Universidad Nacional de Ingeniería
Facultad de Ciencias
Escuela Profesional de Matemática **Ciclo 2022-1**

[Cod: CM4F1 Curso: Análisis y Modelamiento Numérico I]

Segunda Práctica Calificada

---

1. (a) [1 *pto.*] Determine el número positivo normalizado más pequeño que se puede representar utilizando el estándar IEEE-754 (32 bits).

   (b) [1 *pto.*] Determine el número positivo no normalizado más pequeño que se puede representar utilizando el estándar IEEE-754 (32 bits).

   (c) [0.5 *pts.*] Represente en el estándar de coma flotante IEEE-754 de 32 bits los valores 10.25.

   (d) [0.5 *pts.*] Represente en el estándar de coma flotante IEEE-754 de 32 bits los valores 6.75.

   (e) [1 *pto.*] Determine la suma de *c*) y *d*) representados en IEEE-754, indicando los pasos que va realizando en cada momento.

   Solución:

   (a) [1 *pto.*] El número positivo normalizado más pequeño representable en el estándar IEEE-754 (32 bits) es:

$$\underbrace{0}_{\text{positivo}} \quad \underbrace{00000001\ 00000000000000000000000}_{\text{normalizado más pequeño}}$$

   Cuyo valor es:

$$1.0 \cdot 2^{1-127} = 2^{-126}$$

   (b) [1 *pto.*] El número positivo no normalizado más pequeño representable en el estándar es:

$$\underbrace{0}_{\text{positivo}} \quad \underbrace{00000000\ 00000000000000000000001}_{\text{no normalizado más pequeño}}$$

   Cuyo valor es:

$$2^{-23} \cdot 2^{-126} = 2^{-149}$$

   (c) [0.5 *pts.*] Haciendo:
   $10.25 = 1010.01_{(2)} = 1.01001 \times 2^3$

Signo $= 0$

Exponente $= 127 + 3 = 130 = 10000010_{(2)}$

Mantisa $= 010010000 \cdots 00$

En Binario: $010000010010010000 \cdots 00$

(d) [0.5 *pts.*] Haciendo:

$6.75 = 110.11_{(2)} = 1.1011 \times 2^2$

Signo $= 0$

Exponente $= 127 + 2 = 129 = 10000001_{(2)}$

Mantisa $= 10110000 \cdots 00$

En Binario: $010000001101100000 \cdots 00$

(e) [1 *pto.*] De (c) y (d) tenemos:

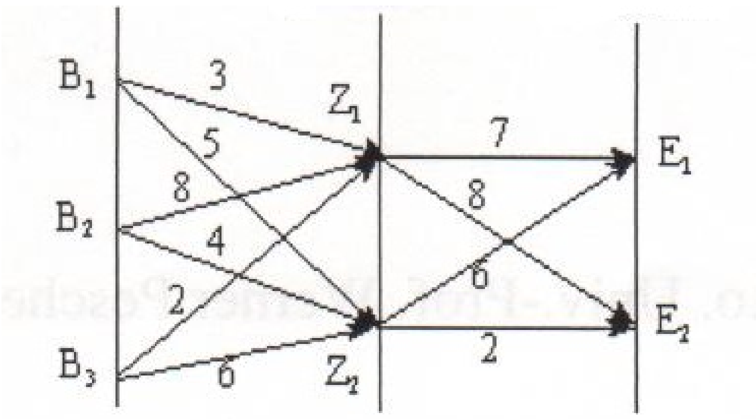
$10.25 = 0\ 10000010\ 1.01001000000000000000000$

$6.75 = 0\ 10000010\ 0.11011000000000000000000$

Suma $= 0\ 10000010\ 10.00100000000000000000000$

Resultado normalizado $= 0\ 10000011\ 00010000000000000000000$

2. Una fábrica utiliza tres productos primarios (B_i) para producir dos productos intermedios (Z_i), que a su vez estos producirán dos productos finales (E_i). Las cantidades requeridas para la producción son:



La demanda externa de unidades es cinco de B_1, diez de B_3, cinco de Z_1, dos de Z_2, tres de E_1 y cuatro de E_2. Determine la demanda total que se necesita por cada producto según los siguientes requerimientos.

(a) [1 *pto.*] Indique las variables.

(b) [1 *pto.*] Modele el sistema.

(c) [1 *pts.*] Determine A^{-1} usando Gauss-Jordan.

(d) [1 *pto.*] Determine el número de condición de A.

Solución:

(a) [1 *pto.*] Sean:

$$\begin{aligned}
x_1 &: \text{Cantidad total de demanda requerido de } B_1.\\
x_2 &: \text{Cantidad total de demanda requerido de } B_2.\\
x_3 &: \text{Cantidad total de demanda requerido de } B_3.\\
x_4 &: \text{Cantidad total de demanda requerido de } Z_1.\\
x_5 &: \text{Cantidad total de demanda requerido de } Z_2.\\
x_6 &: \text{Cantidad total de demanda requerido de } E_1.\\
x_7 &: \text{Cantidad total de demanda requerido de } E_2.
\end{aligned}$$

(b) [1 *pto.*] La demanda total es:

$$\begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{bmatrix} = \begin{bmatrix} 5\\ 0\\ 10\\ 5\\ 2\\ 3\\ 4 \end{bmatrix} + \begin{bmatrix} 0&0&0&3&5&0&0\\ 0&0&0&8&4&0&0\\ 0&0&0&2&6&0&0\\ 0&0&0&0&0&7&8\\ 0&0&0&0&0&6&2\\ 0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0 \end{bmatrix} \cdot \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{bmatrix}$$

