PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

Fundamentos Matemáticos de Inteligencia Artificial

Clase 2

Manuel A. Sánchez

2024.03.25

Algebra Lineal: sistemas lineales

Introducción

Introducción y motivación

Las matrices son conjuntos de números organizados en filas y columnas, y son esenciales para modelar y resolver una variedad de problemas en ciencia de datos e inteligencia artificial. Desde la clasificación de imágenes hasta el procesamiento de lenguaje natural, las matrices proporcionan una forma elegante y poderosa de representar y operar con datos complejos.
Conoceremos a continuacion desde operaciones basicas, algoritmos y aplicaciones que se desarrollan con matrices.

Matrices

Matrices

Definición. Un matriz $A \in \mathbb{R}^{m\times n}$ es un arreglo rectangular de m-filas y n-columnas, formado por números reales.

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

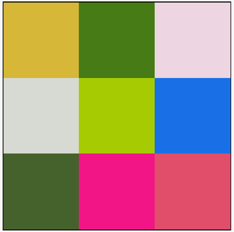
$a_{ij} \in \mathbb{R}$, $1 \leq i \leq m$, $1 \leq j \leq n$
son las componentes de A

Representación por columna.

$$x_1, x_2, \dots, x_n \in \mathbb{R}^m, \text{ entonces } X = \begin{bmatrix} x_1 \,|\, x_2 \,|\, \dots \,|\, x_n \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Ejemplos matrices. Representación de imágenes.

$$R = \begin{bmatrix} 214 & 71 & 237 \\ 215 & 167 & 25 \\ 70 & 242 & 224 \end{bmatrix}, \quad G = \begin{bmatrix} 183 & 123 & 213 \\ 217 & 203 & 112 \\ 98 & 21 & 78 \end{bmatrix}, \quad B = \begin{bmatrix} 55 & 21 & 225 \\ 211 & 2 & 230 \\ 44 & 134 & 106 \end{bmatrix}$$



Ejemplos matrices. Datos de precipitaciones.

Table: Datos de precipitación por ubicación y mes

Ubicación	**Enero**	**Febrero**	**Marzo**	**...**	**Noviembre**	**Diciembre**
Ubicación 1	10	15	20	...	7	5
Ubicación 2	5	8	10	...	4	3
...	...	...	...	...	...	...
Ubicación n	12	18	25	...	5	7

Ejemplos matrices. Historial de compras.

Table: Historial de compras de clientes y productos

Cliente	**Producto A**	**Producto B**	**Producto C**	**...**	**Producto N**
Cliente 1	3	0	1	...	2
Cliente 2	0	2	0	...	1
Cliente 3	1	1	2	...	0
...	...	...	...	...	...
Cliente M	2	0	0	...	3

Matrices conocidas

- La matriz nula. $A = 0 \in \mathbb{R}^{m\times n}$, $a_{ij} = 0$, $1 \leq i \leq m$, $1 \leq j \leq n$.
- La matriz identidad. $I \in \mathbb{R}^{n\times n}$, $I_{ij} = \begin{cases} 1, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$
- Matriz diagonal. $A \in \mathbb{R}^{n\times n}$, $A_{ij} = \begin{cases} a_{ii}, & \text{si } i = j, \\ 0, & \text{si } i \neq j. \end{cases}$
- Matriz triangular. $A \in \mathbb{R}^{m\times n}$.

$$\begin{aligned} \text{Superior:} \quad & a_{ij} = 0, \quad i > j, \\ \text{Inferior:} \quad & a_{ij} = 0, \quad i < j. \end{aligned}$$

- Matriz sparse. $A \in \mathbb{R}^{m\times n}$, nnz$(x)$ « mn.

Operaciones matriciales

Transpuesta. Si $A \in \mathbb{R}^{m\times n}$, entonces su matriz transpuesta $A^\top \in \mathbb{R}^{n\times m}$ se define por

$$(A^\top)_{ji} = (A)_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Además, la matriz $A \in \mathbb{R}^{n\times n}$ se dice **simétrica** si $A = A^\top$.
Adición de matrices. Sean $A, B \in \mathbb{R}^{m\times n}$, entonces la matrix $C = (A+B) \in \mathbb{R}^{m\times n}$ se define por

$$c_{ij} = (A+B)_{ij} = a_{ij} + b_{ij}, \qquad 1 \leq i \leq m, 1 \leq j \leq n.$$

Multiplicación por escalar. Sea $A \in \mathbb{R}^{m\times n}$ and $r \in \mathbb{R}$. Entonces la matriz $D = (rA) \in \mathbb{R}^{m\times}$ se define por:

$$d_{ij} = (rA)_{ij} = ra_{ij}, \qquad 1 \leq i \leq m, 1 \leq j \leq n.$$

Propiedades de la adición matricial

Propiedad de adición	Representación matemática
Conmutatividad	$A + B = B + A$
Asociatividad	$(A + B) + C = A + (B + C)$
Elemento identidad	$(A + 0) = A$
Transpuesta de la suma	$(A + B)^\top = A^\top + B^\top)$

Propiedades multiplicación por escalar

Propiedades multiplicación por escalar	Representación matemática
Conmutatividad	$(rs)A = r(sA)$
Asociatividad	$(r+s)A = rA + rB$
Transpuesta	$(rA)^\top = rA^\top$

Norma matricial

Definición. La norma de Frobenius (una de las normas matriciales) esta definida por:

$$\begin{aligned} \|\cdot\|_F : \mathbb{R}^{m\times n} &\longrightarrow \mathbb{R}_0^+ \\ A &\longrightarrow \|A\|_F := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} \end{aligned}$$

