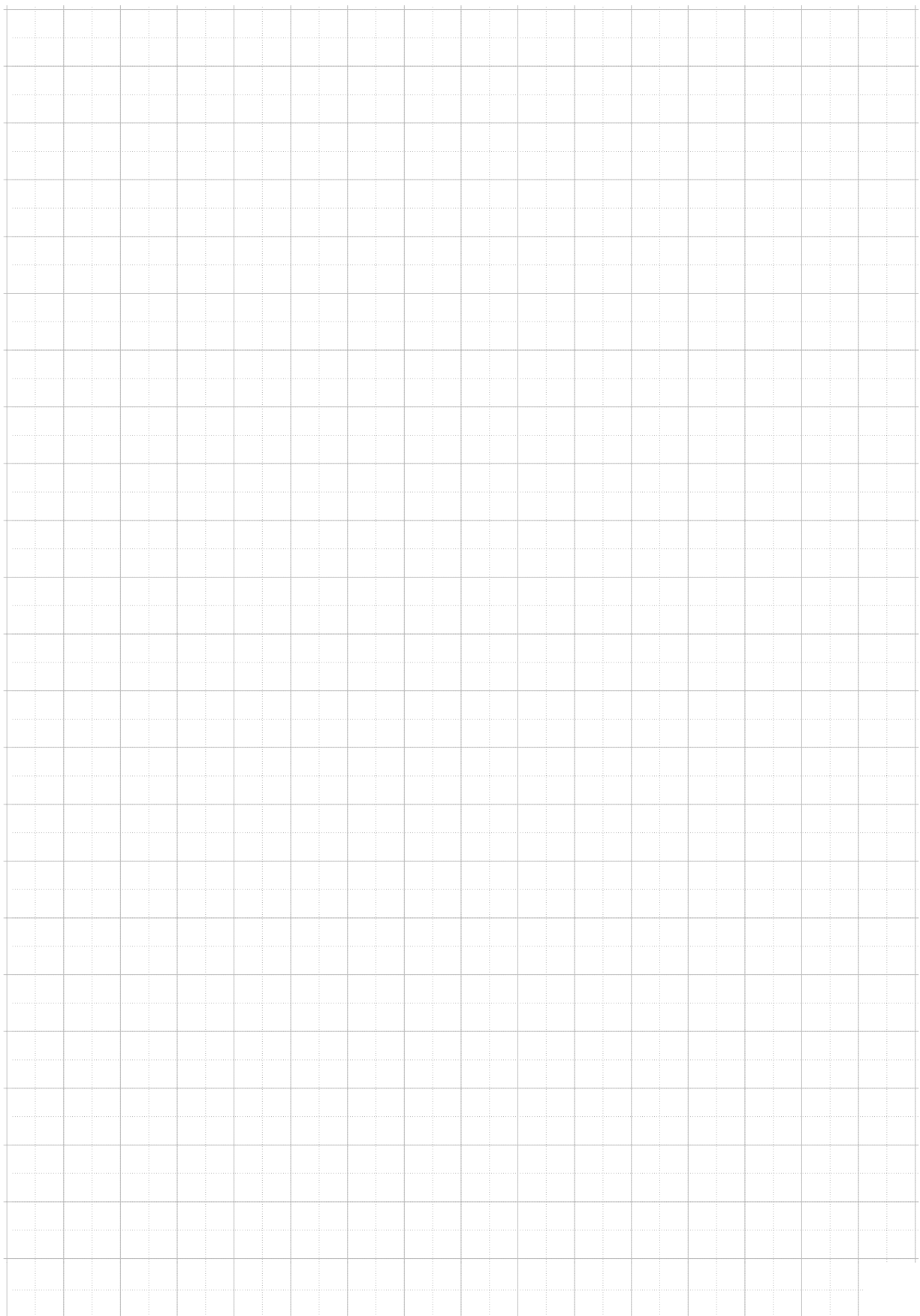
$$\begin{aligned} AP &= A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & \dots & A\vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \dots & \lambda_n\vec{v}_n \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 P\vec{e}_1 & \lambda_2 P\vec{e}_2 & \dots & \lambda_n P\vec{e}_n \end{bmatrix} \\ &= \begin{bmatrix} P(\lambda_1\vec{e}_1) & P(\lambda_2\vec{e}_2) & \dots & P(\lambda_n\vec{e}_n) \end{bmatrix} \\ &= P \begin{bmatrix} \lambda_1\vec{e}_1 & \lambda_2\vec{e}_2 & \dots & \lambda_n\vec{e}_n \end{bmatrix} = P \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} = PD \end{aligned}$$

Recordando que P es invertible, la última igualdad equivale a $P^{-1}AP = D$. Concluimos que A es diagonalizable, D es una diagonal con eigenvalores de A y las columnas de P son eigenvectores.

Nota: La unicidad no se da, ya que el orden de las columnas de P determina el orden de las entradas de D . □

Ejemplo: Demuestre que $A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}$ es diagonalizable y calcule A^{20} .



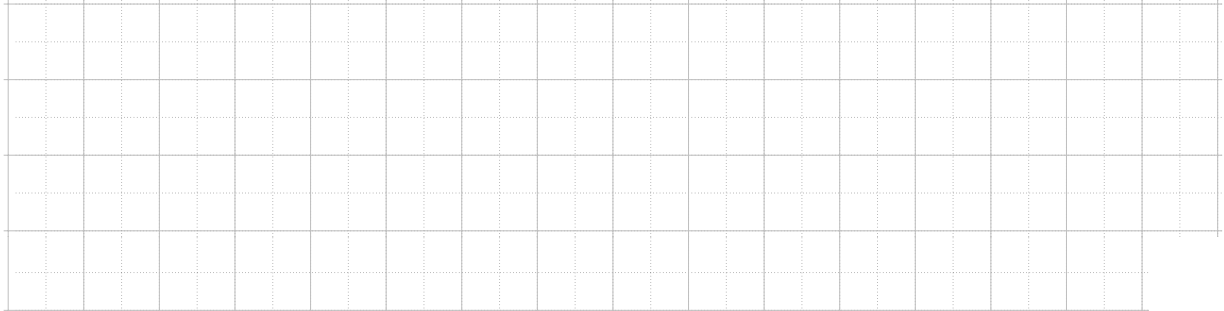


La longitud euclidiana de vectores

La longitud euclidiana de un vector en $\mathbb{R}^n$ generaliza el teorema de Pythagoras, por ejemplo:

Para un vector $\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$ mide la hipotenusa del triángulo $\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix}$, i.e. $\sqrt{x^2 + y^2}$.

Para un vector $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$ mide la hipotenusa del triángulo $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, i.e. $\sqrt{x^2 + y^2 + z^2}$.



Definición 7.3. Longitud del vector $\vec{u} \in \mathbb{R}^n$ es su **norma euclidiana** definida por

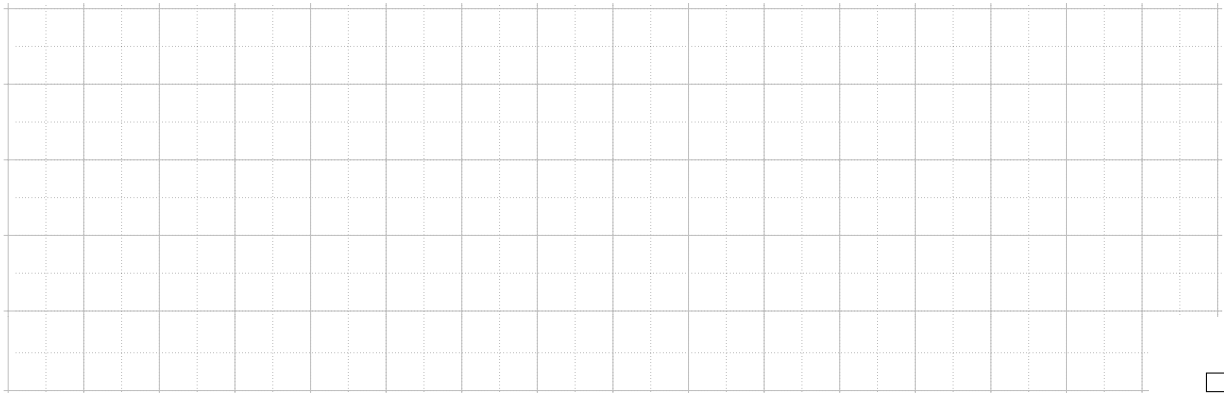
$$\|\vec{u}\|_e = \sqrt{u_1^2 + u_2^2 + \dots + u_n^2} \geq 0.$$

Interpretación: La longitud del vector $\vec{u}$ es $\|\vec{u}\|_e$.

Una propiedad importante:

Si $\vec{u} \neq \vec{0}$, entonces se puede **normalizar** es decir, $\vec{v} = \frac{1}{\|\vec{u}\|_e} \vec{u}$ tiene longitud 1, es decir, $\|\vec{v}\|_e = 1$.

Justificamos.



□

El siguiente teorema es importantísimo en aplicaciones, probabilidad, optimización, ...



Teorema 7.16. Si $A \in \mathbb{R}^{n \times n}$ es una matriz real simétrica, entonces sus eigenvalores son reales (posiblemente iguales) y $\mathbb{R}^n$ tiene una base de eigenvectores de A . Es decir, $AV = V\Lambda$ donde $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ y las columnas de V son eigenvectores de A .

Además, $V^T V$ es una matriz diagonal y V se puede normalizar tal que $V^T V = I$.

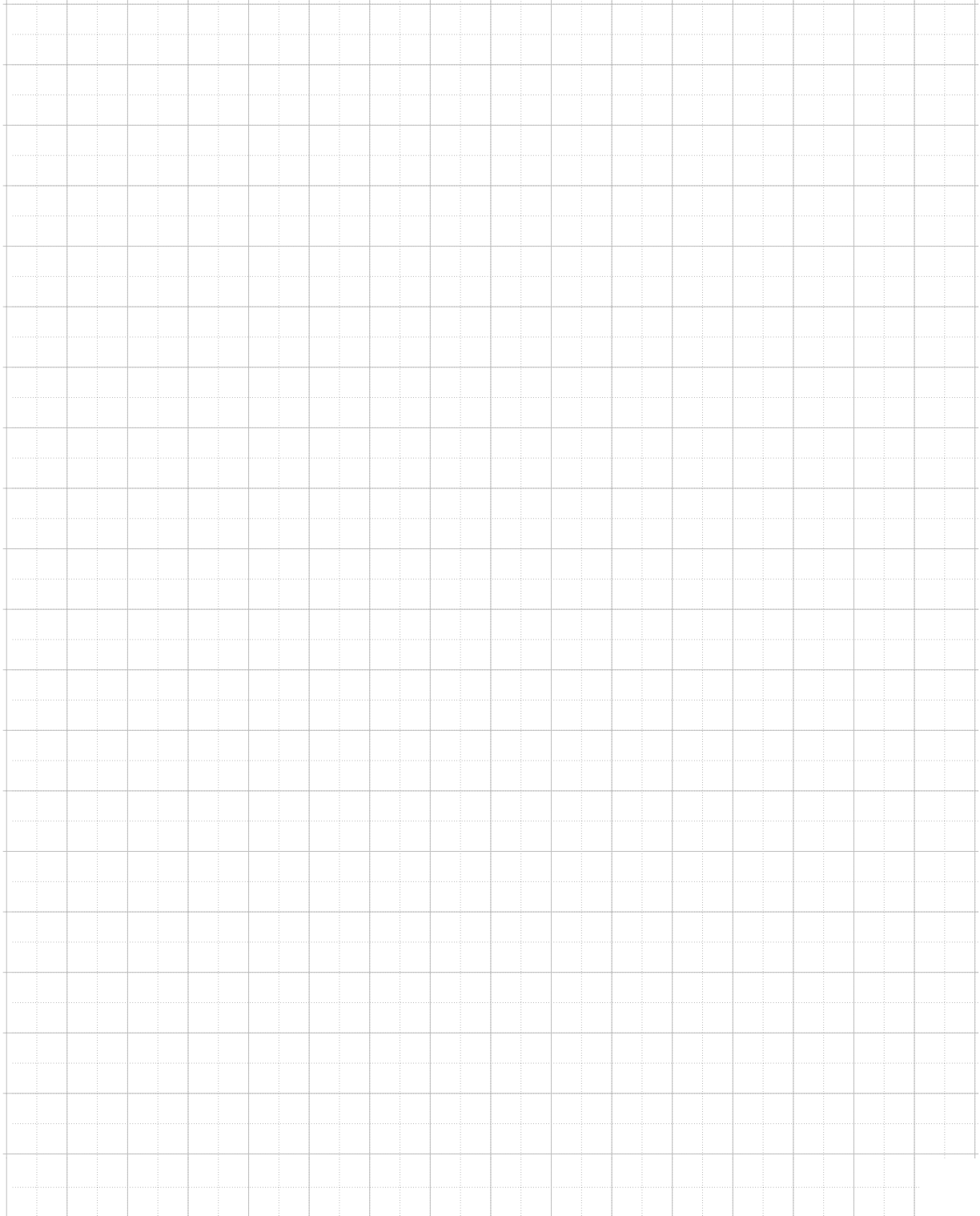
Demostración. Esta prueba queda afuera de nuestro alcance. Ver Álgebra lineal 2.

□

7. Valores y vectores propios y Diagonalización de Matrices

Ejemplo: Diagonaliza la matriz:

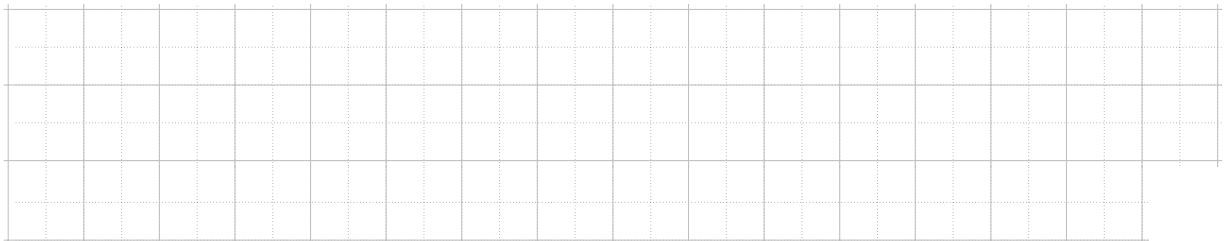
$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 1 & 5 \end{pmatrix}$$



8. Transformaciones Lineales

Este material fue tomado del libro de los profesores Mota & Rumbo y de las notas del Prof. John Truss.

Cuando se define una función entre espacios vectoriales, no solo interesa enviar vectores de un espacio a otro, sino también nos interesa preservar estructura.



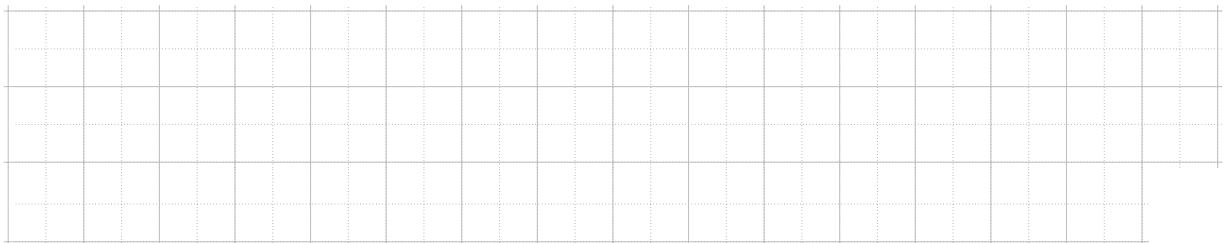
Las transformaciones lineales entre espacios vectoriales se definen como sigue:



Definición 8.1. Sean $(V, +, *_{esc})$ de dim. n & $(W, +, *_{esc})$ de dim. m dos espacios vectoriales sobre el campo $\mathbb{R}$. Una función $T: V \rightarrow W$ es una **Transformación lineal** si cumple:

1. $T(\vec{u} + \vec{v}) = T(\vec{u}) + T(\vec{v})$,
2. $T(\alpha\vec{u}) = \alpha T(\vec{u})$.

