PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE

UC | Chile

UC | Chile

Algebra Lineal Aplicada para Ciencia de Datos

Clase 5. Eliminación Gaussiana y factorización LU

UC | Chile

Introducción

Introducción y motivación

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

$$\begin{aligned} A_{11}x_1 + A_{12}x_2 + \cdots + A_{1n}x_n &= b_1 \\ A_{21}x_1 + A_{22}x_2 + \cdots + A_{2n}x_n &= b_2 \\ &\vdots \\ A_{m1}x_1 + A_{m2}x_2 + \cdots + A_{mn}x_n &= b_m \end{aligned}$$

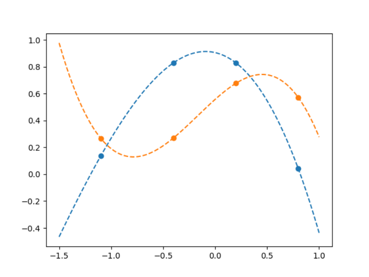
expresado en forma matricial por

$$Ax = b$$

Ejemplo: Interpolación polinomial

Considere el problema de encontrar los coeficientes del polinomio cubico 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

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

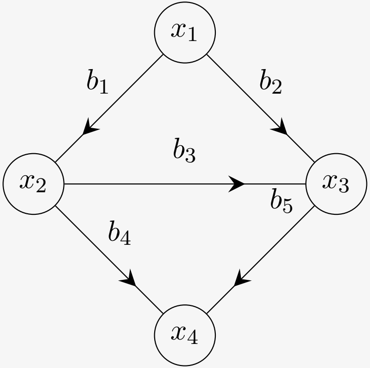


Encontrar los valores de los coeficientes es equivalente a resolver el sistema lineal

$$Ac = b$$

Ejemplo: grafos direccionados

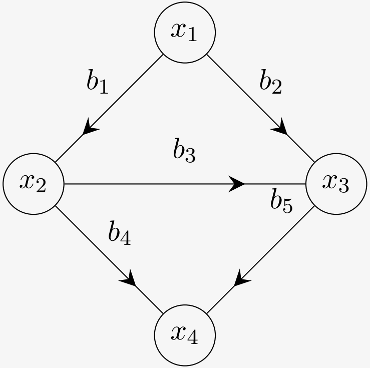
Matriz de incidencia de un grafo. $A_{ij} = \begin{cases} 1 & \text{arista } i \text{ apunta hacia nodo } j, \\ -1 & \text{arista } i \text{ apunta desde nodo } j, \\ 0 & \text{otherwise.} \end{cases}$



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

Ejemplo: grafos direccionados

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



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

Tema 1: 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

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

- 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 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:

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

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

Tema 2: Sistemas triangulares

Sistema triangular

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

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

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 de sustitución progresiva

Algorithm 1: Sustitución progresiva

Data: $L \in \mathbb{R}^{n\times n}$, triangular inferior invertible, $b \in \mathbb{R}^n$.
Result: $x \in \mathbb{R}^n$, solución de $Lx = b$.
for $i \leftarrow 1$ **to** n **do**
  $x_i = (b_i - L_{i,i-1}x_{i-1} - \cdots - L_{i,1}x_1)/L_{ii}$
return x

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 de sustitución regresiva

Algorithm 2: Sustitución regresiva

Data: $R \in \mathbb{R}^{n \times n}$, triangular superior invertible, $b \in \mathbb{R}^n$.
Result: $x \in \mathbb{R}^n$, solución de $Rx = b$.
for $i \leftarrow n$ **to** 1 **do**
  State $x_i = (b_i - R_{i,i+1}x_{i+1} - \cdots - R_{i,n}x_n)/R_{ii}$
return x

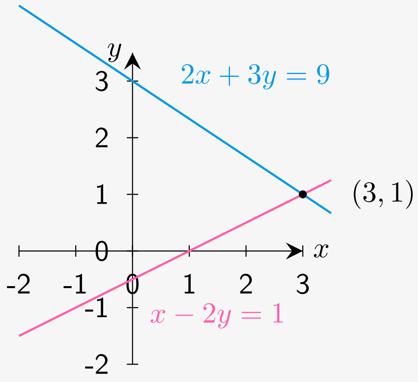
Tema 3: Eliminación Gaussiana

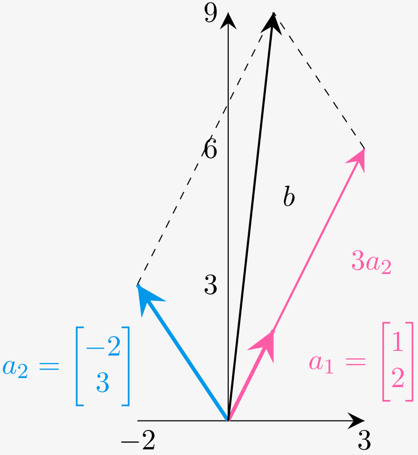
Ejemplo: interpretación por vector filas vs. columna

Considere el siguiente sistema lineal de 2×2.

$$\begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix} \iff \begin{array}{rcrcl} x_1 & - & 2x_2 & = & 1 \\ 2x_1 & + & 3x_2 & = & 9 \end{array} \iff \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$$

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.

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

Pivoteo parcial/Intercambio de filas

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

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 $PA = LU$?

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

Algoritmo de eliminación Gaussiana

Algorithm 3: Eliminación Gaussiana con Pivotes Parciales y Factorización $PA = LU$

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

```
Inicializar L y P como la matriz identidad, U como una copia de A;
for k ← 1 to n − 1 do
    p ← arg max_{i=k}^{n} |U_ik|;                  // Encuentra el pivote parcial
    if p ≠ k then
        Intercambiar fila k con fila p en U;        // Intercambia filas si es necesario
        Intercambiar fila k con fila p en L;
        Intercambiar fila k con fila p en P;
    for i ← k + 1 to n do
        m ← U_ik / U_kk;
        L_ik ← m;
        for j ← k to n do
            U_ij ← U_ij − m · U_kj;
return P, L, U;
```

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 $\mathbf{Im}(A)$.

UC | Chile