Ejercicio. Muestre que:

$$\begin{aligned} &\|A\|_F = \|A^\top\|_F \\ &\|A\|_F^2 = \|a_1\|^2 + \|a_2\|^2 + \ldots + \|a_n\|^2 \end{aligned}$$

Producto matriz-vector

Definición. Sea $A \in \mathbb{R}^{m\times n}$ y $x \in \mathbb{R}^n$, entonces el producto matriz-vector se define por:

$$A = [a_1 \,|\, a_2 \,|\, \dots \,|\, a_n], y = Ax \in \mathbb{R}^m$$

$$y_i = \sum_{j=1}^{n} a_{ij}x_j = a_1x_1 + ... + a_nx_n$$

Ejemplo.

$$\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 4 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$

Transformaciones lineales

Transformaciones lineales

Propiedad del producto matriz-vector	Representación matemática
Distributividad con respecto a la suma de vectores	$A(x+y) = Ax + Ay$
Distributividad con respecto a la suma de matrices	$(A+B)x = Ax + Bx$
Compatibilidad con la multiplicación por escalar	$A(\alpha x) = \alpha Ax = (\alpha A)x$

Transformaciones lineales. Representaciones

Sea $A \in \mathbb{R}^{m \times n}$ y $x \in \mathbb{R}^n$. Entonces:

- representación por columnas

$$A = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \quad \text{con} \quad a_1, \dots, a_n \in \mathbb{R}^m \quad \Longrightarrow \quad Ax = x_1 a_1 + \dots + x_n a_n$$

- representación por filas

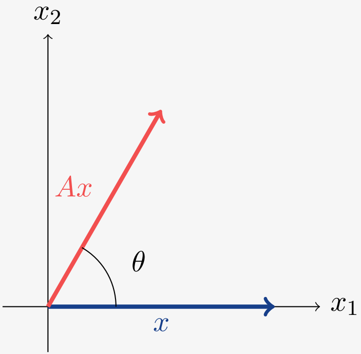
$$A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix} \quad \text{con} \quad a_1, \dots, a_m \in \mathbb{R}^n \quad \Longrightarrow \quad Ax = \begin{bmatrix} a_1^\top x \\ \vdots \\ a_m^\top x \end{bmatrix} = \begin{bmatrix} a_1 \cdot x \\ \vdots \\ a_m \cdot x \end{bmatrix}$$

Transformaciones lineales. Ejemplo

En $\mathbb{R}^2$ la matriz

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

representa una **rotación en un ángulo** θ en sentido antihorario



Ejemplo matriz de diferencias

Sea $D \in \mathbb{R}^{n-1 \times n}$, $b \in \mathbb{R}^n$ definidas por

$$D_{ij} = \begin{cases} -1, & \text{si } i = j, \\ 1, & \text{si } j = i+1, \\ 0, & \text{en otro caso}, \end{cases}, \qquad b_i = f\left(\frac{i-1}{n-1}\right) = \sin\left(2\pi \frac{i-1}{n-1}\right), \quad 1 \leq i \leq n$$

El espacio nulo y el espacio imagen de una matriz

Definición. El **espacio nulo** de A se define por:

$$\mathbf{nul}(A) := \{x \in \mathbb{R}^n : \; Ax = 0\} \subseteq \mathbb{R}^n.$$

Si $\mathbf{nul}(A) = \{0\}$, las columnas de A son **linealmente independientes**.

Definición. El **espacio imagen** de A se define por

$$\mathbf{im}(A) := \{Ax : \; x \in \mathbb{R}^n\} = \{\alpha_1 a_1 + \ldots + \alpha_n x_n : \; \alpha_1, \ldots, \alpha_n \in \mathbb{R}\} \subseteq \mathbb{R}^m$$

Si $\mathbf{im}(A) = \mathbb{R}^m$, decimos que A tiene **rango completo**.

Teorema rango-nulidad

Observación.

Si $A \in \mathbb{R}^{m\times n}$, sabemos que $\mathbf{nul}(A) \subset \mathbb{R}^n$ y que $\mathbf{im}(A) \subset \mathbb{R}^m$ y así

$$\dim \mathbf{nul}(A) \leq n \quad \text{y} \quad \dim \mathbf{im}(A) \leq m$$

El **teorema de rango-nulidad** nos dice que

$$n = \dim(\mathbf{nul}(A)) + \dim(\mathbf{im}(A)).$$

Ejemplo matriz-vector

Sean los vectores $x_1, x_2, ..., x_m \in \mathbb{R}^n$, aquí el índice de los vectores indica los estados de una variable en el tiempo.

Sistema dinámico lineal: $x_{t+1} = A_t x_t, \quad t = 1, 2, 3, ..$

Extensiones:

- $+$ input $\quad x_{t+1} = A_t x_t + B_t u_t + C_t, \quad t = 1, 2, 3, ..$
- Modelo de K-Markov

$$x_{t+1} = A_1 x_t + A_2 x_{t-1} + ... + A_K x_{t-K+1}, \quad t = K, K+1, ..$$

Dinámicas de población

Describir la evolución de la distribución de edad en una población,

$$x_i \in \mathbb{R}^{100},\ (x_t)_i : \quad \text{número de personas con edad } i-1 \text{ en año } t.$$

Tenemos la tasa de natalidad y mortalidad $b \in \mathbb{R}^{100}$, $d_t \in \mathbb{R}^{100}$

$(b)_i$: número promedio de nacimientos por persona de edad $(i-1)$.

$(d_t)_i$: fracción de las personas de edad $(i-1)$ que morirán ese año.