Nota: de la definición anterior se sigue que:

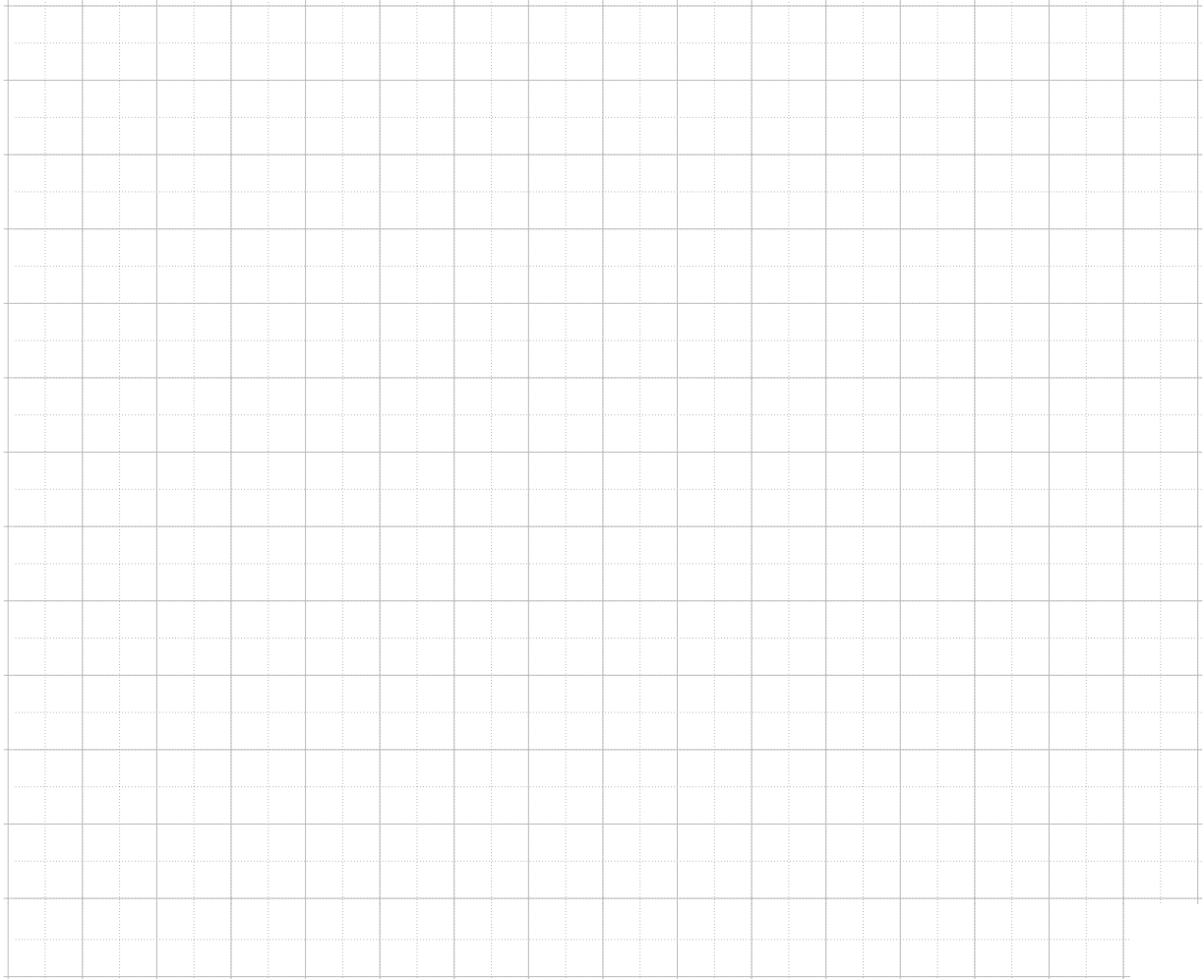


Ejemplo 1: Sea $T: \mathbb{R} \rightarrow \mathbb{R}$ definida por $T(x) = 8x$. Demuestre que es una transformación lineal.

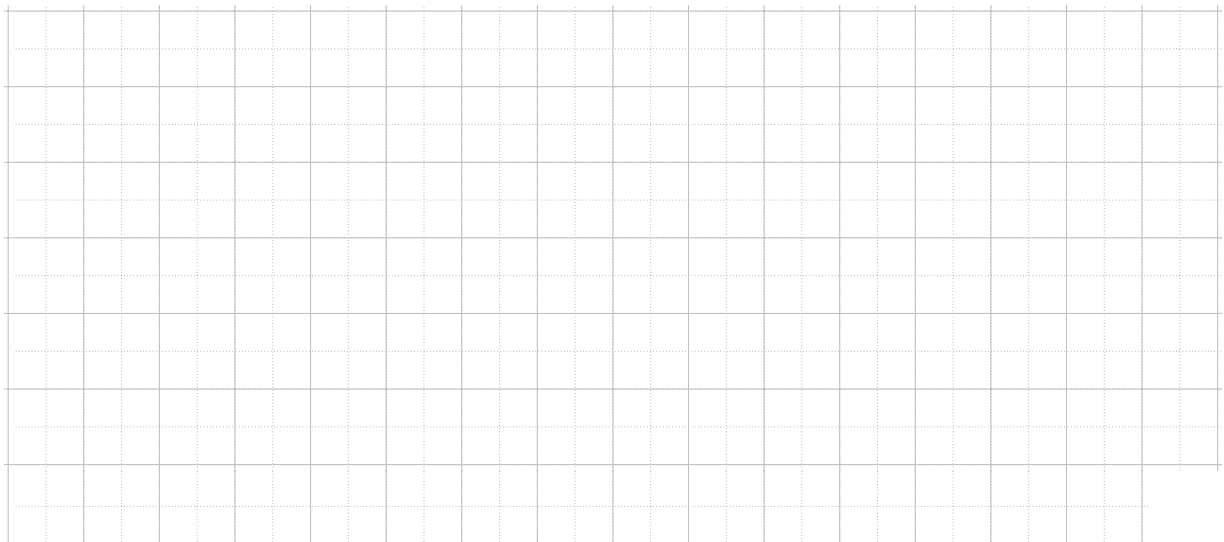


8. Transformaciones Lineales

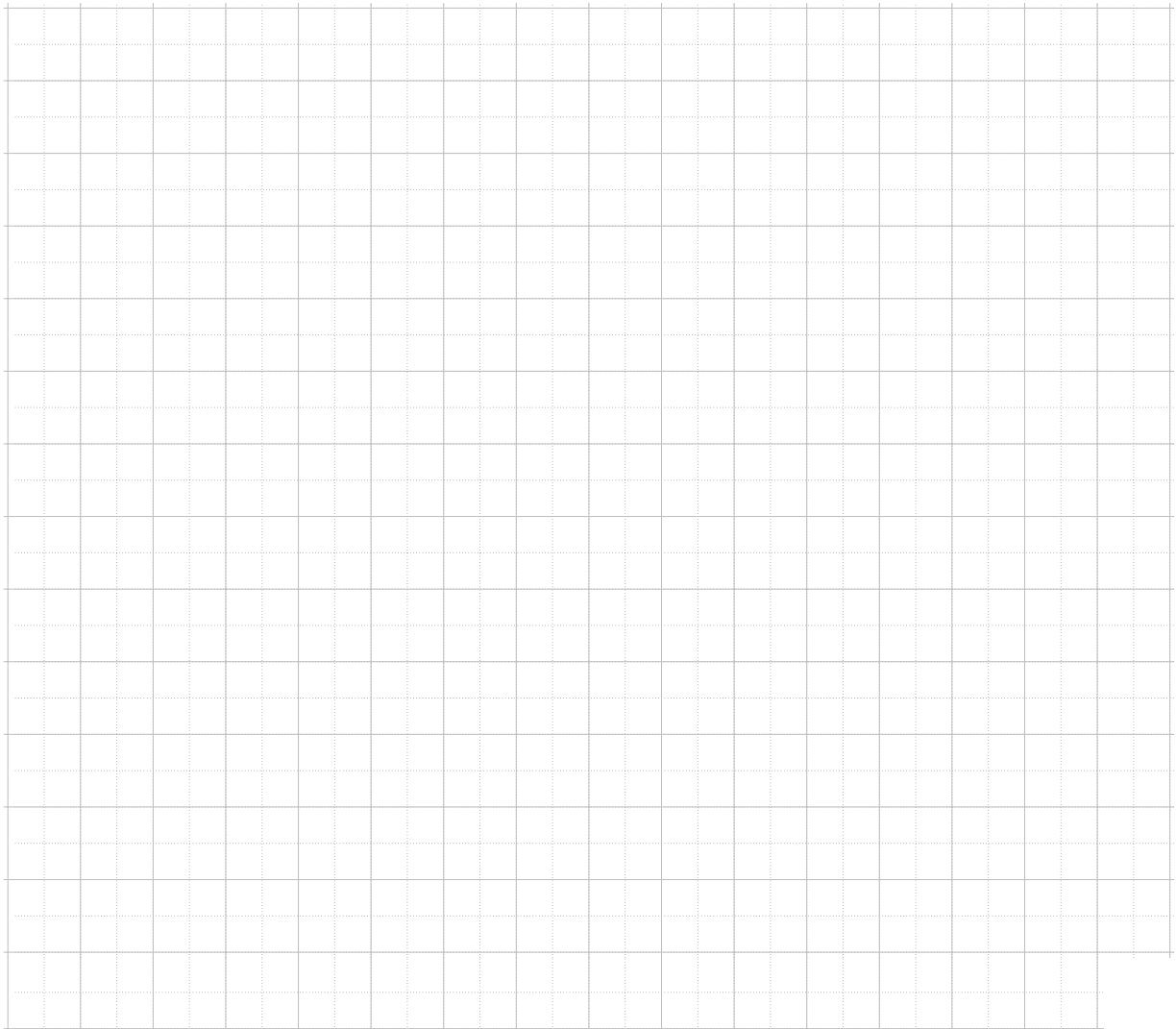
Ejemplo 2: Demuestre que $T: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dada por $T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 3x_2 \\ x_2 - 2x_3 \end{pmatrix}$ es una transformación lineal.



Ejemplo 3: La diferenciación nos da una transformación lineal entre el espacio de polinomios de grado n y $n - 1$, con la suma y la multiplicación por escalar $(P_n, +, *_{esc})$ y $(P_{n-1}, +, *_{esc})$.



Ejemplo 4: Compruebe que $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ dada por $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ 2x-y \\ 3x+4y \end{pmatrix}$ es una transformación lineal.



 **Lema 8.1.** Sea A una matriz de $m \times n$, entonces la función $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ dada por $T(\vec{x}) = A\vec{x}$ para todo $\vec{x} \in \mathbb{R}^n$ es una transformación lineal.

Demostración.

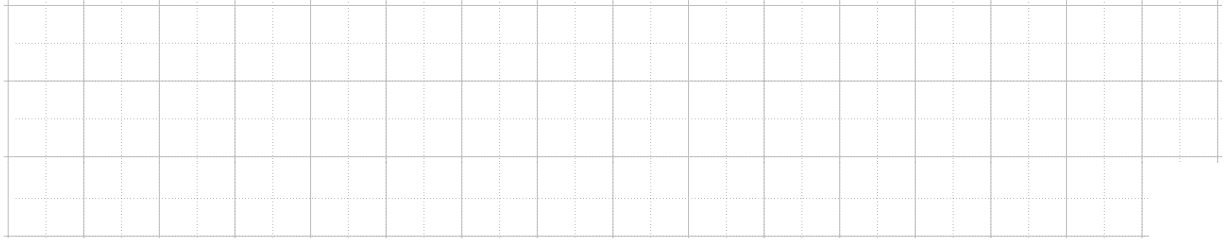


□

8. Transformaciones Lineales

Interpretación: El lema anterior nos dice que cada matriz determina una transformación lineal.

A continuación veremos que la conversa a este lema también es válido. Para desarrollar la demostración ocuparemos la siguiente identidad:

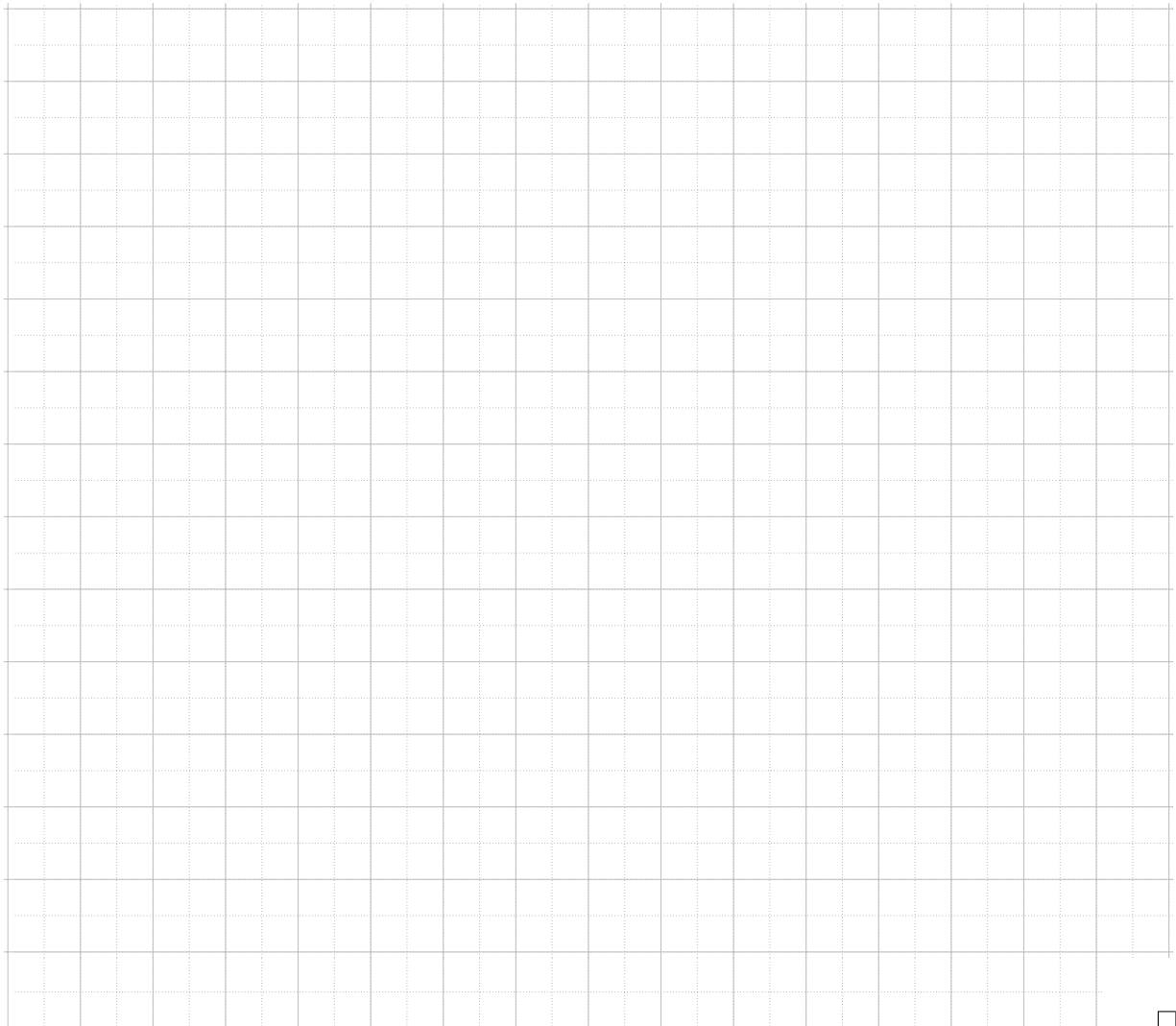


Lema 8.2. Si $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es una transformación lineal, entonces existe una matriz $A \in \mathbb{R}^{m \times n}$ tal que $T(\vec{x}) = A\vec{x}$, donde A esta dada por

$$A = \begin{bmatrix} T(\vec{e}_1) & T(\vec{e}_2) & \dots & T(\vec{e}_n) \end{bmatrix}.$$

y se conoce como la **matriz estándar** debido a que $\{\vec{e}_1, \dots, \vec{e}_n\}$ es la base **estándar** de $\mathbb{R}^n$.