El sistema resulta:

$$\begin{bmatrix} 1&0&0&-3&-5&0&0\\ 0&1&0&-8&-4&0&0\\ 0&0&1&-2&-6&0&0\\ 0&0&0&1&0&-7&-8\\ 0&0&0&0&1&-6&-2\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1 \end{bmatrix} \cdot \begin{bmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{bmatrix} = \begin{bmatrix} 5\\ 0\\ 10\\ 5\\ 2\\ 3\\ 4 \end{bmatrix}$$

(c) [1 *pto.*] Por las transformaciones de Gauss-Jordan, tenemos:

$$A^{-1} = \begin{bmatrix} 1&0&0&3&5&51&34\\ 0&1&0&8&4&80&72\\ 0&0&1&2&6&50&28\\ 0&0&0&1&0&7&8\\ 0&0&0&0&1&6&2\\ 0&0&0&0&0&1&0\\ 0&0&0&0&0&0&1 \end{bmatrix}$$

(d) [1 *pto.*] El número de condición es:

$$Cond_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 16 \times 165 = 2640.$$

3. Indique y justifique la veracidad (V) o falsedad (F) de cada una de las siguientes afirmaciones:

(a) Dados las sucesiones $x_n = o(1)$ y $z_n = o(1)$, se cumple que $x_n = O(z_n)$ y $z_n = O(x_n)$ (1.5 ptos)

(b) Es posible obtener una 2×2 matriz invertible con entradas de 4 digitos en su mantiza, tal que el cálculo de su determinante, usando 4 digitos en la mantisa y redondeo, sea 0. (1.5 ptos)

(c) Sea B la matriz que resulta de intercambiar las columnas de A entonces se cumple que $cond_2(B) = cond_2(A)$. (1 ptos)

Solución:

(a) (Falso) considerar $x_n = 1/n$ y $z_n = 1/n^2$, pero no existe $C \in I\!R$ tal que $1/n \leq C(1/n^2)$, pues de lo contrario C sería una cota superior de $\mathbb{N}$.

(b) (Verdadero) Considerar la matriz con entradas de 4 digitos en su mantiza $A = \begin{bmatrix} 0.1001 * 10^1 & 0.1902 * 10^1 \\ 0.1 * 10^1 & 0.19 * 10^1 \end{bmatrix}$ el cálculo exacto de la $det(A)$ es -10^{-4}, pero operando usando 4 digitos en la mantisa y redondeo, se tiene

$$[(0.1001 * 10^1) * (0.19 * 10^1)] - [(0.1902 * 10^1) * (0.1 * 10^1)] = fl(0.19019 * 10^1) - 0.1902 * 10^1 = 0$$

(c) (Verdadero) cada intercambio de dos columnas de A, es obtenida como AE_1, siendo E_1 una matriz ortogonal. Por tanto $B = AU$, siendo U ortogonal ($UU^t = I$). Luego $cond_2(B) = \sqrt{\rho_{max}(AU(AU)^t)}\sqrt{\rho_{max}([(AU)^{-1}]^t(AU)^{-1})} = cond_2(A)$

4. (a) Determinar cuantas cifras significativas se pierden al restar 456 y 440, en su normalizada forma flotante base 2. Luego compararlo con el resultado que se obtiene al aplicar el teorema sobre las cifras significativas. (2 ptos)

(b) Sea B la matriz que resulta al multilpicar una de las filas de A, por $\lambda > 0$. Mostrar que se cumple la desigualdad: $cond_2(B) \leq \max\{\lambda, \lambda^{-1}\}\, cond_2(A)$. Dar ejemplos donde se cumple la desigualdad estricta y la igualdad de la desigualdad anterior. (2 ptos)

Solución:

(a) Se cumple que $456 = (111001000)_2 = (0.111001)_2 * 2^9$ y $440 = (110111000)_2 = (0.110111)_2 * 2^9$, luego su diferencia es $(0.000010)_2 * 2^9$, observandose la perdida de 4 cifras significativas. Ahora acotaremos $1 - (440/456) = 16/456$ para aplicar el teorema: se tiene que $2^{-5} \leq 16/456 \leq 2^{-4}$, por tanto el teorema afirma que se perdera por lo menos 4 y lo sumo 5 digitos.

(b) Se tiene que $B = EA$, sinedo $E = diag\{1, 1, \cdots, \lambda, \cdots, 1, 1\}$, luego

$$cond_2(B) = ||EA||_2\, ||A^{-1}E^{-1}||_2 \leq ||A||_2\, ||A^{-1}||_2\, ||E||_2\, ||E^{-1}||_2 \leq cond_2(A) \max\{1, \lambda\} \max\{1, \lambda^{-1}\}$$

obteniendo la desigual al ser $\max\{1, \lambda\} \max\{1, \lambda^{-1}\} = \max\{\lambda, \lambda^{-1}\}$. Por otro lado considerando $A = I_n$, se tiene que $cond_2(B) = \max\{1, \lambda\}$ cumpliendo la igualdad y desigualda estrictita, en la desigualdad mostrada, para $\lambda \geq 1$ y $\lambda < 1$.

5. [4 *pts.*] Realizó su presentación en la práctica dirigida.

18 de Mayo del 2022