Construimos el **modelo**:

$$\begin{aligned}(x_{t+1})_1 &= b^\top x_t \\ (x_{t+1})_{i+1} &= (1-d)(x_t)_i, \quad 1 \le i \le 99\end{aligned}$$

Dinámicas de población

$$(x_{t+1})_1 = b^\top x_t$$
$$(x_{t+1})_{i+1} = (1-d)(x_t)_i, \quad 1 \le i \le 99$$

$$\boxed{x_{t+1} = Ax_t}$$

$$A = \begin{bmatrix} b_1 & b_2 & \dots & b_{99} & b_{100} \\ 1-d_1 & 0 & 0 & \dots & 0 \\ 0 & 1-d_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 1-d_{99} & 0 \end{bmatrix}$$

Dinámicas de población

$$
\begin{aligned}
(x_{t+1})_1 &= b^\top x_t \\
(x_{t+1})_{i+1} &= (1-d)(x_t)_i, \quad 1 \leq i \leq 99
\end{aligned}
$$

$$\boxed{x_{t+1} = Ax_t}$$

$$
A = \begin{bmatrix}
b_1 & b_2 & \dots & b_{99} & b_{100} \\
1-d_1 & 0 & 0 & \dots & 0 \\
0 & 1-d_2 & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & 0 \\
0 & \dots & 0 & 1-d_{99} & 0
\end{bmatrix}
$$

Predecir:

- población
- niños en edad escolar
- adultos en edad de jubilación

Dinámica de epidemias

Modelo de fracción de la población en: $x_t \in \mathbb{R}^4 \begin{cases} \text{Susceptible} \\ \text{Infectado} \\ \text{Recuperado} \\ \text{Fallecidos} \end{cases}$

Suponemos en nuestro modelo simple que cada día:

- 5% de la población susceptible tendrá la enfermedad
- 1% de la población infectada fallecerá
- 10% de la población infectada se recuperará y quedará inmune
- 4% de la población infectada se recuperará pero no quedará inmune

Dinámica de epidemias

$$x_{t+1} = Ax_t \quad \iff \quad x_{t+1} = \begin{bmatrix} 0.95 & 0.04 & 0 & 0 \\ 0.05 & 0.85 & 0 & 0 \\ 0 & 0.1 & 1 & 0 \\ 0 & 0.01 & 0 & 1 \end{bmatrix} x_t$$

Aproximación lineal por expansión de Taylor

Sea la función $f : \mathbb{R}^n \to \mathbb{R}^m$. La expansión lineal en serie de Taylor en $x = z \in \mathbb{R}^n$ de $f(x) = (f_1(x), ..., f_m(x))$

$$\hat{f}_i(x) \approx f_i(z) + \nabla f_i(z)^\top (x - z)$$

donde el gradiente de f_i está definido por el vector

$$\nabla f_i(z) = \begin{bmatrix} \frac{\partial f_i}{\partial x_1}(z) \\ \vdots \\ \frac{\partial f_i}{\partial x_n}(z) \end{bmatrix}, \qquad J(z) = \nabla f(z) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(z) & \dots & \frac{\partial f_1}{\partial x_n}(z) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(z) & \dots & \frac{\partial f_m}{\partial x_n}(z) \end{bmatrix}$$

$$\boxed{f(x) = f(z) + J(z)(x - z)}$$

Producto de matrices

Definición. Sea $A \in \mathbb{R}^{m\times p}$ y $B \in \mathbb{R}^{p\times n}$, entonces el producto matricial de A y B es una matriz $C \in \mathbb{R}^{m\times n}$, tal que

$$c_{ij} = \sum_{k=1}^{p} a_{ik}b_{kj}, \quad 1 \leq i \leq m,\, 1 \leq j \leq n.$$

Ejemplo.

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

Propiedades del producto matricial

Propiedad del producto matricial	Representación matemática
Asociatividad	$A(BC) = (AB)C$
Distributividad con respecto a la suma de matrices por la izquierda	$A(B + C) = AB + AC$
Distributividad con respecto a la suma de matrices por la derecha	$(A + B)C = AC + BC$
Compatibilidad con la multiplicación por escalar	$A(rB) = rAB = (rA)B$
Transpuesta del producto	$(AB)^\top = B^\top A^\top$

Producto exterior

Definición. Sean los vectores $A \in \mathbb{R}^m$ y $b \in \mathbb{R}^n$. Entonces el producto $(ab^\top) \in \mathbb{R}^{m\times n}$ se conoce como el producto exterior de a y b definido por

$$ab^\top = \begin{bmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_n \\ a_2b_1 & a_2b_2 & \dots & a_2b_n \\ \vdots & \vdots & \ddots & \vdots \\ a_mb_1 & a_mb_2 & \dots & a_mb_n \end{bmatrix}$$

Producto matricial

Es el producto matricial conmutativo?

$$AB = BA?$$

Producto matricial. Representaciones

Sean las matrices $A \in \mathbb{R}^{m \times p}$ y $B \in \mathbb{R}^{p \times n}$ tiene múltiples representaciones

Si representamos B en términos de sus **columnas**

$$B = \begin{bmatrix} b_1 & \dots & b_n \end{bmatrix} \quad \text{con} \quad b_1, \dots, b_n \in \mathbb{R}^p \quad AB = \begin{bmatrix} Ab_1 & \dots & Ab_n \end{bmatrix}$$

Si representamos A en términos de sus **filas**

$$A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix} \quad \text{con} \quad a_1, \dots, a_m \in \mathbb{R}^p, \quad AB = \begin{bmatrix} (B^\top a_1)^\top \\ \vdots \\ (B^\top a_m)^\top \end{bmatrix}$$