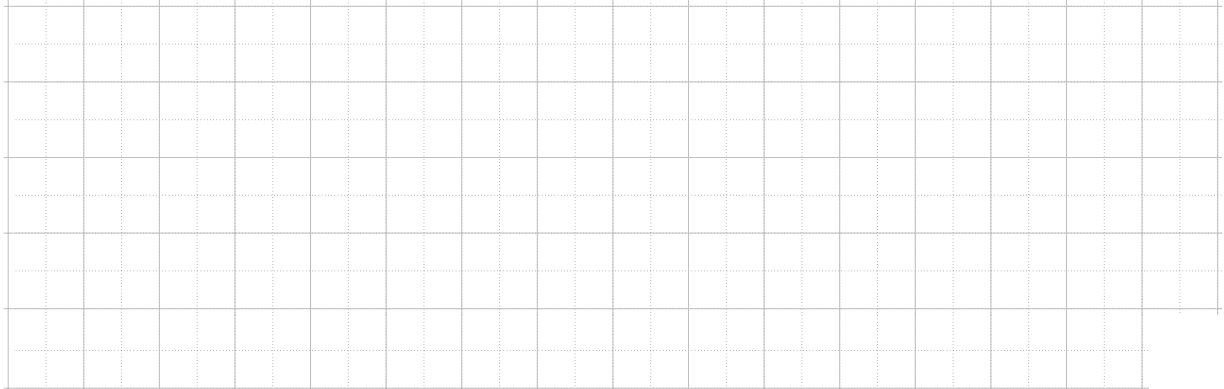
Demostración.



□

Ejemplo 5: Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x \\ -y \end{pmatrix}$.

- Demuestre que T es transformación lineal (ejercicio).
- Encuentre la matriz estándar asociada a esta transformación lineal.



Ejemplo 6: Sea $T_\alpha: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la transformación lineal definida por

$$T_\alpha \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x \cos(\alpha) - y \sin(\alpha) \\ x \sin(\alpha) + y \cos(\alpha) \end{pmatrix}.$$

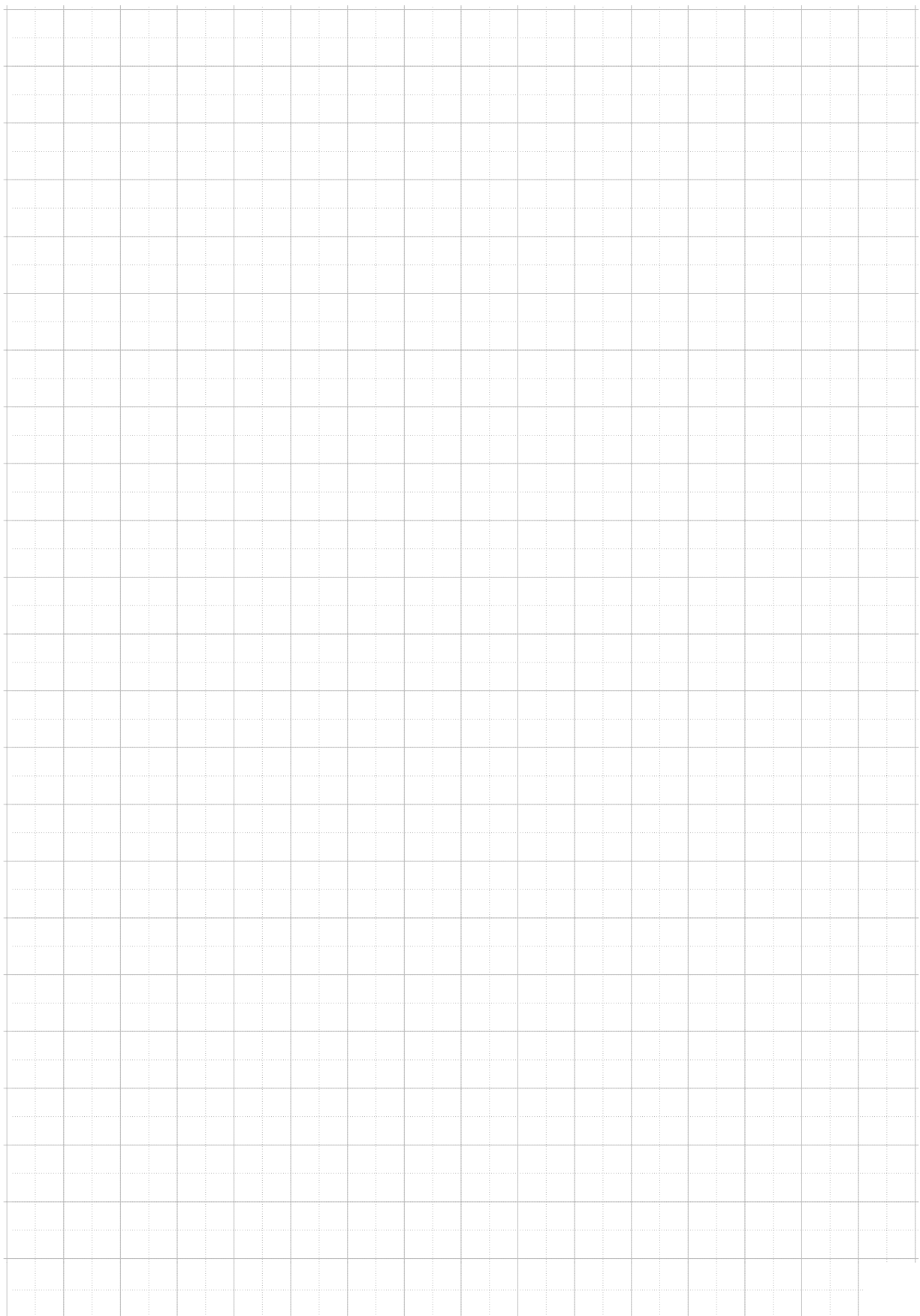
Encuentre la matriz estándar asociada a T_α :



Ejemplo 7: Sea $T: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ la transformación lineal definida por

$$T \left(\begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x \\ 0 \\ y \\ 2x + 3y \end{pmatrix}.$$

Encuentre la matriz estándar tal que $T(\vec{u}) = A\vec{u}$ para $\vec{u} \in \mathbb{R}^2$. Después calcule $T\left(\begin{pmatrix} 4 \\ -5 \end{pmatrix}\right)$.



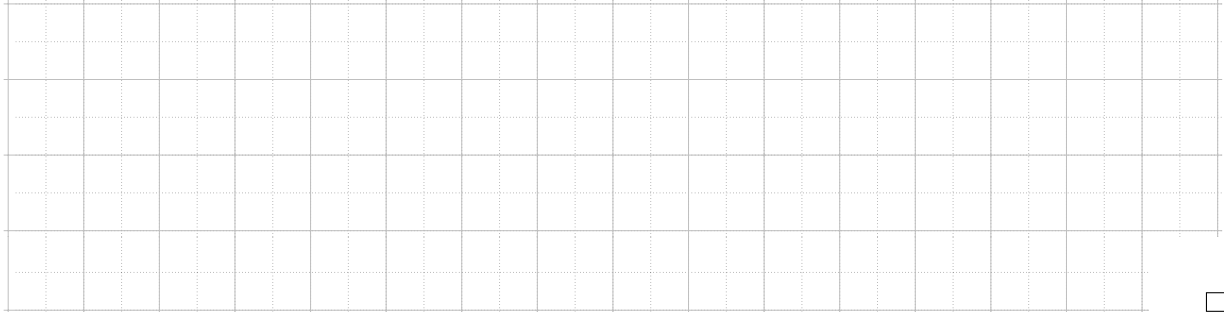
8.1. Propiedades



Lema 8.3. Si $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una transformación lineal, entonces T mapea al vector $\vec{0}_n \in \mathbb{R}^n$ al vector $\vec{0}_m \in \mathbb{R}^m$, es decir, se cumple $T(\vec{0}_n) = \vec{0}_m$.

Demostración. Se puede mostrar con lema 8.2 o usando propiedades de espacios vectoriales.

Sean $\vec{0}_n \in \mathbb{R}^n$ el elemento neutral de la suma en $\mathbb{R}^n$. Entonces para todo $\vec{v} \in \mathbb{R}^n$ se tiene $\vec{v} + \vec{0}_n = \vec{v}$, en particular se cumple para $\vec{v} = \vec{0}_n$, i.e., se obtiene $\vec{0}_n + \vec{0}_n = \vec{0}_n$ por lo cual



□

Ejercicio 1:

- Justifique porque la función $S: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ dada por $S\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \begin{pmatrix} x \\ 1 \\ y \\ 2x+3y \end{pmatrix}$ no es una transformación lineal.
- Verifique si la transformación $T: \mathbb{R} \rightarrow \mathbb{R}$ dada por $T(x) = 2x + 8$ es una transformación lineal.



Lema 8.4. Si $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ elementos de un espacio vectorial $(\mathcal{V}, +, \cdot)$ sobre los reales, entonces $T: \mathbb{R}^n \rightarrow \mathcal{V}$ dada por

$$T(\vec{x}) = T\left(\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right) = x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_n\vec{v}_n$$

es una transformación lineal.

Demostración. La parte $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ se deja de ejercicio.

A continuación mostraremos que si $\gamma \in \mathbb{R}$ & $\vec{x} \in \mathbb{R}^n$, entonces



□

Recuerde:

- Por lemas 8.1 y 8.2 cada transformación lineal $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ es representable por una matriz $A \in \mathbb{R}^{m \times n}$.
- Por lema 8.3 transformaciones lineales preservan elementos neutrales de la suma.
- Por lema 8.4 una combinación lineal es resultado de una transformación lineal.

8.2. Rango (Imagen) y núcleo

Definición 8.2. Sean $(\mathcal{V}, +, \cdot)$ & $(\mathcal{W}, +, \cdot)$ dos espacios vectoriales.

Además, sea $T: \mathcal{V} \rightarrow \mathcal{W}$ una transformación lineal. Entonces

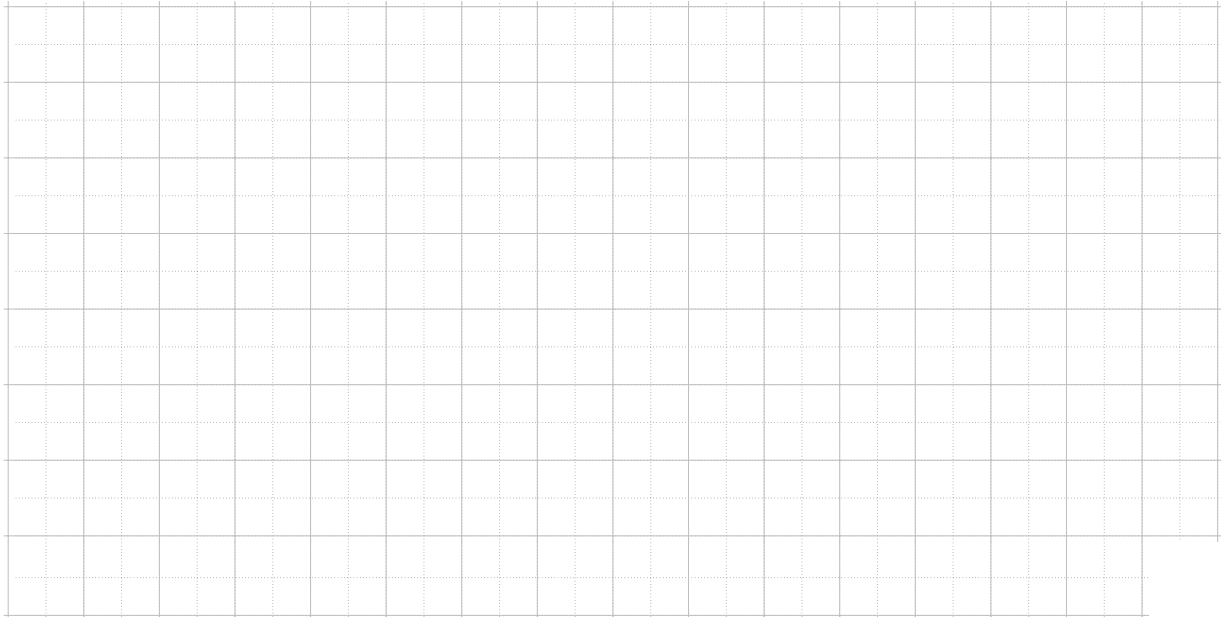
- El **dominio** de T es $\mathcal{V}$, *i.e.* “el conjunto inicial”.
- El **codominio** (también contra-dominio) de T es $\mathcal{W}$, *i.e.* “el conjunto final”.
- La **imagen** (ingles: **image, range**) de T se define como

$$\text{Img}(T) := \left\{ \vec{b} \in \mathcal{W} : T(\vec{x}) = \vec{b} \text{ para algún } \vec{x} \in \mathcal{V} \right\} = \{T(\vec{x}) : \vec{x} \in \mathcal{V}\}.$$

- El **núcleo** de T , es también conocido como **kernel** $\ker(T)$ y esta dado por

$$\text{Ker}(T) := \left\{ \vec{x} \in \mathcal{V} : T(\vec{x}) = \vec{0}_{\mathcal{W}} \right\}.$$

que es el conjunto de vectores en $\mathcal{V}$ que van a dar al vector $\vec{0}$.

**Caso especial:**

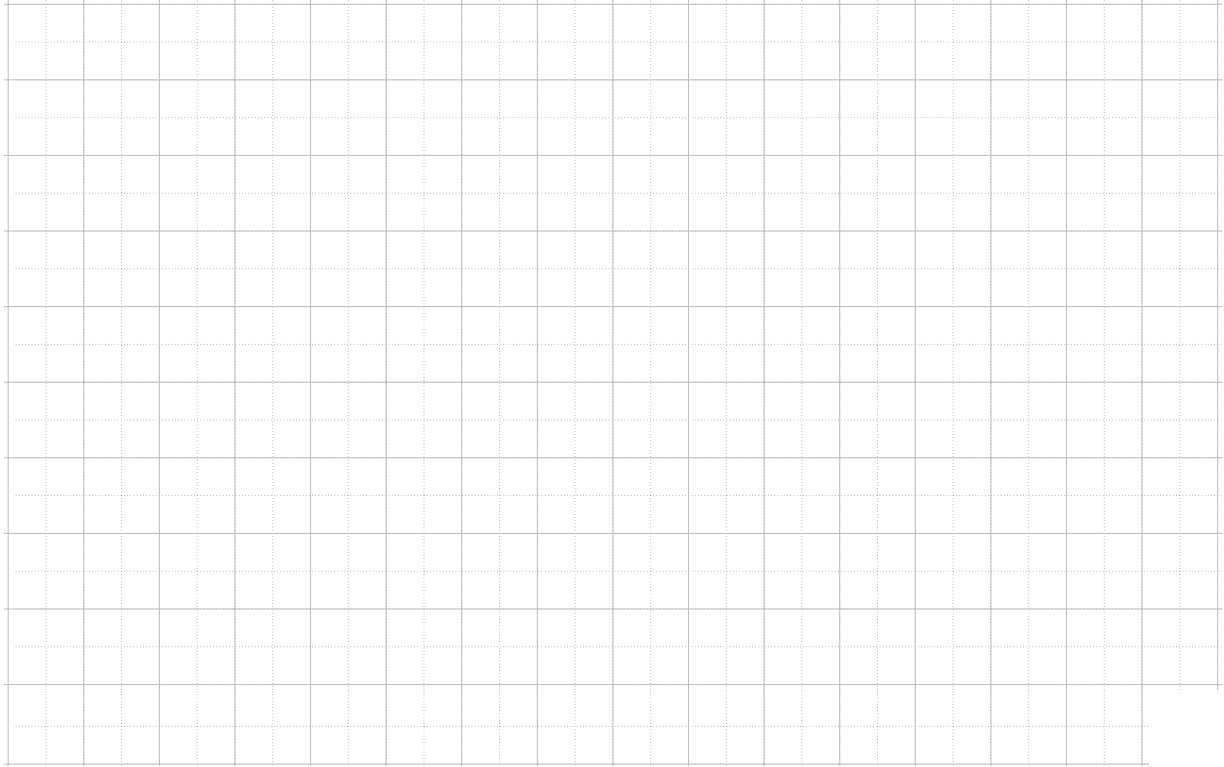
Si $\mathcal{V} = \mathbb{R}^n$, $\mathcal{W} = \mathbb{R}^m$ y $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$, entonces por lema 8.2 se tiene $T(\vec{v}) = A\vec{v}$. Por lo tanto:

$$\text{Img}(T) = \{A\vec{x} : \vec{x} \in \mathbb{R}^n\} = \text{Img}(A) \quad \& \quad \text{Ker}(T) = \left\{ \vec{x} \in \mathbb{R}^n : A\vec{x} = \vec{0} \right\} = \text{Null}(A).$$

Ejemplo sombra: Sea $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por la proyección del espacio al plano $x - y$, i.e. la sombra de vectores con el “sol en el cenit”

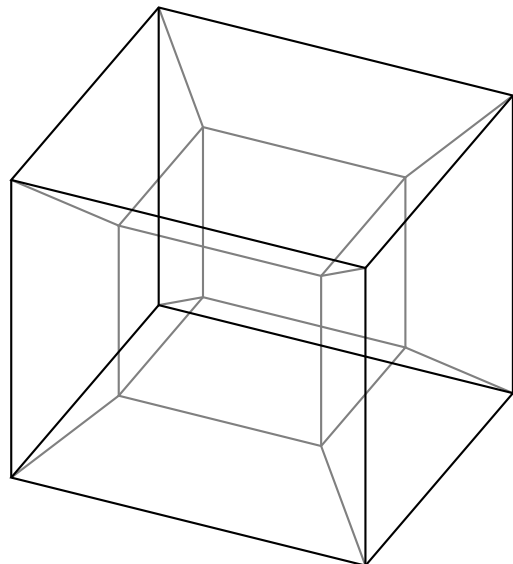
$$T \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}.$$

Calcule la imagen & el kernel de la esa transformación.



Ejemplo: Si consideramos la transformación lineal $T: \mathbb{R}^2 \rightarrow \mathbb{R}^4$ dada por $T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ x+y \\ 0 \\ 0 \end{pmatrix}$.

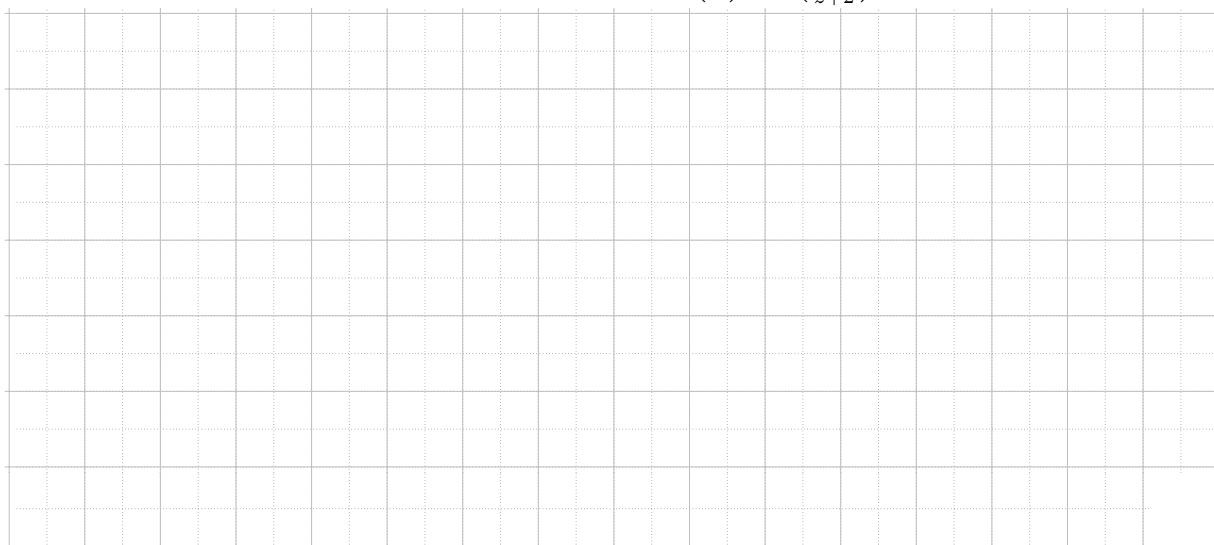
Calcule la imagen y el kernel de esta transformación.





8.2.1. Transformaciones lineales son muy restrictivas

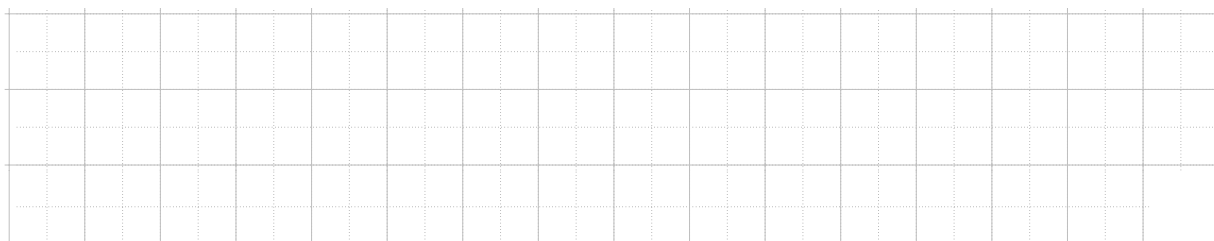
Ejemplo: Sea $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ una traslación dada por $T \begin{pmatrix} x \\ y \\ z \end{pmatrix} := \begin{pmatrix} x+2 \\ y+2 \\ z+2 \end{pmatrix}$.



Ejemplo: Sea $T: \mathbb{R}^n \rightarrow [0, \infty)$ la longitud del vector $\vec{v} \in \mathbb{R}^n$, es decir $T(\vec{v}) = \|\vec{v}\|_e$ donde

$$\|\vec{v}\|_e = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

Entonces, T NO es una transformación lineal ya que:



8.2.2. Transformaciones lineales preservan propiedades

Una transformación lineal $T: \mathcal{V} \rightarrow \mathcal{W}$ mapea vectores en vectores, puntos igualmente espaciados (en una recta) en puntos igualmente espaciados y mapea subespacios vectoriales de $\mathcal{V}$ a subespacios vectoriales de $\mathcal{W}$.

Ejemplo: Si tenemos puntos $\vec{u}, \vec{w}, \vec{v}$ igualmente espaciados en una recta en $\mathcal{V}$, entonces los puntos $T(\vec{u}), T(\vec{w}), T(\vec{v})$ son igualmente espaciados. Solo tenemos que ver si el punto medio $\vec{p}_m = \frac{1}{2}(\vec{u} + \vec{w}) = \frac{1}{2}\vec{u} + \frac{1}{2}\vec{w}$ será mapeado al punto medio entre $T(\vec{u})$ y $T(\vec{w})$.



Ejemplo: Un **producto punto** es una transformación lineal. Por ejemplo:

Si $T: \mathbb{R}^3 \rightarrow \mathbb{R}$ es definido como $T(\vec{u}) = \vec{a}^\top \vec{u}$ donde $\vec{a}^\top = (1, 2, 3) \in \mathbb{R}^{1 \times 3}$, entonces

$$T(\vec{u}) = 1u_1 + 2u_2 + 3u_3.$$

Además, si $T(\vec{u}) \neq 0$, entonces $\{T(\alpha\vec{u}) : \alpha \in \mathbb{R}\} = \{\alpha T(\vec{u}), \alpha \in \mathbb{R}\} = \mathbb{R}$, es decir, T mapea el espacio vectorial $(\mathbb{R}^3, +, \cdot)$ al espacio $(\mathbb{R}, +, *)$.

Se verán más propiedades pero algunas se basan en los conceptos de transformaciones lineales inyectivas, sobreyectivas y biyectivas.

8.3. Transformaciones inyectivas, sobres y biyectivas.



Definición 8.3. Sea T una transformación lineal $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$.

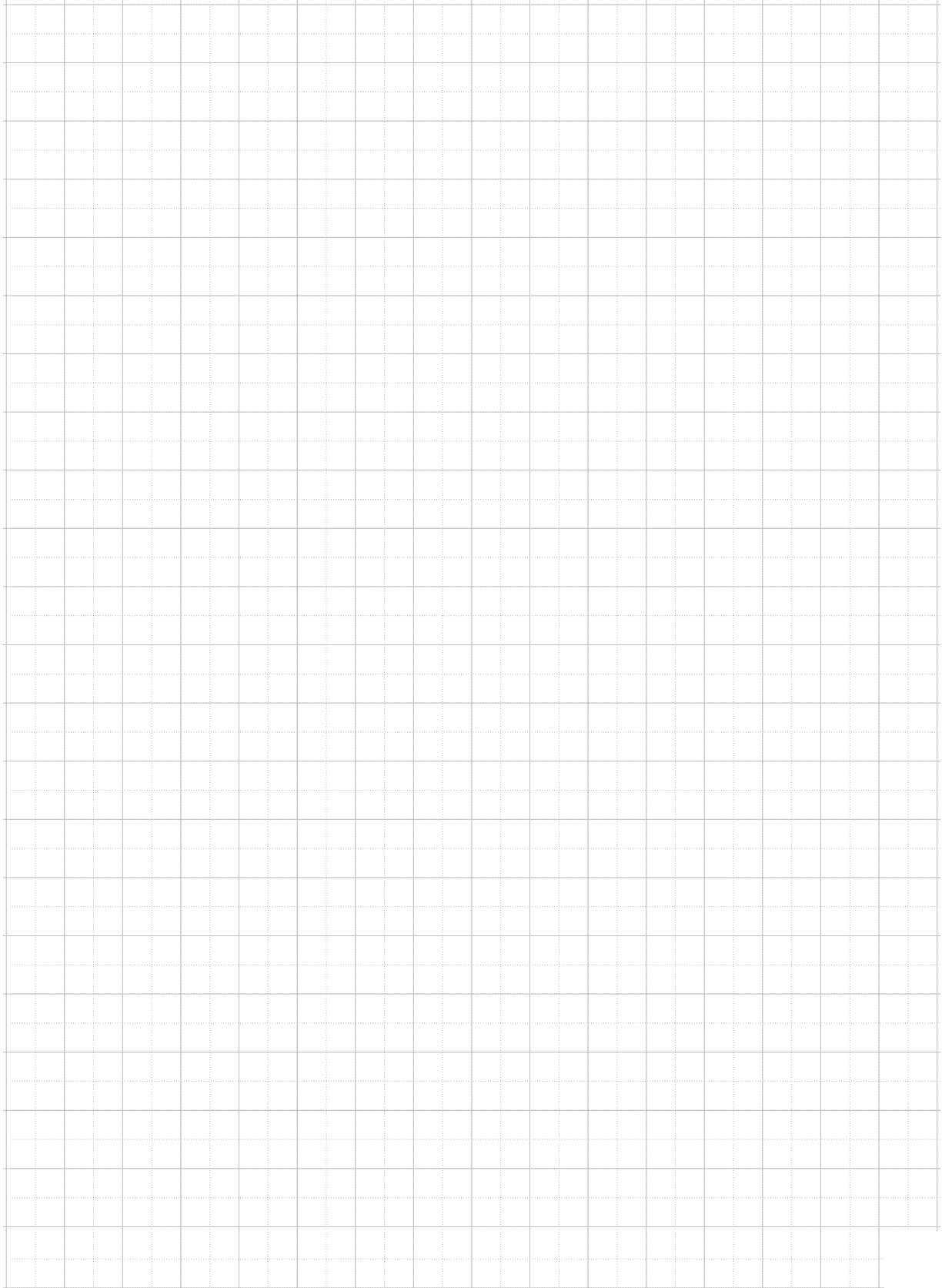
Se dice que T es **sobre** (o **suprayectiva**) si se cumple $\text{Img}(T) = \mathbb{R}^m$.

Se dice que T es **inyectiva** (o **uno-a-uno**) si y solo si $\vec{x} \neq \vec{y}$ implica $T(\vec{x}) \neq T(\vec{y})$.

Alternativamente, T es inyectiva si para cualesquiera dos vectores $\vec{x}, \vec{y} \in \mathbb{R}^n$ se cumple que $T(\vec{x}) = T(\vec{y})$ implica $\vec{x} = \vec{y}$.