Producto matricial. Representaciones

Si representamos $A \in \mathbb{R}^{m \times p}$ en términos de sus **filas** y $B \in \mathbb{R}^{p \times n}$ de sus **columnas**

$$A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} b_1 & \dots & b_n \end{bmatrix} \quad \text{entonces} \quad AB = \begin{bmatrix} a_1 \cdot b_1 & \dots & a_1 \cdot b_n \\ \vdots & \ddots & \vdots \\ a_m \cdot b_1 & \dots & a_m \cdot b_n \end{bmatrix}$$

Si representamos $A \in \mathbb{R}^{m \times p}$ en términos de sus **columnas** y $B \in \mathbb{R}^{p \times n}$de sus **filas**

$$A = \begin{bmatrix} a_1 & \dots & a_k \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} b_1^\top \\ \vdots \\ b_k^\top \end{bmatrix} \quad \text{entonces} \quad AB = a_1 b_1^\top + \dots + a_k b_k^\top$$

Ilustración de producto matricial

Composición de funciones lineales.

$$\left.\begin{array}{rrcl} f: & \mathbb{R}^p & \rightarrow & \mathbb{R}^m \\ & x & \rightarrow & f(x) = Ax \\ g: & \mathbb{R}^n & \rightarrow & \mathbb{R}^p \\ & x & \rightarrow & g(x) = Bx \end{array}\right\} \qquad \begin{array}{rrcl} g: & \mathbb{R}^n & \rightarrow & \mathbb{R}^m \\ & x & \rightarrow & h(x) = f(g(x)) = ABx \end{array}$$

Ilustración de producto matricial

Matriz de diferencias de orden 2. Consideramos la matriz $D_n \in \mathbb{R}^{n-1\times n}$, y $D_{n-1} \in \mathbb{R}^{n-2\times n-1}$. Considere la ecuación diferencial:

$$\begin{cases} -u''(x) &= \sin(2\pi x), \quad x \in (0,1), \\ u(0) &= 0, \\ u(1) &= 0. \end{cases}$$

Calcule la solución del sistema lineal

$$\frac{1}{(n-2)^2}(D_{n-1}D_n)u(x) = b$$

Matriz Inversa

Definición de matriz Inversa por la izquierda y derecha

Sea $A \in \mathbb{R}^{m\times n}$. Entonces

- Una matriz $X \in \mathbb{R}^{n\times m}$ que satisface $XA = I \in \mathbb{R}^{n\times n}$, es llamada una **inversa por la izquierda**. Decimos que la matriz A es invertible por la izquierda si una inversa por la izquierda de A existe.
- Una matriz $X \in \mathbb{R}^{m\times n}$ que satisface $AX = I \in \mathbb{R}^{m\times m}$, es llamada una **inversa por la derecha**. Decimos que la matriz A es invertible por la derecha si una inversa por la derecha de A existe.

Propiedades de la matriz inversa

Si A tiene una inversa por la izquierda X, entonces las **columnas** de A son **linealmente independientes**. En efecto, si suponemos que las columnas de A no son linealmente independientes entonces existe un vector $x \in \mathbb{R}^n$, no nulo, tal que $Ax = 0$, así

$$0 = X(Ax) = (XA)x = Ix = x$$

lo que es una contradicción.
El recíproco también es cierto, es decir, si las columnas de una matriz son linealmente independientes, entonces existe una matriz inversa por la izquierda de A.

Propiedades de la matriz inversa

- Si A tiene una inversa por la derecha B, entonces $B^\top$ es una inversa por la izquierda de $A^\top$. En efecto, $\quad AB = I \quad \Longrightarrow \quad (B^\top A^\top) = (AB)^\top = I$.
- Si A tiene una inversa por la izquierda C, entonces $C^\top$ es una inversa por la derecha de $A^\top$. En efecto, $\quad A^\top C^\top = (CA)^\top = I$.
- Una matriz es invertible por la derecha si y sólo si sus filas son linealmente independiente.
- Una matriz alta no puede tener una inversa por la derecha. Solo matrices cuadradas o anchas pueden ser invertibles por la derecha.

Ejemplo: La inversa por la izquierda no es única.

Considere la siguiente matriz $A \in \mathbb{R}^{3\times 2}$,

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

Las siguientes dos matrices son inversas por la izquierda de A

$$B = \frac{1}{9}\begin{bmatrix} -11 & -10 & 16 \\ 7 & 8 & -11 \end{bmatrix}, \qquad C = \frac{1}{2}\begin{bmatrix} 0 & -1 & 6 \\ 0 & 1 & -4 \end{bmatrix}$$

Definición de la matriz inversa

Si una matriz A es invertible por la izquierda y por la derecha entonces las inversas por la izquierda y por la derecha son iguales, y estas son únicas. Además decimos en este caso que la matriz es **invertible** (o no singular) y la matriz inversa se denota por A^{-1}. Una matriz cuadrada que no es invertible se dice **singular**.

$$AA^{-1} = A^{-1}A = I$$

Condiciones de invertibilidad

Si $A \in \mathbb{R}^{n\times n}$ entonces invertibilidad por la izquierda, invertibilidad por la derecha e invertibilidad son equivalentes.
En efecto, suponga que $A \in \mathbb{R}^{n\times n}$ es una matriz invertible por la izquierda, entonces existen b_i tales que

$$Ab_i = e_i \quad \Longleftarrow \quad AB = A[b_1, b_2, ..., b_n] = [e_1, e_2, ..., e_n] = I,$$

lo que implica que B es la inversa por la derecha de A, por lo tanto

$$\text{invert. por la izquierda} \quad \Longrightarrow \quad \text{indep. de columnas} \quad \Longrightarrow \quad \text{invert.por la derecha}$$

Equivalentemente podemos mostrar que

$$\text{invert. por la derecha} \quad \Longrightarrow \quad \text{indep. de filas} \quad \Longrightarrow \quad \text{invert.por la izquierda}$$