8. Transformaciones Lineales

Ejemplos:





Teorema 8.5. Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una transformación lineal. Entonces

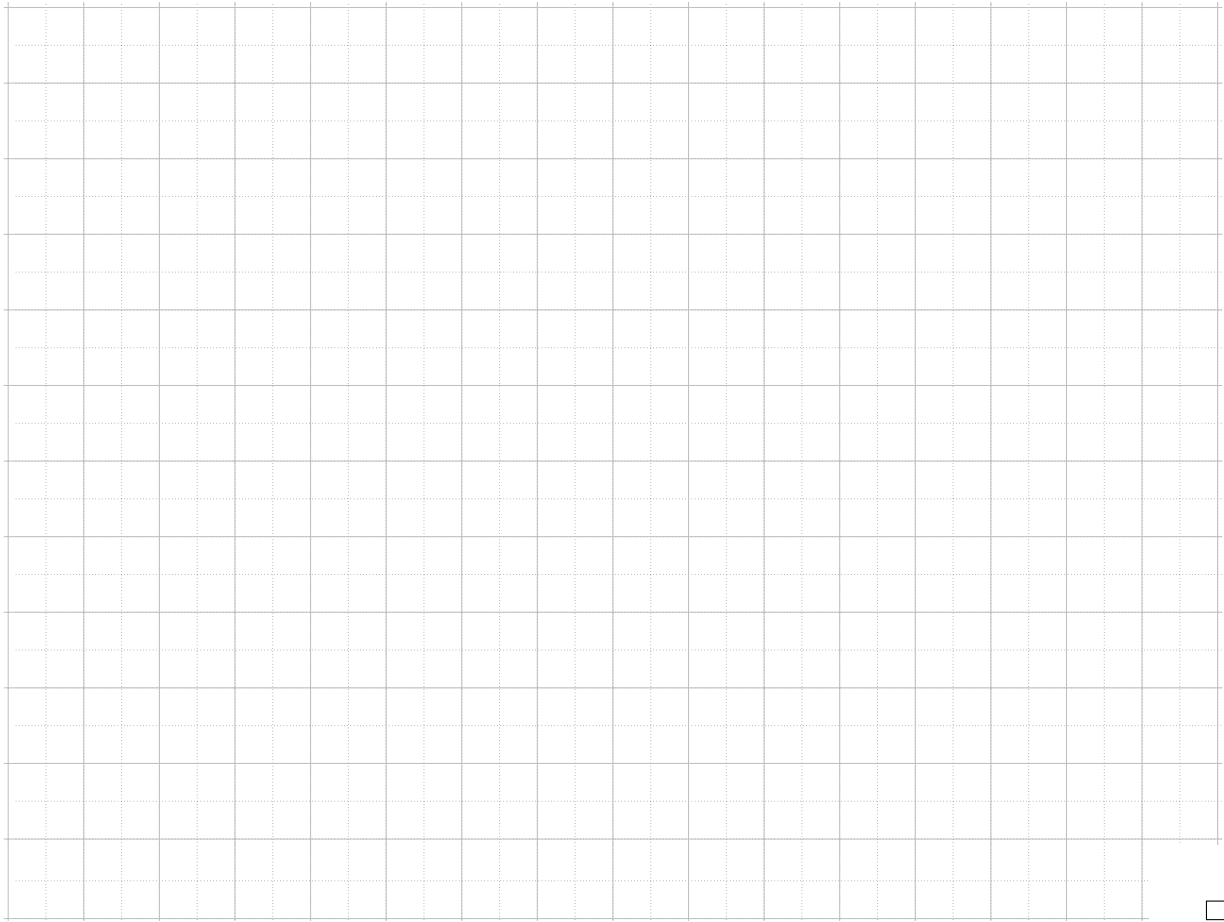
$$T \text{ es inyectiva} \iff \text{Ker}(T) = \{\vec{0}\}.$$

Con la matriz estándar $A \in \mathbb{R}^{m \times n}$ de T , eso se puede reescribir como:

$$T \text{ es inyectiva} \iff \text{Null}(A) = \{\vec{0}\}.$$

Interpretación: Inyectividad requiere que las columnas de $A \in \mathbb{R}^{m \times n}$ son linealm. independientes. Eso solo es posible cuando $m \geq n$.

Demostración.



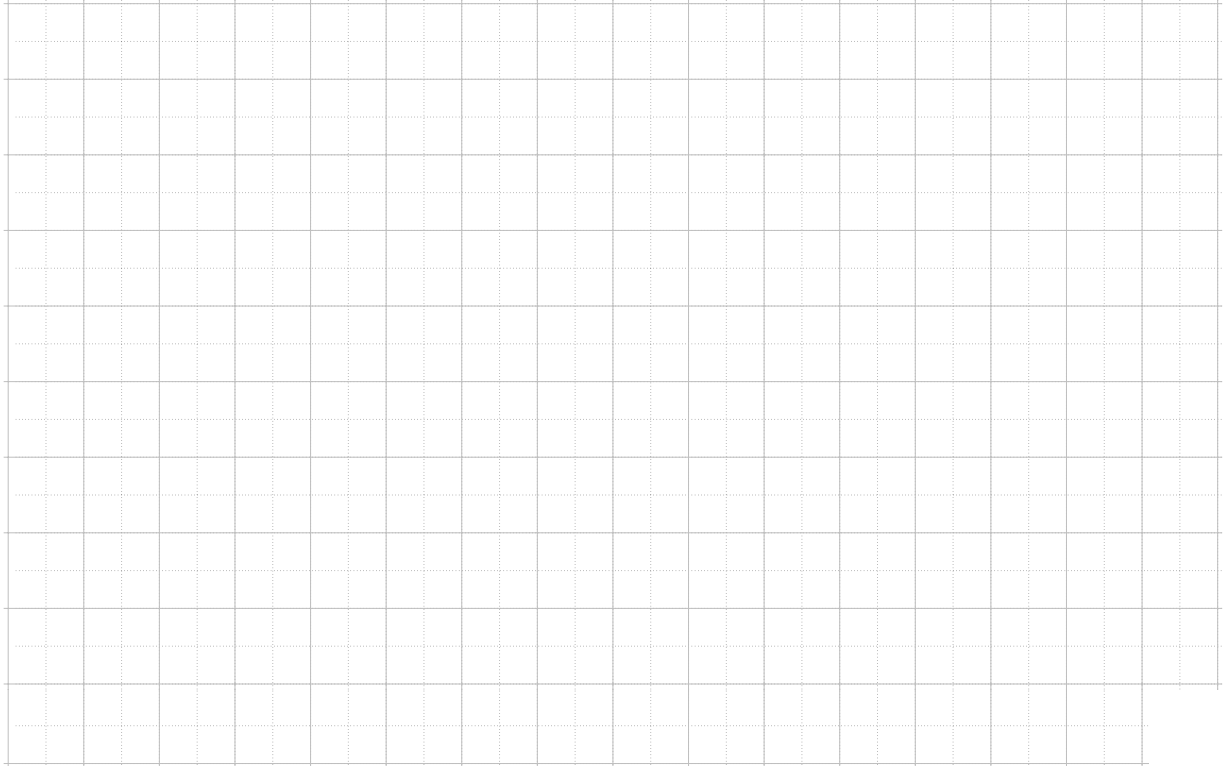
□

Interpretaciones: Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una transformación lineal. Entonces

- T es suprayectiva cuando los renglones de su matriz estándar son L.I.
- T es suprayectiva cuando las columnas de su matriz estándar generan $\mathbb{R}^m$.
- T no es suprayectiva cuando existe algún vector $\vec{b} \in \mathbb{R}^m$ tal que la ecuación $T(\vec{x}) = \vec{b}$ no tenga solución.
- T es inyectiva cuando las columnas de su matriz estándar son L.I.
- T no es inyectiva cuando existe algún vector $\vec{b} \in \mathbb{R}^m$ tal que la ecuación $T(\vec{x}) = \vec{b}$ tenga dos (o más) soluciones.

8. Transformaciones Lineales

Ejemplo: Sea $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x-y \\ 2x+y \end{pmatrix}$. Determine si T es inyectiva.



Ejemplo: Sea $T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x-y \\ -2x+2y \end{pmatrix}$. Determine si T es inyectiva.

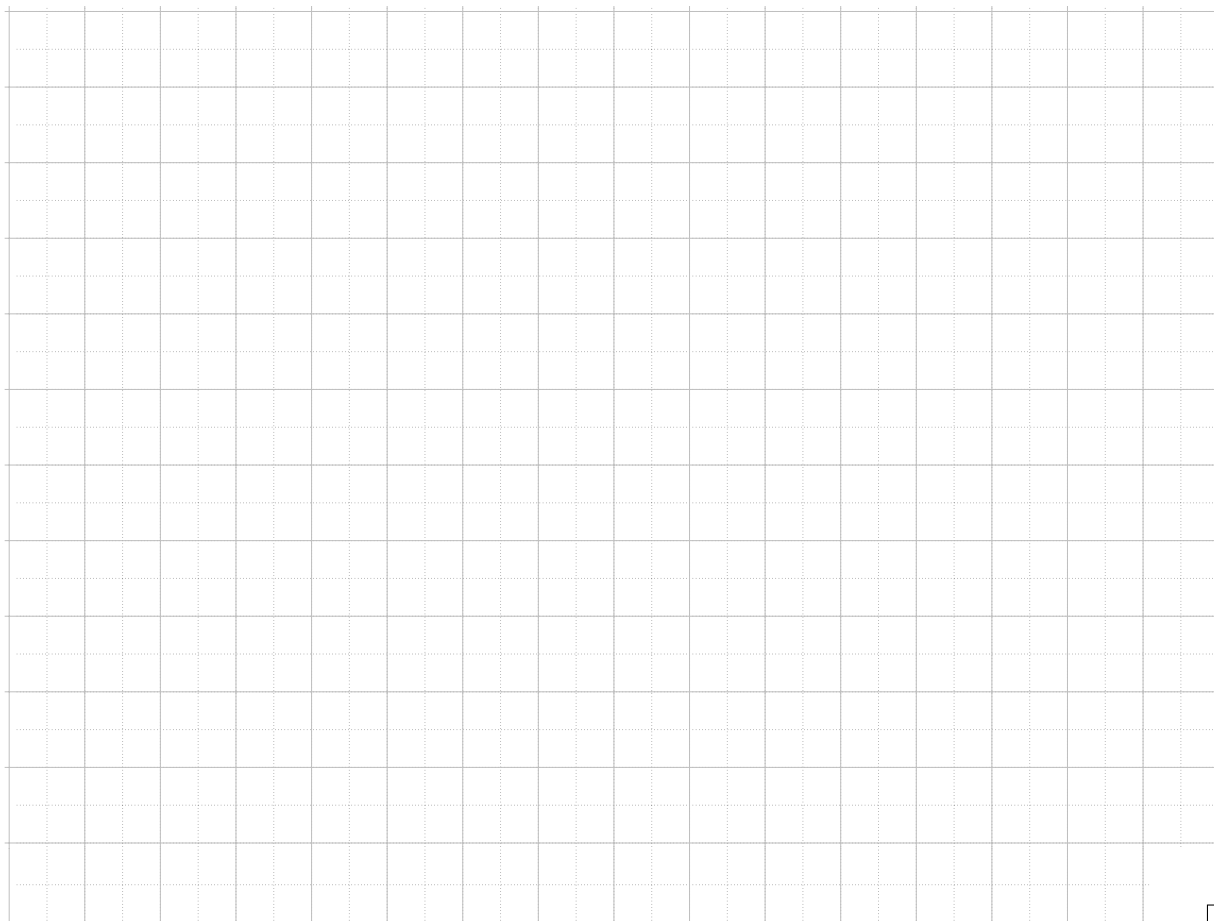




Lema 8.6. Sean $m \geq n \geq p \geq 1$ y $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ una transformación lineal inyectiva.

Si $\{\vec{v}_1, \dots, \vec{v}_p\} \subset \mathbb{R}^n$ es un conjunto L.I., entonces $\{T(\vec{v}_1), \dots, T(\vec{v}_p)\}$ es L.I.

Demostración.



□

Ejemplo:



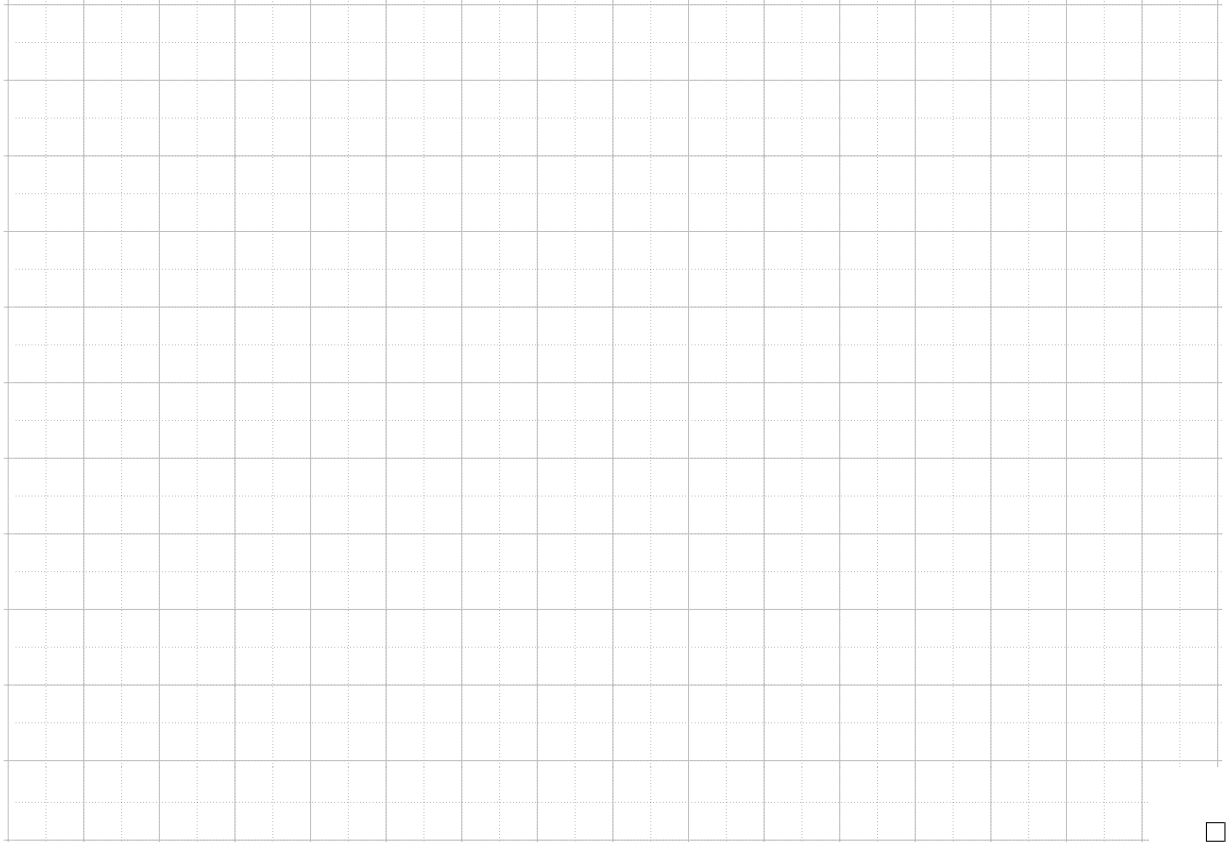


Teorema 8.7. Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ con $m = n$ una transformación lineal.

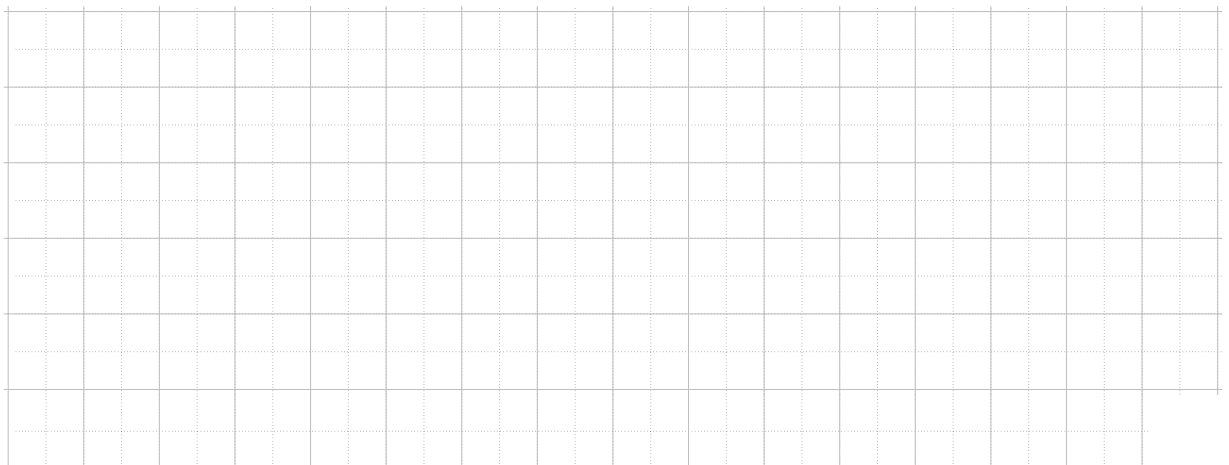
1. Si T es inyectiva, entonces T es sobre (supra).
2. Si T es sobre, entonces T es inyectiva.