Ejemplos básicos de cálculo de inversa

1. La inversa de la matriz identidad I es la misma matriz identidad, es decir, $I^{-1} = I$.
2. La inversa de una matriz diagonal A con entradas diagonales distintas de cero es la matriz diagonal con entradas diagonales el inverso de las entradas de A, es decir

$$A^{-1} = \begin{bmatrix} 1/a_{11} & 0 & \cdots & 0 \\ 0 & 1/a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1/a_{nn} \end{bmatrix}$$

o tambien lo escribimos como

$$A = \mathbf{diag}(a_{11}, a_{22}, ..., a_{nn})^{-1} = \mathbf{diag}(a_{11}^{-1}, a_{22}^{-1}, ..., a_{nn}^{-1}).$$

Ejemplos básicos de cálculo de inversa

1. Considere la matriz A y su inversa A^{-1} dadas por

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 0 & 2 & 2 \\ -3 & -4 & -4 \end{bmatrix}, \qquad A^{-1} = \frac{1}{30}\begin{bmatrix} 0 & -20 & -10 \\ -6 & 5 & -2 \\ 6 & 10 & 2 \end{bmatrix}$$

   estas verifican que $AA^{-1} = I$.

2. Formula de la inversa para matrices de 2×2. Una matriz $A \in \mathbb{R}^{2\times 2}$ es invertible si y sólo si $a_{11}a_{22} - a_{12}a_{21} =\neq 0$, y su inversa está dada por:

$$A^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

Ejemplo de inversa

Decida si las siguientes matrices son o no invertibles.

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

Resolución de sistema lineal por la inversa

Consideremos un sistema de ecuaciones lineales con n variables $Ax = b$, y asumamos que A es invertible, entonces para cualquier $b \in \mathbb{R}^n$, la solución es

$$x = A^{-1}b$$

Un sistema cuadrado de ecuaciones lineales $Ax = b$, con A una matriz invertible, tiene una única solución $x = A^{-1}b$, para cualquier vector b.

Otras propiedades

- Si A es invertible, su matriz transpuesta $A^\top$ es tambien invertible y su inversa es $(A^{-1})^\top = A^{-\top}$.
- Si A y B son invertibles y del mismo tamaño, entonces el producto AB es invertible y $(AB)^{-1} = B^{-1}A^{-1}$.

Clasificación lineal

Motivación

Considere una librería que tiene una lista de n clientes $\{\boldsymbol{c}\}$ donde para cada cliente $\boldsymbol{c}$ tenemos registrados n-atributos $\boldsymbol{c} = (c_1, c_2, ..., c_n)$. Entonces, con esta información, se quiere determinar a que clientes recomendar un determinado libro L.

Hipótesis de linealidad. Asumimos que la decisión de recomendar o no recomendar un determinado libro L al cliente $\boldsymbol{c}$ dependende de forma lineal de los atributos del cliente $\boldsymbol{c}$.

Este supuesto lo podemos interpretar matemáticamente como

$$w_1 c_1 + w_2 c_2 + ... + w_n c_n \geq w_0$$

si y sólo si se recomienda el libro L a $\boldsymbol{c}$.

Problema.

$$w^\top \boldsymbol{c} \geq w_0$$

Como determinar w?

Datos

Buscamos en la base de datos ejemplos de clientes que compraron o no el libro

Ejemplos positivos $\{\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_m\}$.

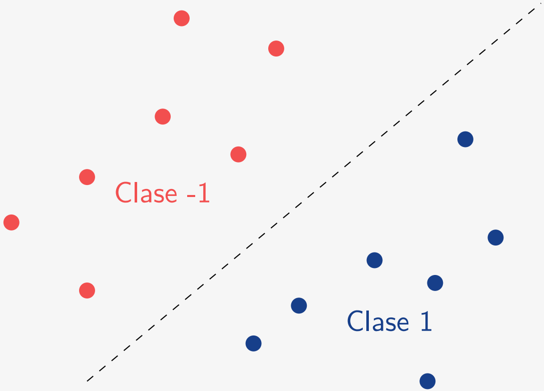
$$\begin{array}{rcl} w_1 p_{11} + w_2 p_{12} + ... + w_n p_{1n} & \geq & w_0 \\ w_1 p_{21} + w_2 p_{22} + ... + w_n p_{2n} & \geq & w_0 \\ \vdots & \geq & \vdots \\ w_1 p_{m1} + w_2 p_{m2} + ... + w_n p_{mn} & \geq & w_0 \end{array}$$

Ejemplos negativos $\{\boldsymbol{q}_1, \boldsymbol{q}_2, ..., \boldsymbol{q}_r\}$.

$$\begin{array}{rcl} w_1 q_{11} + w_2 q_{12} + ... + w_n q_{1n} & \geq & w_0 \\ w_1 q_{21} + w_2 q_{22} + ... + w_n q_{2n} & \geq & w_0 \\ \vdots & \geq & \vdots \\ w_1 q_{r1} + w_2 q_{r2} + ... + w_n q_{rn} & \geq & w_0 \end{array}$$

El problema de **separar** un conjunto de datos usando hiperplanos se conoce como clasificación lineal.

Problema de clasificación lineal



Clasificación lineal

Observación.

- Un conjunto de datos no es necesariamente separable linealmente.
- Un conjunto de datos puede transformarse para que si pueda ser linealmente separable.
- Algoritmos para resolver el problema:

Ejemplos positivos

$$\begin{aligned} P \in \mathbb{R}^{m\times n}, \quad P\boldsymbol{w} &\geq w_0 \\ \longrightarrow \quad \hat{P}\boldsymbol{y} &\geq 0 \end{aligned}$$

Ejemplos negativos

$$\begin{aligned} Q \in \mathbb{R}^{m\times n}, \quad Q\boldsymbol{w} &< w_0 \\ \longrightarrow \quad \hat{Q}\boldsymbol{y} &< 0 \end{aligned}$$

Clasificación lineal

Observación.