Interpretación: Si la dimensión del dominio coincide con la del codominio, entonces una condición es suficiente.

Demostración.



Ejemplo:

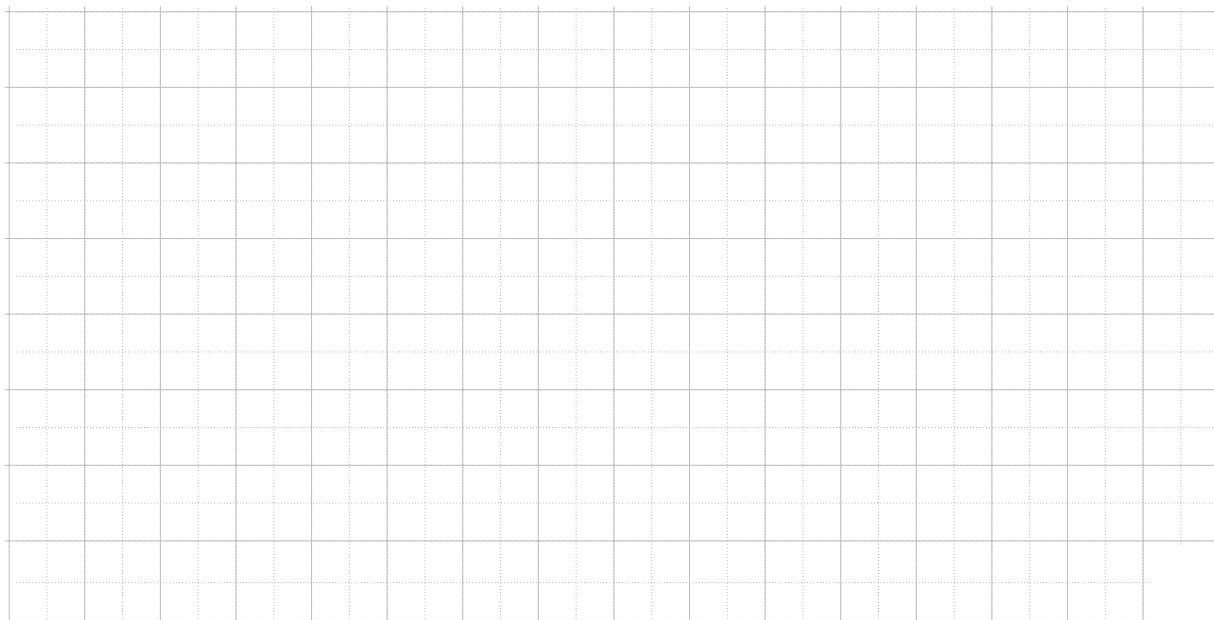




Definición 8.4. Sean $(\mathcal{V}, +, \cdot)$ y $(\mathcal{W}, +, \cdot)$ dos espacios vectoriales.

Una transformación lineal $T: \mathcal{V} \rightarrow \mathcal{W}$ es **biyectiva** si T es inyectiva & sobre.

Ejemplo:



Resumen:

Sea $A \in \mathbb{R}^{n \times n}$ una matriz. Entonces, las siguientes afirmaciones son equivalentes.

- A tiene inversa, es decir, A es invertible.
- $\det(A) \neq 0$
- renglones L.I.
- columnas L.I.
- $A\vec{x} = \vec{0}$ tiene solución única, la trivial $\vec{x} = \vec{0}$ ($\text{Null}(A) = \{\vec{0}\}$).
- $A\vec{x} = \vec{b}$ tiene solución para todo $\vec{b} \in \mathbb{R}^n$ ($\text{Img}(A) = \mathbb{R}^n$).
- existe una sucesión de operaciones elementales por filas que convierte A en la matriz identidad I , es decir, $FER(A) = I$.
- La transformación lineal $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ dada por $T(\vec{x}) = A\vec{x}$ es inyectiva, supra. y biyectiva.

8.4. Composición e inversas

Aplicaciones de composiciones de transformaciones lineales:

- “cifrar y descifrar” sucesiones de valores con transformaciones lineales (biyectivas).
- En robótica una composición describe como “mover un brazo de un robot”.
- Aprendizaje de máquina.
- Definir la multiplicación de matrices.



Definición 8.5. Sean $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $S: \mathbb{R}^m \rightarrow \mathbb{R}^p$ dos transformaciones lineales. Entonces, la **composición** S después de T es la función $(S \circ T): \mathbb{R}^n \rightarrow \mathbb{R}^p$ que para cada $\vec{x} \in \mathbb{R}^n$ da el resultado

$$(S \circ T)(\vec{x}) := S(T(\vec{x})).$$

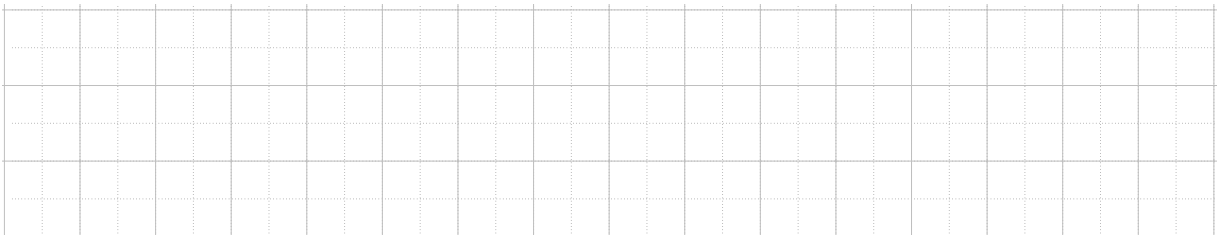
En palabras: “ $S \circ T$ ” se lee “ S después de T ”.

Dibujo:



Ejemplo: Determine $S \circ T$ para $T: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ y $S: \mathbb{R}^3 \rightarrow \mathbb{R}^4$ definidas por

$$T(\vec{x}) = \begin{pmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{pmatrix} \quad \& \quad S(\vec{u}) = \begin{pmatrix} 2u_1 + u_3 \\ 3u_2 - u_3 \\ u_1 - u_2 \\ u_1 + u_2 + u_3 \end{pmatrix}$$





Ejemplo: Compruebe que la matriz estándar de $S \circ T$ es dada por el producto $A_S A_T$ donde A_S, A_T son las matrices estándar de las transformaciones S y T , respectivamente.





Lema 8.8. Sean $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ y $S: \mathbb{R}^m \rightarrow \mathbb{R}^p$ dos transformaciones lineales. Entonces, la composición $(S \circ T): \mathbb{R}^n \rightarrow \mathbb{R}^p$ es una transformación lineal.

Demostración.



□

8.4.1. Inversas



Definición. La **identidad** es la función $I: \mathcal{V} \rightarrow \mathcal{V}$ que da $I(\vec{x}) = \vec{x}$ para toda $\vec{x} \in \mathcal{V}$.

Notar que esto requiere que el dominio es igual al codominio.



Lema 8.9. Sea $T: \mathbb{R}^n \rightarrow \mathbb{R}^n$ una transformación lineal biyectiva y $A_T \in \mathbb{R}^{n \times n}$ su matriz estándar. Entonces, A_T es invertible y la transformación lineal $S: \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por $S(\vec{x}) = A_T^{-1}\vec{x}$ cumple $(S \circ T) = I = (T \circ S)$, es decir, la composición de S con T da la identidad.

Nota: La función S también se denota por T^{-1} ya que es la transformación **inversa** de T .

Demostración.



□

Ejemplo: Encuentre la transformación inversa S asociada a $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dada por

$$T(\vec{x}) = \begin{pmatrix} \cos \alpha & -\operatorname{sen} \alpha \\ \operatorname{sen} \alpha & \cos \alpha \end{pmatrix} \vec{x}.$$



Ejemplo: Encuentre la transformación inversa de $T: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ dada por

$$T(\vec{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \cos \alpha & -\operatorname{sen} \alpha \\ 0 & \operatorname{sen} \alpha & \cos \alpha \end{pmatrix} \vec{x}.$$



8.4.2. Ap. Transporte seguro de datos binarios

[David Poole] “[Poo11] Álgebra lineal. Una Introducción moderna”, 3a edición (páginas 252–255).

El *código de Hamming* consiste de 3 transformaciones lineales:

- El transmisor usa $G: \mathbb{Z}_2^k \rightarrow \mathbb{Z}_2^n$ con $n > k$ para codificar un vector binario con k entradas. El resultado $\vec{c}_e = G\vec{x}$ se transmite y será recibido en otra máquina (receptor).
- El receptor tiene una transformación $P: \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^{n-k}$ (para verificar datos recibidos) y otra $Q: \mathbb{Z}_2^n \rightarrow \mathbb{Z}_2^k$ (para recuperar la $\vec{x}$).
- La transformación P permite detectar a un solo error en el vector recibido $\vec{c}_r$.



8.4.3. Ap. Mover brazos en robótica

Para interesados véase las [Poo11, paginas 232–235].



El producto punto en $\mathbb{R}^n$ es compatible con la norma euclidiana, ya que la longitud de cada vector se puede medir con el producto punto, es decir para $\vec{u} \in \mathbb{R}^n$ tenemos

$$\|\vec{u}\|_e = \sqrt{\sum_{i=1}^n (u_i)^2} = \sqrt{\langle \vec{u}, \vec{u} \rangle}. \quad (9.1)$$

Interpretaciones:

- $\|\vec{u}\|_e$ es la longitud de la línea que conecta el origen $\vec{0}$ con la ubicación $\vec{u}$.
- La **distancia** entre dos vectores $\vec{u}$ & $\vec{v}$ es la longitud de $(\vec{u} - \vec{v})$:

$$\|\vec{u} - \vec{v}\|_e = \sqrt{(u_1 - v_1)^2 + \dots + (u_n - v_n)^2}$$

- Para $\vec{u} \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$ la definición implica que $\|\alpha \vec{u}\|_e = |\alpha| \cdot \|\vec{u}\|_e$.

Ejemplo: $\vec{u} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$.



9.1. Ángulos y ortogonalidad



Teorema 9.3 (Cauchy-Schwarz-Bunyakovsky). Para cualesquiera dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ se tiene:

$$|\langle \vec{u}, \vec{v} \rangle| \leq \|\vec{u}\|_e \|\vec{v}\|_e$$

Demostración. La prueba se divide en 2 casos.

Caso 1: Sean $\vec{u}, \vec{v}$ L.D., es decir, $\vec{u} = \alpha \vec{v}$ para un escalar $\alpha \in \mathbb{R}$. Entonces

$$|\langle \vec{u}, \vec{v} \rangle| = |\langle \alpha \vec{v}, \vec{v} \rangle| = |\alpha \langle \vec{v}, \vec{v} \rangle| = |\alpha| \|\vec{v}\|_e^2 = |\alpha| \|\vec{v}\|_e \|\vec{v}\|_e = \|\alpha \vec{v}\|_e \|\vec{v}\|_e = \|\vec{u}\|_e \|\vec{v}\|_e$$



Caso 2: Sean $\vec{u}, \vec{v}$ L.I., es decir, $\vec{u} \neq \alpha \vec{v}$ para cualesquier escalar $\alpha \in \mathbb{R}$

$$\begin{aligned}
 0 < \|\vec{u} + \alpha \vec{v}\|_e^2 &= \langle \vec{u} + \alpha \vec{v}, \vec{u} + \alpha \vec{v} \rangle \\
 &= \langle \vec{u}, \vec{u} + \alpha \vec{v} \rangle + \langle \alpha \vec{v}, \vec{u} + \alpha \vec{v} \rangle \\
 &= \langle \vec{u}, \vec{u} \rangle + \langle \vec{u}, \alpha \vec{v} \rangle + \langle \alpha \vec{v}, \vec{u} \rangle + \langle \alpha \vec{v}, \alpha \vec{v} \rangle \\
 &= \langle \vec{u}, \vec{u} \rangle + 2\alpha \langle \vec{u}, \vec{v} \rangle + \alpha^2 \langle \vec{v}, \vec{v} \rangle \quad (\star)
 \end{aligned}$$

El lado derecho es un polinomio de grado 2 en α que siempre es mayor que cero. Su mínimo lo toma donde su derivada es cero, es decir, el valor α_* debe cumplir

$$0 = 2 \langle \vec{u}, \vec{v} \rangle + 2\alpha_* \langle \vec{v}, \vec{v} \rangle \iff \alpha_* = -\frac{\langle \vec{u}, \vec{v} \rangle}{\langle \vec{v}, \vec{v} \rangle}.$$

En la desigualdad $(\star)$, sustituimos α con α_* para llegar a:

$$0 < \langle \vec{u}, \vec{u} \rangle - \frac{\langle \vec{u}, \vec{v} \rangle^2}{\langle \vec{v}, \vec{v} \rangle} \iff \langle \vec{u}, \vec{v} \rangle^2 < \|\vec{u}\|_e^2 \|\vec{v}\|_e^2, .$$

□



Corolario 9.4. Dados dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ existe un único ángulo $\alpha \in [0, \pi]$ tal que

$$\langle \vec{v}, \vec{u} \rangle = \langle \vec{u}, \vec{v} \rangle = \cos(\alpha) \|\vec{u}\|_e \|\vec{v}\|_e .$$

Demostración.

Por el teorema 9.3 sabemos que

$$-1 \leq \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|_e \|\vec{v}\|_e} \leq 1 .$$

Por $\cos(0) = 1$, $\cos(\pi) = -1$ y dado que $\cos: [0, \pi] \rightarrow [-1, 1]$ es continua, el teorema del valor intermedio justifica que

$$\forall z \in [-1, 1] \exists \alpha \in [0, \pi] \quad \text{tal que} \quad \cos(\alpha) = z.$$

Además, el α es único ya que la función $\cos$ es estrictamente monótona. Tomar $z = \frac{\langle \vec{u}, \vec{v} \rangle}{\|\vec{u}\|_e \|\vec{v}\|_e}$ muestra el resultado. □

9. Un espacio vectorial con producto punto

Ahora que tenemos ángulos entre vectores podemos definir ortogonalidad.

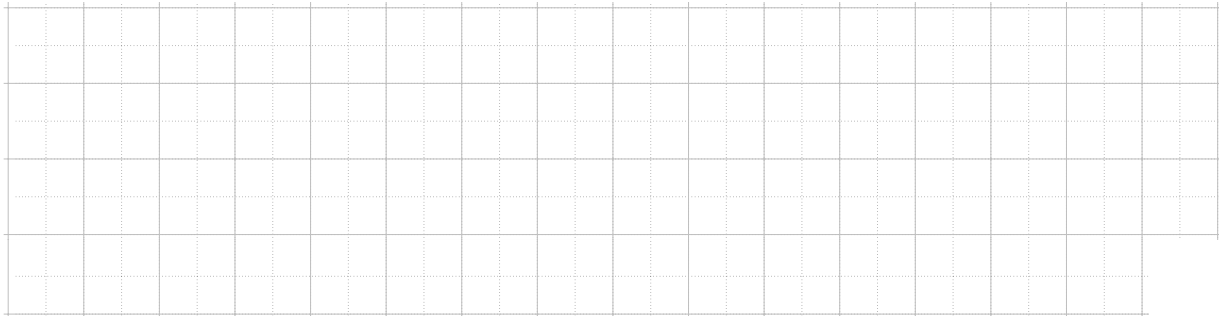


Definición 9.2. Dos vectores $\vec{u}, \vec{v} \in \mathbb{R}^n$ son **ortogonales** si

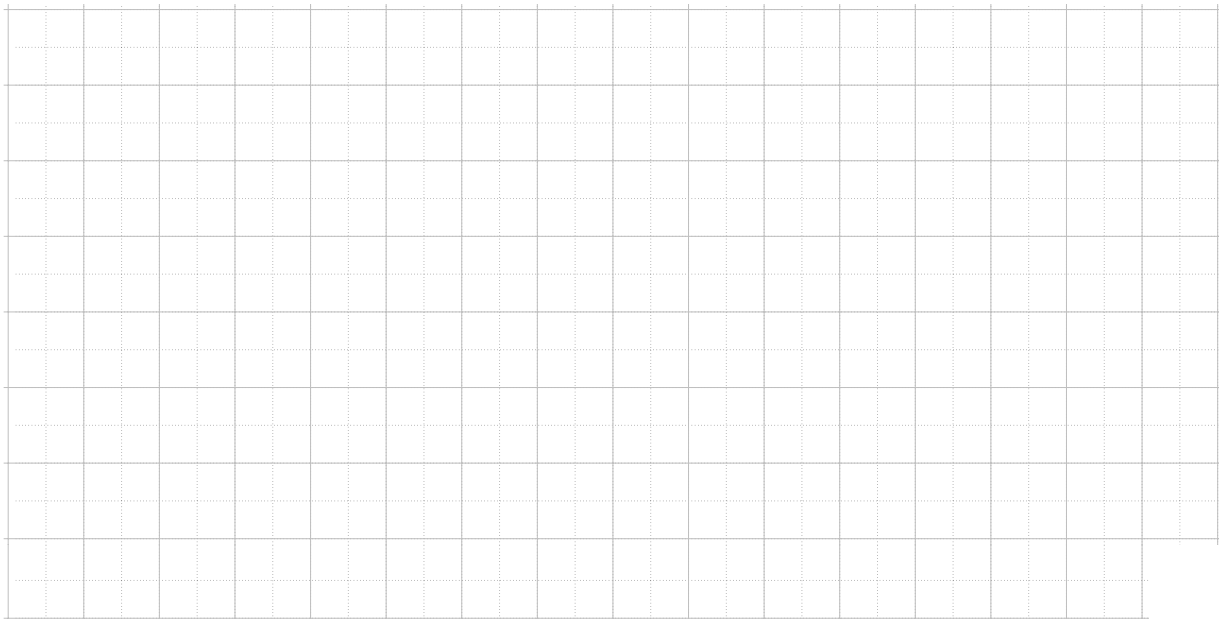
$$\langle \vec{u}, \vec{v} \rangle = 0$$

Si $\vec{u}, \vec{v} \in \mathbb{R}^n \setminus \{\vec{0}\}$, entonces $\alpha \in [0, \pi]$ y $\cos \alpha = 0 \implies \alpha = \pi/2$.

Ejemplo: La base canónica:



Ejemplo: Eigenvectores de una matriz simétrica (para distintos eigenvalores).



Definición 9.3. Un vector $\vec{w} \in \mathbb{R}^n$ es un vector **unitario** (o **normalizado**) si tiene longitud uno, i.e. $\|\vec{w}\| = 1$.

Observación: Cada vector $\vec{u} \in \mathbb{R}^n \setminus \{\vec{0}\}$ es producto de su longitud y un vector unitario.



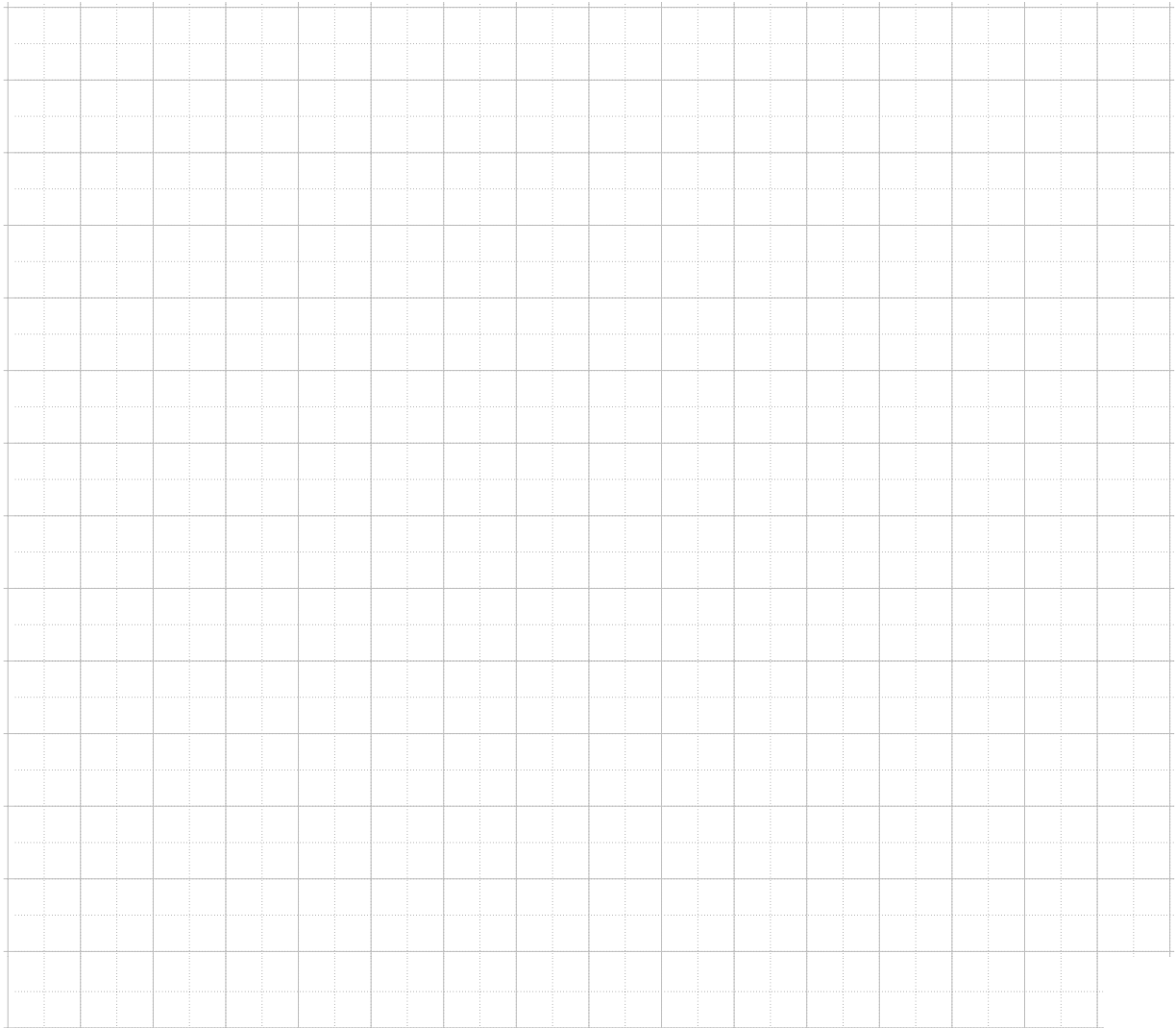


Definición 9.4. Una matriz $Q \in \mathbb{R}^{n \times n}$ es **ortogonal** si su transpuesta es su inversa, *i.e.*

$$Q^T Q = I \quad \& \quad Q Q^T = I$$

En palabras: Las columnas de Q son ortogonales y vectores unitarios (tienen longitud uno).

Ejemplo: Eigenvectores normalizados de una matriz simétrica.



Interpretación: Ahora podemos entender el teorema 7.16.

El teorema dice que cada matriz simétrica en $\mathbb{R}^{n \times n}$ tiene eigenvalores reales, es diagonalizable y que existe una matriz ortogonal de eigenvectores, es decir $AV = VD$ y $V^T V = I$.

En palabras: Las columnas de V son ortogonales y tienen longitud uno (unitarias).



9.2. Transformaciones rígidas

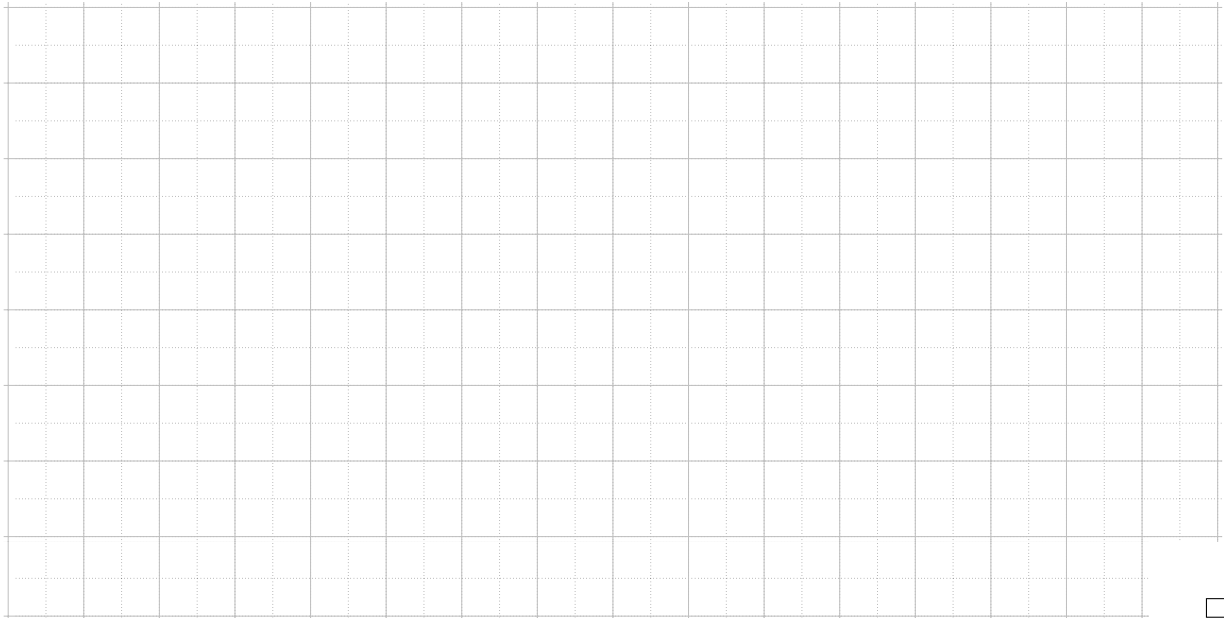
Las matrices ortogonales preservan distancias y ángulos.

Esto se debe a $\langle \vec{u}, \vec{u} \rangle = \|\vec{u}\|_e^2$.

Propiedades: De matrices ortogonales $Q \in \mathbb{R}^{n \times n}$, es decir $Q^\top Q = I$.

🍃 | “ Q preserva distancias en la norma euclidiana”, es decir $\|Q\vec{x} - Q\vec{y}\|_e = \|\vec{x} - \vec{y}\|_e$:

Demostración.

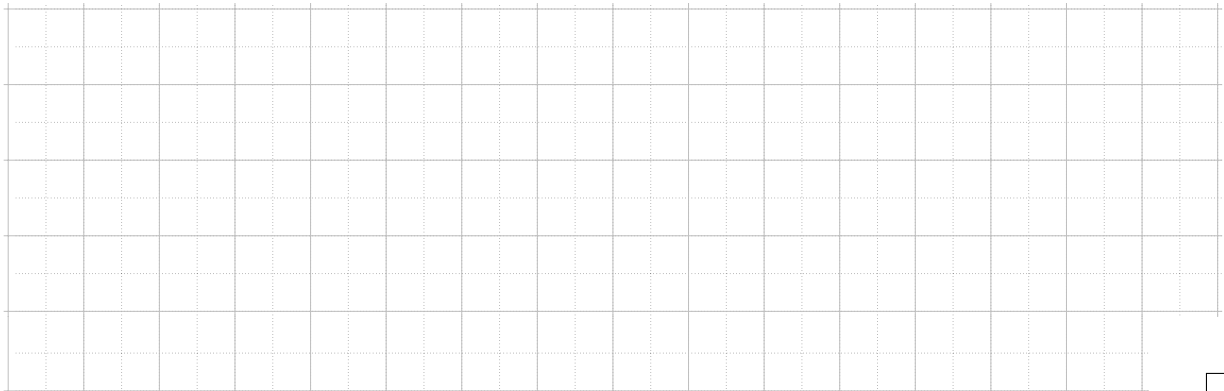


□

🍃 | “ Q preserva la distancia al origen: $\|Q\vec{x}\|_e = \|\vec{x}\|_e$.”

🍃 | “ Q preserva ángulos”:

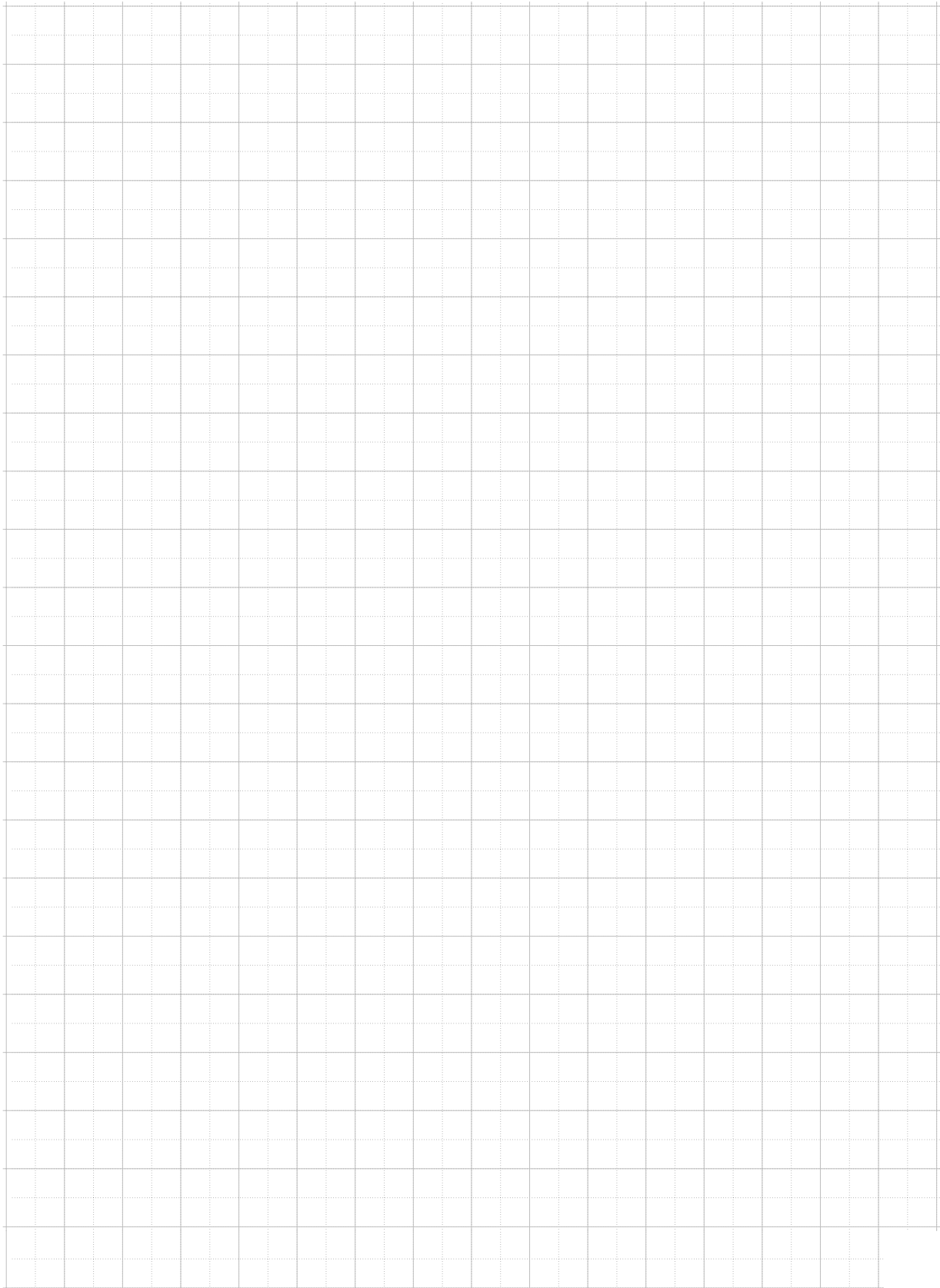
Demostración.