- Un conjunto de datos no es necesariamente separable linealmente.
- Un conjunto de datos puede transformarse para que si pueda ser linealmente separable.
- Algoritmos para resolver el problema: regresion lineal, perceptron, SVM.

Ejemplos positivos

$$P \in \mathbb{R}^{m\times n}, \quad P\boldsymbol{w} \geq w_0$$
$$\longrightarrow \quad \hat{P}\boldsymbol{y} \geq 0$$

Ejemplos negativos

$$Q \in \mathbb{R}^{m\times n}, \quad Q\boldsymbol{w} < w_0$$
$$\longrightarrow \quad \hat{Q}\boldsymbol{y} < 0$$

Algoritmo del perceptron

Algorithm Perceptron

Input: e.
Output: $y \in \mathbb{R}^n$

$y_0 = 0 \in \mathbb{R}^n$
while y no sea solución **do**
Escoja e mal clasificado por y
Actualice $y = \eta \begin{cases} (y+e) & \text{si } e \text{ es positivo} \\ (y-e) & \text{si } e \text{ es negativo} \end{cases}$
end while

Mas aplicaciones de clasificadores Booleanos

Control de Acceso. Utilizados para tomar decisiones de acceso en sistemas de seguridad. Ejemplo: Permitir el acceso solo si el titular de la tarjeta (A) y el código PIN (B) son correctos: $Y = A \land B$.

Detección de Fraudes. Identificación de transacciones sospechosas basadas en reglas lógicas. Ejemplo: Alertar si tanto la ubicación de la transacción (A) como la cantidad (B) son inusuales: $Y = A \land B$.

Automatización de Procesos. Controlar el flujo de trabajo en sistemas automatizados.Ejemplo: Iniciar una tarea solo si dos condiciones son cumplidas: $Y = A \land B$.

Diagnóstico Médico Simple. Clasificación de pacientes en función de síntomas simples. Ejemplo: Diagnosticar una enfermedad si tanto la presión arterial (A) como la temperatura (B) están fuera de los límites normales: $Y = A \lor B$.

Ver Capítulo 14 Boyd & Vanderbergue

Sistemas lineales

Introducción y motivación

En general queremos resolver un sistema de m ecuaciones lineales en n variables o incógnitas $x_1, x_2, ..., x_n$

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}\right\} \iff \quad Ax = b$$

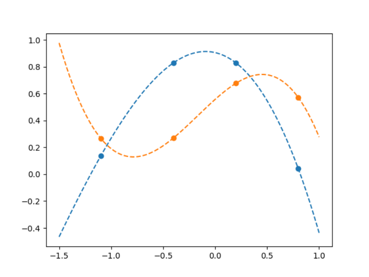
Definición. El sistema de ecuaciones lineales se dice:

- **Sobredeterminado**, si $m > n$,
- **Subdeterminado**, si $m < n$,
- **Cuadradp**, si $m = n$.

Ejemplo: Interpolación polinomial

Considere el problema de encontrar los coeficientes del polinomio cúbico que $p(x) = c_1 + c_2 x + c_3 x^2 + c_4 x^3$ que interpola los valores b_1, b_2, b_3, b_4 en los puntos $x = -1.1, -0.4, 0.2, 0.8$. Resolvemos generando la matriz de Vandermonde

$$A = \begin{bmatrix} 1 & -1.1 & (-1.1)^2 & (-1.1)^3 \\ 1 & -0.4 & (-0.4)^2 & (-0.4)^3 \\ 1 & 0.2 & (0.2)^2 & (0.2)^3 \\ 1 & 0.8 & (0.8)^2 & (0.8)^3 \end{bmatrix}$$

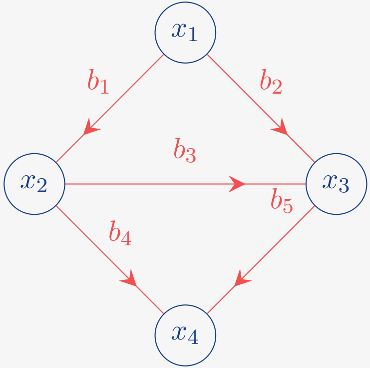


Encontrar los valores de los coeficientes es equivalente a resolver el sistema lineal

$$Ac = b$$

Ejemplo: grafos direccionados

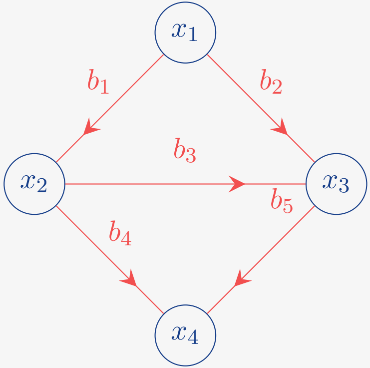
Matriz de incidencia de un grafo. $a_{ij} = \begin{cases} 1 & \text{arista } i \text{ apunta hacia nodo } j, \\ -1 & \text{arista } i \text{ apunta desde nodo } j, \\ 0 & \text{otherwise.} \end{cases}$



$$\underbrace{\begin{array}{rrrrcl} -x_1 & +x_2 & & & = & b_1 \\ -x_1 & & +x_3 & & = & b_2 \\ & -x_2 & +x_3 & & = & b_3 \\ & -x_2 & & +x_4 & = & b_4 \\ & & -x_3 & +x_4 & = & b_5 \end{array}}_{Ax=b}$$