□

🍃 | *Conclusión:* Aplicar una transformación lineal y ortogonal no cambia ángulos ni distancias.

Ejemplos de matrices: (Permutaciones / reflexiones / rotaciones)



9.3. Matrices positivas definidas

Aplicaciones en: Estadística, probabilidad, optimización, etc..



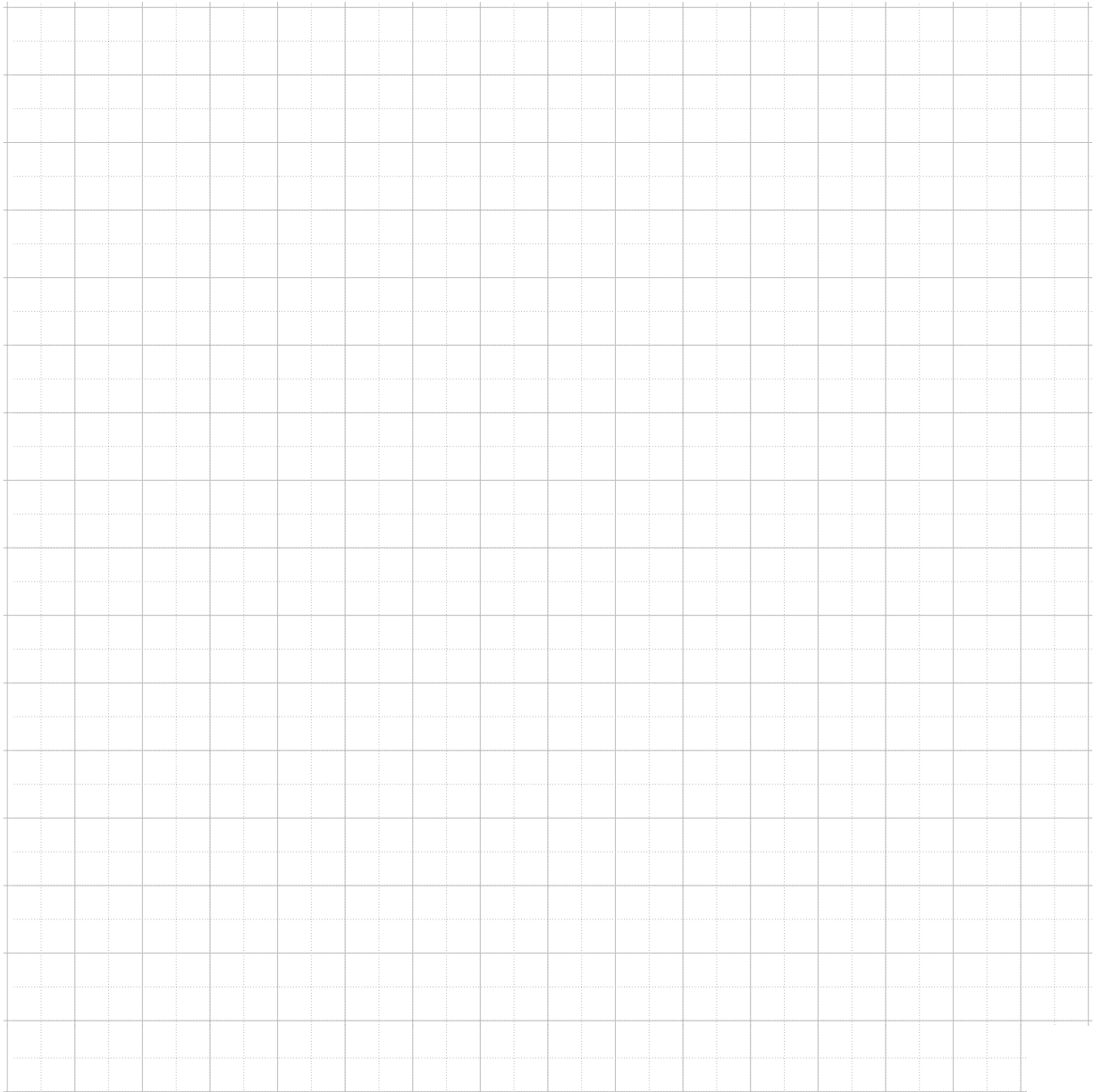
Definición. Una matriz real cuadrada $B \in \mathbb{R}^{n \times n}$ es **positiva definida** cuando

$$\vec{x}^\top B \vec{x} > 0 \quad \text{para todo} \quad \vec{x} \in \mathbb{R}^{n \times 1} \setminus \{\vec{0}\}$$

y **positiva semi-definida** cuando

$$\vec{x}^\top B \vec{x} \geq 0 \quad \text{para todo} \quad \vec{x} \in \mathbb{R}^{n \times 1}.$$

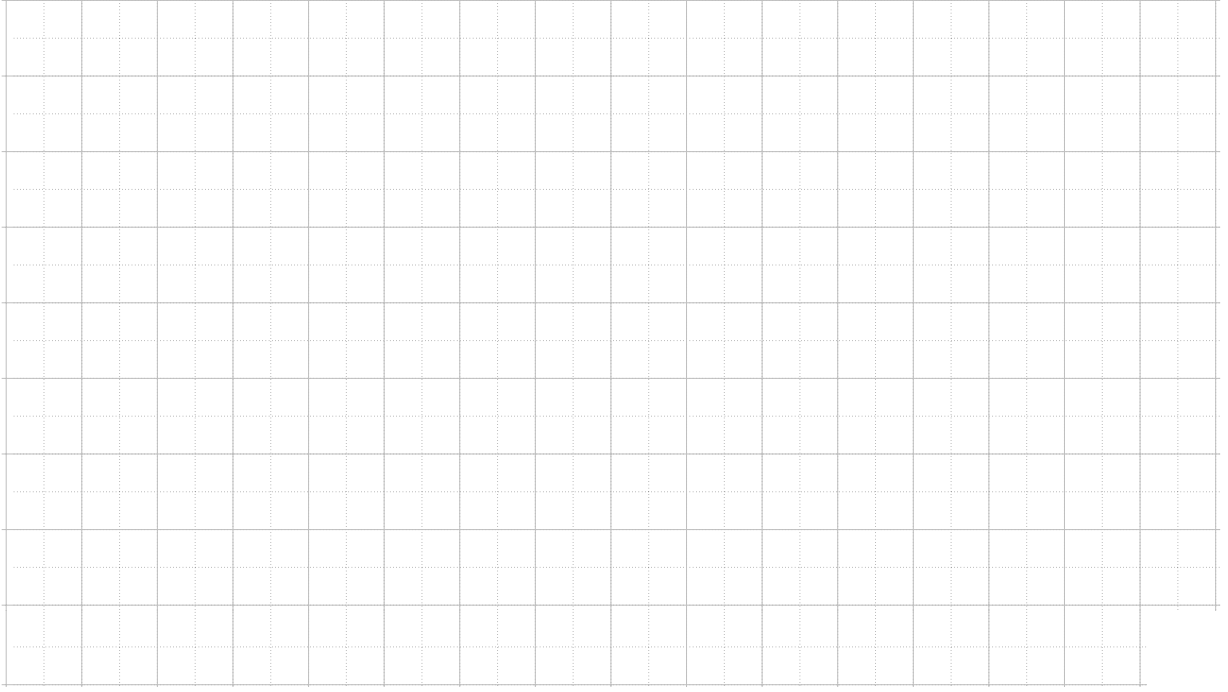
Ejemplos “+”:



 | **Lema 9.5.** Si una matriz $B \in \mathbb{R}^{n \times n}$ es positiva definida, entonces B es invertible.

Demostración. Primero mostramos que el sistema $B\vec{x} = \vec{0}$ tiene solución única $\vec{x} = \vec{0}$. Procedemos por contradicción: Supongamos que el sistema tiene otra solución $\vec{z} \neq \vec{0}$. Esto contradice que B es positiva definida, ya que entonces $0 < \vec{z}^\top B\vec{z} = \vec{z}^\top (B\vec{z}) = \vec{z}^\top (\vec{0}) = 0$. Concluimos que el sistema $B\vec{x} = \vec{0}$ tiene solución única y por lo tanto teorema 5.2 implica que B tiene inversa. $\square$

Ejemplos “–”:



Proposición 9.6. Sea $B \in \mathbb{R}^{n \times n}$ una matriz simétrica. Entonces, B es positiva definida si y solo si todos sus eigenvalores son positivos.

Demostración. Por el teorema 7.16 los eigenvalores de B son reales.

Supongamos primero que B es positiva definida.

Sea $\vec{v} \in \mathbb{R}^{n \times 1} \setminus \{\vec{0}\}$ un eigenvector de B tal que $\|\vec{v}\|_e = 1$.

$$0 < \vec{v}^\top B\vec{v} = \vec{v}^\top (B\vec{v}) = \vec{v}^\top (\lambda\vec{v}) = \lambda\vec{v}^\top \vec{v} = \lambda\|\vec{v}\|_e^2 = \lambda.$$

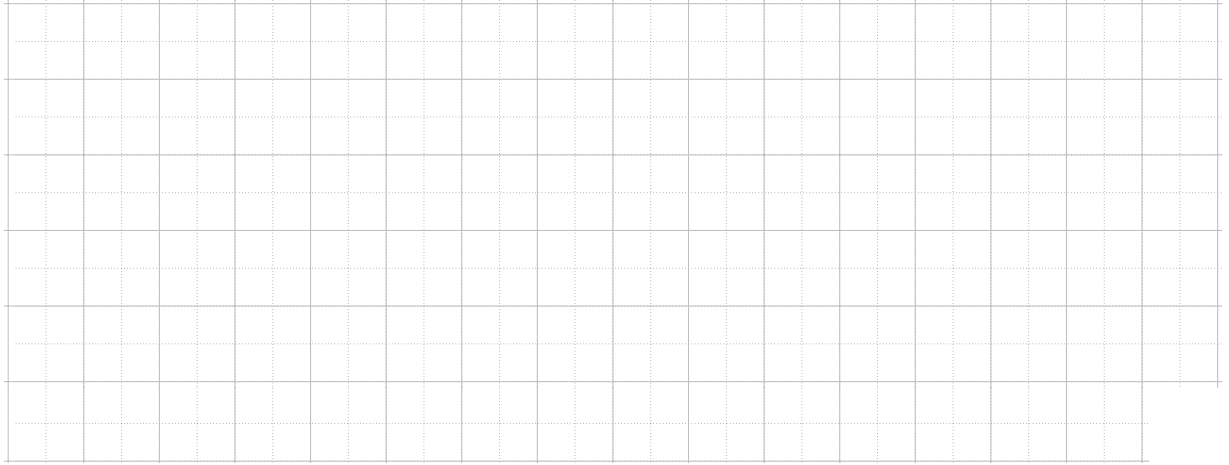
Concluimos que cada eigenvalor es real y positivo.

Por otro lado, supongamos que todos los eigenvalores de B son positivos. La simetría de B y el teorema 7.16 implican que en $\mathbb{R}^n$ existe una base de eigenvectores de B , es decir, cualquier vector $\vec{x} \neq \vec{0}$ es combinación lineal de ellos: $\vec{x} = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \dots + \alpha_n\vec{v}_n = V\vec{\alpha}$ con $\vec{\alpha} \neq \vec{0}$. El mismo teorema permite escribir $B = V\Lambda V^\top$ ya que $V^{-1} = V^\top$. Por lo tanto,

$$\vec{x}^\top B\vec{x} = (V\vec{\alpha})^\top V\Lambda V^\top (V\vec{\alpha}) = \vec{\alpha}^\top V^\top V\Lambda V^\top V\vec{\alpha} = \vec{\alpha}^\top \Lambda \vec{\alpha} = \sum_{i=1}^n \lambda_i \alpha_i^2 > 0.$$

Concluimos que B es positiva definida. $\square$

Ejemplos “ $\pm$ ”:



El producto punto es un caso especial del siguiente lema con $B = I$.



Lema 9.7. Sea $B \in \mathbb{R}^{n \times n}$ simétrica positiva definida, entonces el producto interior dado por $\langle \vec{x}, \vec{y} \rangle_B := \vec{x}^\top B \vec{y}$ satisface las propiedades de los lemas 9.1, 9.2 y

$$|\langle \vec{x}, \vec{y} \rangle_B| \leq \sqrt{\langle \vec{x}, \vec{x} \rangle_B} \sqrt{\langle \vec{y}, \vec{y} \rangle_B}$$

Demostración. Por teorema 7.16 la matriz simétrica B tiene una matriz de eigenvectores V que es ortogonal pues $V^\top V = I$ y $B = V \Lambda V^\top$. Por proposición 9.6 los eigenvalores de B son positivos y los podemos acomodar en la matriz diagonal $\Lambda = \Lambda^{1/2} \Lambda^{1/2}$. Como consecuencia B tiene una “raíz” simétrica, es decir

$$B = V \Lambda V^\top = (V \Lambda^{1/2} V^\top)(V \Lambda^{1/2} V^\top) = B^{1/2} B^{1/2}.$$

Usamos esta raíz y el teorema 9.3 para concluir:

$$\begin{aligned} \langle \vec{x}, \vec{y} \rangle_B &= \vec{x}^\top B \vec{y} = \vec{x}^\top B^{1/2} B^{1/2} \vec{y} = (B^{1/2} \vec{x})^\top B^{1/2} \vec{y} \\ &= \langle B^{1/2} \vec{x}, B^{1/2} \vec{y} \rangle \\ &\leq \|B^{1/2} \vec{x}\|_e \|B^{1/2} \vec{y}\|_e \\ &= \sqrt{\langle B^{1/2} \vec{x}, B^{1/2} \vec{x} \rangle} \sqrt{\langle B^{1/2} \vec{y}, B^{1/2} \vec{y} \rangle} \\ &= \sqrt{\vec{x}^\top B^{1/2} B^{1/2} \vec{x}} \sqrt{\vec{y}^\top B^{1/2} B^{1/2} \vec{y}} \\ &= \sqrt{\vec{x}^\top B \vec{x}} \sqrt{\vec{y}^\top B \vec{y}} = \sqrt{\langle \vec{x}, \vec{x} \rangle_B} \sqrt{\langle \vec{y}, \vec{y} \rangle_B} \end{aligned}$$

Con eso se concluye la demostración. □



“This is the end. Hold your breath and count to ten.” [Adele, 2012]

Bibliografía

- [Lay07] David C. Lay. *Álgebra Lineal y sus Aplicaciones*. 3a edición. 2007.
- [MR20] Miguel Angel Mota G. e I. Beatriz Rumbos P. *Álgebra Lineal en $\mathbb{R}^n$* . 1a edición. 2020.
- [Poo11] David Poole. *Álgebra Lineal. Una Introducción moderna*. 3a edición. 2011.