Ejemplo: grafos direccionados

Matriz de incidencia de un grafo. $a_{ij} = \begin{cases} 1 & \text{arista } i \text{ apunta hacia nodo } j, \\ -1 & \text{arista } i \text{ apunta desde nodo } j, \\ 0 & \text{otherwise.} \end{cases}$



$$\begin{bmatrix} -1 & +1 & 0 & 0 \\ -1 & 0 & +1 & 0 \\ 0 & -1 & +1 & 0 \\ 0 & -1 & 0 & +1 \\ 0 & 0 & -1 & +1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$

Sistemas triangulares

Sistema triangular

Consideremos el sistema triangular inferior para $L \in \mathbb{R}^{n\times n}$, triangular inferior

$$
\begin{array}{ccccccc}
L_{11}x_1 & 0 & \cdots & 0 & = & b_1 \\
L_{21}x_1 & +L_{22}x_2 & \ddots & \vdots & = & b_2 \\
\vdots & \ddots & \ddots & 0 & = & \vdots \\
L_{n1}x_1 + & \cdots & +L_{nn-1}x_{n-1} & +L_{nn}x_n & = & b_n
\end{array}
$$

El sistema es invertible si las columnas de L son linealmente independientes, es decir que $Lx = 0$ es solo posible si $x = 0$..

Algoritmo de sustitución progresiva

Consideramos un algoritmo para resolver un conjunto de ecuaciones lineales $Lx = b$, donde $n \times n$ matriz L es **triangular inferior** con entradas en la diagonal no cero, así es invertible.

$$\begin{bmatrix} L_{11} & 0 & \cdots & 0 \\ L_{21} & L_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ L_{n1} & \cdots & L_{nn-1} & L_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Algoritmo para sistema triangular inferior

Algorithm Sustitución progresiva

Input: Matriz $L \in \mathbb{R}^{n\times n}$, triangular inferior invertible, y $b \in \mathbb{R}^n$.
Output: Vector $x \in \mathbb{R}^n$, solución de $Lx = b$.

for i=1:n **do**

$$x_i = (b_i - L_{i,i-1}x_{i-1} - \cdots - L_{i,1}x_1)/L_{ii}$$

end for

Algoritmo de sustitución regresiva

Algoritmo para resolver un conjunto de ecuaciones lineales $Rx = b$, donde $n \times n$ matriz R es **triangular superior** con entradas en la diagonal no cero, así es invertible.

$$\begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ 0 & R_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & R_{n-1n} \\ 0 & \cdots & 0 & R_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}.$$

Algoritmo para sistema triangular superior

Algorithm Sustitución regresiva

Input: Matriz $R \in \mathbb{R}^{n \times n}$, triangular superior invertible, $b \in \mathbb{R}^n$.
Output: $x \in \mathbb{R}^n$, solución de $Rx = b$.

for i=n:1 **do**
  $x_i = (b_i - R_{i,i+1}x_{i+1} - \cdots - R_{i,n}x_n)/R_{ii}$
end for

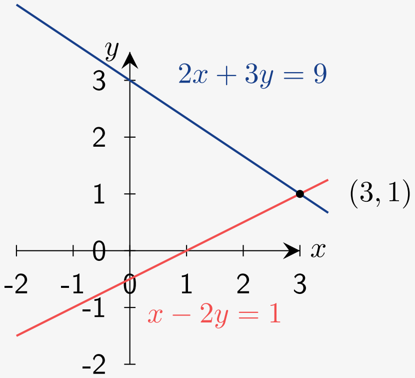
Eliminación Gaussiana

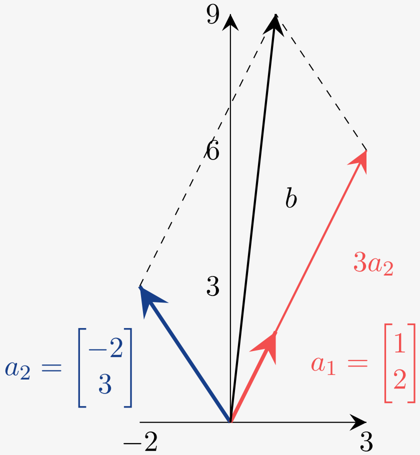
Ejemplo: interpretación por vector filas vs. columna

Considere el siguiente sistema lineal de 2×2.

$$\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix} \quad \Longleftrightarrow \quad \begin{array}{rcrcr} x_1 & - & 2x_2 & = & 1 \\ 2x_1 & + & 3x_2 & = & 9 \end{array} \quad \Longleftrightarrow \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

Ejemplo: interpretación por vector filas vs. columna





Eliminación Gaussiana

La elimininación estándar tiene el siguiente orden

Columna 1:	Escoger como pivot el elemento de esta columna que corresponde a la primera ecuación. Usar la ecuación 1 para crear ceros bajo el primer pivot. (Los pivots no pueden ser cero.)
Columna 2:	Use como pivot el elemento de esta columna y la segunda ecuación. Usar la segunda ecuación para crear ceros bajo el segundo pivot.
Columna 3 a n:	Continuar con el procedimiento hasta encontrar la matriz triangular superior U.

Eliminación Gaussiana

$$\begin{bmatrix} \boldsymbol{x} & x & x & x \\ \boldsymbol{x} & x & x & x \\ \boldsymbol{x} & x & x & x \\ \boldsymbol{x} & x & x & x \end{bmatrix} \longrightarrow \begin{bmatrix} x & x & x & x \\ 0 & \boldsymbol{x} & x & x \\ 0 & \boldsymbol{x} & x & x \\ 0 & \boldsymbol{x} & x & x \end{bmatrix} \longrightarrow \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & \boldsymbol{x} & x \\ 0 & 0 & \boldsymbol{x} & x \end{bmatrix} \longrightarrow \begin{bmatrix} x & x & x & x \\ 0 & x & x & x \\ 0 & 0 & x & x \\ 0 & 0 & 0 & x \end{bmatrix}$$

Paso 1 Paso 2 Paso 3 Paso 4

Eliminación Gaussiana

En el Paso 1, usamos el pivot A_{11} para hacer ceros bajo el pivot en la columna 1. Así debemos restar a las columnas 2,3,y 4 la columna 1 multiplicado por:

$$\text{Multiplicadores:} \quad \ell_{21} = \frac{A_{21}}{A_{11}}, \quad \ell_{31} = \frac{A_{31}}{A_{11}}, \quad \ell_{41} = \frac{A_{41}}{A_{11}},$$

En el Paso 2, usamos el pivot de la columna 2 y fila 2 de la matriz actualizada y hacemos ceros bajo el pivot calculando los multiplicadores ℓ_{32} y ℓ_{42}. Finalmente, en el tercer paso calculamos el multiplicador ℓ_{43}.

Ejemplo de eliminación Gaussiana

Consideramos el proceso de eliminación Gaussiana

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 2 & 7 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 3 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

Ejemplo de eliminación Gaussiana

Interpretación equivalente por columnas

$$\begin{bmatrix}1 & 2 & 3\\2 & 5 & 7\\2 & 7 & 8\end{bmatrix} - \underbrace{\begin{bmatrix}1 & 2 & 3\\2 & 4 & 6\\2 & 4 & 6\end{bmatrix}}_{\ell_1 u_1^\top} = \begin{bmatrix}0 & 0 & 0\\0 & 1 & 1\\0 & 3 & 2\end{bmatrix}$$

$$\begin{bmatrix}0 & 0 & 0\\0 & 1 & 1\\0 & 3 & 2\end{bmatrix} - \underbrace{\begin{bmatrix}0 & 0 & 0\\0 & 1 & 1\\0 & 3 & 3\end{bmatrix}}_{\ell_2 u_2^\top} = \underbrace{\begin{bmatrix}0 & 0 & 0\\0 & 0 & 0\\0 & 0 & -1\end{bmatrix}}_{\ell_3 u_3^\top}$$

Ejemplo de eliminación Gaussiana

$$
\begin{aligned}
A &= \underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 4 & 6 \end{bmatrix}}_{\ell_1 u_1^\top} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 3 & 2 \end{bmatrix}}_{\ell_2 u_2^\top} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{\ell_3 u_3^\top} \\
&= \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 3 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} = LU
\end{aligned}
$$

Pivoteo parcial/Intercambio de filas

No solo es necesario intercambiar filas cuando nos encontramos con un pivot que puede ser cero, sino que además es necesario hacerlo por razones de estabilidad.

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 7 \\ \mathbf{2} & 4 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & \mathbf{1} & 1 \\ {\color{blue}0} & 1 & 3 \\ 2 & 4 & 8 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 & -1 \\ 0 & {\color{blue}0} & 2 \\ 2 & 4 & 8 \end{bmatrix}$$

Pivoteo parcial/Intercambio de filas

$$A = \underbrace{\begin{bmatrix} 0 \\ 1/2 \\ 1 \end{bmatrix}}_{\ell_1} \underbrace{\begin{bmatrix} 2 & 4 & 8 \end{bmatrix}}_{u_1^\top} + \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\ell_2} \underbrace{\begin{bmatrix} 0 & 1 & 1 \end{bmatrix}}_{u_2^\top} + \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}_{\ell_3} \underbrace{\begin{bmatrix} 0 & 0 & 2 \end{bmatrix}}_{u_3^\top}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Pivoteo parcial/Intercambio de filas

Sea la matriz de permutación

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

Entonces

$$PA = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 & 8 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

Observación: La inversa de una matriz de permutación P es su transpuesta $P^{\top}$.

Resolución del sistema lineal

¿Como resolvemos el sistema lineal $Ax = b$ usando la factorización $PA = LU$?

$$Ax = b \iff PAx = Pb \iff LUx = Pb \iff \begin{cases} Ly = Pb \\ Ux = y \end{cases}$$

Algorithm Eliminación Gaussiana con Pivotes Parciales y Factorización $PA = LU$

Input: Matriz $A \in \mathbb{R}^{n \times n}$.
Output: Matrices L (triang. inferior), U (triang. superior) y P (permutación)

```
for k = 1 : n − 1 do
    p ← arg max_{i=k}^{n} |U_ik|
    if p ≠ k then
        Intercambiar fila k con fila p en U, en L, y en P
    end if
    for i = k + 1 : n do
        m ← U_ik / U_kk
        L_ik ← m
        for j = k : n do
            U_ij ← U_ij − m · U_kj
        end for
    end for
end for
```

Forma de echelon

Se dice que una matriz $A \in \mathbb{R}^{m \times n}$ tiene forma de echelon o forma de escalón si

1. Todas las filas que contienen solo ceros están en la parte inferior de la matriz.
2. El primer elemento no nulo de cada fila no nula está a la derecha del primer elemento no nulo de la fila anterior.

Ejemplo:

$$A = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & -1 & -4 & 8 \\ -1 & 1 & 3 & -5 \\ -1 & 2 & 5 & -6 \\ -1 & -2 & -3 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Las columnas 1, 2, y 4 de A forman la base del subespacio $\mathbf{Im}(A)$.

Independencia lineal

Pregunta. Dada una colección de vectores $a_1, ..., a_k \in \mathbb{R}^n$ y un vector $b \in \mathbb{R}^n$. Comopodemos determinar si b es combinación lineal de los vectores $\{a_1, ..., a_k\}$.

PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE