Notas de Métodos Numéricos

Carlos Martínez-Rodríguez
Academia de Matemáticas
Casa Libertad
carlos.martinez@uacm.edu.mx

Agosto 2025

El curso se llevará a cabo tres veces a la semana, con sesiones de 1.5 horas, de las cuales una de ellas se realizará en el laboratorio de cómputo 3.
El curso por su naturaleza implica que la/el estudiante implemente los distintos algoritmos que se revisan en el curso, por lo tanto es importante que demuestre que efectivamente puede implementar, revisar y mejorar los algoritmos ya existentes.

Hay dos maneras de certificar la materia: a) Portafolio y b) Examen de certificación. Al menos una semana antes de que termine el curso las y los estudiantes tendrán conocimiento de sus calificaciones parciales y por tanto de la calificación promedio obtenida al momento, para que sea el/la mismo(a) estudiante quién decida si certifica por la modalidad de portafolio, o por la modalidad de examen de certificación.

UACM Universidad Autónoma de la Ciudad de México Nada humano me es ajeno

El portafolio se conforma de evaluaciones (40 %), programas, (40 %) tareas (10 %), tareitas (5 %)y trabajos adicionales (5 %). Mientras que la certificación es un examen elaborado por el comité de certificación y que será presentado por todas y todos los estudiantes que se inscriban en esta modalidad en las fechas establecidas por la Coordinación de Certificación y Registro. En cualquiera de las dos modalidades es indispensable que el/la estudiante se registre a este proceso para que su calificación pueda ser asignada al final del proceso de Certificación.

UACM
Universidad Autónoma de la Ciudad de México
Nada humano me es ajeno

El curso tiene un fuerte sustento en la programación constante, sin embargo, es importante resaltar que los conceptos teóricos deben ser dominados por las y los estudiantes, por lo tanto las evaluaciones y las tareas tendrán estas dos componentes principales. Al contrario de lo que pueda pensarse la asistencia al curso es obligatoria pero no influye directamente en la calificación obtenida. Sin embargo, hay que mencionar que si la asistencia se realiza de manera intermitente es probable que cueste un poco de trabajo reincorporarse a la dinámica de trabajo que se irá construyendo con el grupo con el transcurso de las clases.

En este curso estudiaremos los elementos básicos de los métodos numéricos, utilizando el programa de distribución libre *R*. Antes de iniciar propiamente con el estudio de los métodos numéricos realizaremos un breve repaso de algunos conceptos de álgebra lineal, cálculo diferencial mismos que son fundamentales en esta materia y de la que se supone las y los estudiantes se encuentran familiarizados con ellos.

Comenzamos el capítulo con un repaso de algunos aspectos importantes del cálculo que son necesarios a lo largo del texto. Suponemos que los estudiantes que lean este texto conocen la terminología, la notación y los resultados que se dan en un curso típico de cálculo.

El límite de una función nos dice qué tan cerca se encuentran las imágenes de una función si dos elementos de su dominio se encuentran lo suficientemente cerca, es decir, decimos que una función $f(x)$ definida en el intervalo (a,b) tiene **límite** L en el punto $x = x_0$, lo que denotamos por

$$\lim_{x \to x_0} f(x) = L,$$

si para cualquier $\varepsilon > 0$, existe un número real $\delta > 0$ tal que $|f(x) - L| < \varepsilon$ siempre que $0 < |x - x_0| < \delta$. Es decir, que los valores de la función estarán cerca de L siempre que x esté suficientemente cerca de x_0.

Se dice que una función f es continua en a si cuando x se aproxima al valor de a, entonces también $f(x)$ se aproxima a $f(a)$, es decir, $f(x)$ es **continua** en el punto $x = a$ si

$$\lim_{x \to a} f(x) = f(a),$$

y se dice que f es **continua en el conjunto** (a, b) si es continua en cada uno de los puntos del intervalo. Denotaremos el conjunto de todas las funciones f que son continuas en $E = (a, b)$ por $C(E)$.

Se dice que una sucesión $\{x_n\}_{n=1}^{\infty}$ **converge** a un número x, si $\lim_{n\to\infty} x_n = x$ o bien, $x_n \to x$ cuando $n \to \infty$, si para cualquier $\varepsilon > 0$, existe un número natural $N(\varepsilon)$ tal que $|x_n - x| < \varepsilon$ para cada $n > N(\varepsilon)$. Cuando una sucesión tiene límite, se dice que la **sucesión converge**.

Si $f(x)$ es una función definida en un conjunto S de números reales y $x_0 \in S$, entonces las siguientes afirmaciones son equivalentes:

1. $f(x)$ es continua en $x = x_0$,
2. Si $\{x_n\}_{n=1}^{\infty}$ es cualquier sucesión en S que converge a x_0, entonces $\lim_{n\to\infty} f(x_n) = f(x_0)$.

UACM Universidad Autónoma de la Ciudad de México Nada humano me es ajeno

Teorema (Teorema del valor intermedio o de Bolzano.)

Si $f \in C[a,b]$ y ℓ es un número cualquiera entre $f(a)$ y $f(b)$, entonces existe al menos un número $c \in (a,b)$ tal que $f(c) = \ell$. Véase la Figura 1.

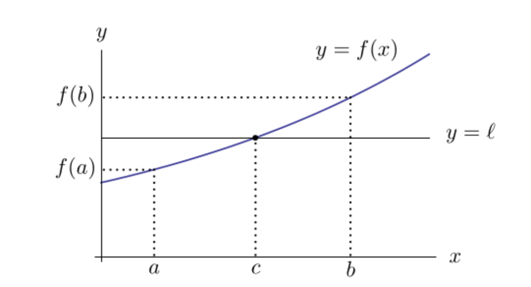


Figura: *Teorema del valor intermedio o de Bolzano*

Nota

Todas las funciones con las que se van a trabajar en este curso de métodos numéricos serán continuas, ya que esto es lo mínimo que debemos exigir para asegurar que la conducta de un método se puede predecir.

Si $f(x)$ es una función definida en un intervalo abierto que contiene un punto x_0, entonces se dice que $f(x)$ es **derivable** en $x = x_0$ cuando existe el límite

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

El número $f'(x_0)$ se llama **derivada** de f en x_0 y coincide con la pendiente de la recta tangente a la gráfica de f en el punto $(x_0, f(x_0))$, Figura 2.

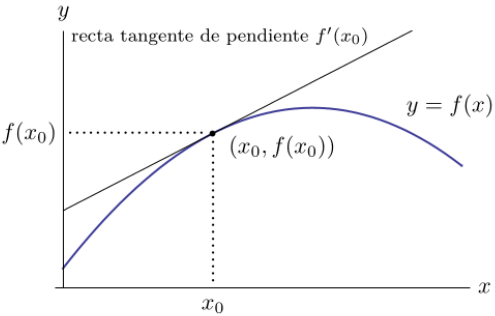


Figura: Derivada de una función en un punto

Derivabilidad implica continuidad. Si la función $f(x)$ es derivable en $x = x_0$, entonces $f(x)$ es continua en $x = x_0$. El conjunto de todas las funciones que admiten n derivadas continuas en S se denota por $C^n(S)$, mientras que el conjunto de todas las funciones indefinidamente derivables en S se denota por $C^\infty(S)$. Las funciones polinómicas, racionales, trigonométricas, exponenciales y logarítmicas están en $C^\infty(S)$, siendo S el conjunto de puntos en los que están definidas.

Teorema (Teorema del valor medio o de Lagrange.)

Si $f \in C[a, b]$ y es derivable en (a, b), entonces existe un punto $c \in (a, b)$ tal que

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Geométricamente hablando, Figura 3, el teorema del valor medio dice que hay al menos un número $c \in (a, b)$ tal que la pendiente de la recta tangente a la curva $y = f(x)$ en el punto $(c, f(c))$ es igual a la pendiente de la recta secante que pasa por los puntos $(a, f(a))$ y $(b, f(b))$.

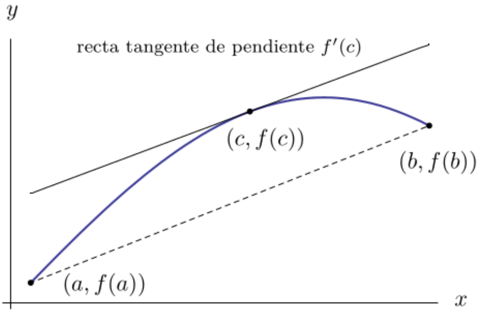


Figura: Teorema del valor medio o de Lagrange

Teorema (Teorema de los valores extremos.)

Si $f \in C[a,b]$, entonces existen c_1 y c_2 en (a,b) tales que $f(c_1) \leq f(x) \leq f(c_2)$ para todo $x \in [a,b]$. Si además, f es derivable en (a,b), entonces los puntos c_1 y c_2 están en los extremos de $[a,b]$ o bien son puntos críticos.

Teorema (Primer teorema fundamental o regla de Barrow.)

Si $f \in C[a,b]$ y F es una primitiva cualquiera de f en $[a,b]$ (es decir, $F'(x) = f(x)$), entonces

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Teorema (Segundo teorema fundamental.)

Si $f \in C[a,b]$ y $x \in (a,b)$, entonces

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

Teorema (Teorema del valor medio para integrales.)

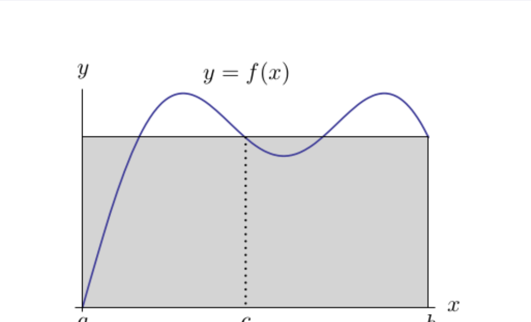
Si $f \in C[a,b]$, g es integrable en $[a,b]$ y $g(x)$ no cambia de signo en $[a,b]$, entonces existe un punto $c \in (a,b)$ tal que

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx.$$

Nota

Cuando $g(x) = 1$, véase la Figura 4, este resultado es el habitual teorema del valor medio para integrales y proporciona el valor medio de la función f en el intervalo $[a, b]$, que está dado por

$$f(c) = \frac{1}{b-a}\int_a^b f(x)\,dx.$$



Teorema (Teorema de Taylor.)

Supongamos que $f \in C^{(n)}[a,b]$ *y que* $f^{(n+1)}$ *existe en* $[a,b]$. *Sea* x_0 *un punto en* $[a,b]$. *Entonces, para cada* x *en* $[a,b]$, *existe un punto* $\xi(x)$ *entre* x_0 *y* x *tal que*

$$f(x) = P_n(x) + R_n(x),$$

donde

$$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2!}h^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}h^n = \sum_{k=0}^{n} \frac{f^{(k)}(x_0}{k!} \\ R_n(x) &= \frac{f^{(n+1)}(\xi(x))}{(n+1)!}h^{n+1}, \quad y \quad h = x - x_0. \end{aligned}$$

UACM Universidad Autónoma de la Ciudad de México Nada humano me es ajeno

El polinomio $P_n(x)$ se llama **n-ésimo polinomio de Taylor** de f alrededor de x_0 (véase la Figura 5). $R_n(x)$ se llama **error de truncamiento** (o *resto de Taylor*) asociado a $P_n(x)$. Como el punto $\xi(x)$ en el error de truncamiento $R_n(x)$ depende del punto x en el que se evalúa el polinomio $P_n(x)$, podemos verlo como una función de la variable x.

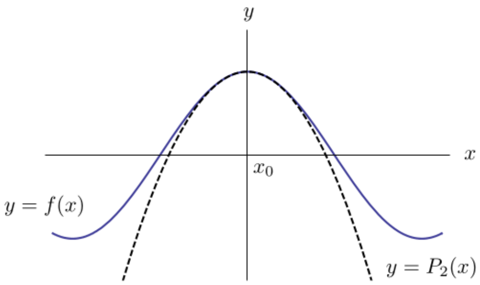


Figura: Gráficas de $y = f(x)$ y de su polinomio de Taylor $y = P_2(x)$ alrededor de x_0.

Nota

La serie infinita que resulta al tomar límite en la expresión de $P_n(x)$ cuando $n \to \infty$ se llama ***serie de Taylor*** *de f alrededor de x_0.*
Cuando $x_0 = 0$, el polinomio de Taylor se suele denominar ***polinomio de Maclaurin****, y la serie de Taylor se llama* ***serie de Maclaurin****.*

La denominación error de truncamiento en el teorema de Taylor se refiere al error que se comete al usar una suma truncada al aproximar la suma de una serie infinita.

A continuación enunciaremos algunas de las definiciones y teoremas básicos que utilizaremos a lo largo de estas notas.

Definición

f es de clase C^1 en el intervalo $[a; b]$ si f' es continua en $[a; b]$.

Definición

f es de clase C^n en el intervalo $[a; b]$ si $f^{(n)}$ es continua en $[a; b]$.

Definición

f es de clase C^∞ en el intervalo I si f es infinitas veces derivable y continua en I.

Teorema

(Teorema de los valores intermedios). Sea f continua en el intervalo $[a; b]$. Si $k \in \mathbb{R}$ es un número comprendido entre $f(a)$ y $f(b)$, entonces existe al menos un punto ξ perteneciente al intervalo $(a; b)$ tal que $f(\xi) = k$.

Teorema (Bolzano)

Si f es continua en el intervalo $[a; b]$ y $f(a) \cdot f(b) < 0$, entonces existe un ξ perteneciente al $(a; b)$ tal que $f(\xi) = 0$.

Teorema (Teorema de acotabilidad)

Si $f : [a; b] \mapsto \mathbb{R}$ es continua en $[a; b]$, entonces f está acotada en $[a; b]$.

Teorema (Teorema de Weierstrass)

Si $f : [a; b] \mapsto \mathbb{R}$ *es continua en* $[a; b]$, *entonces* f *tiene un máximo global y un mínimo global en* $[a; b]$.

Teorema (Teorema Generalizado de Rolle)

Si f *continua en* $[a; b]$, *y existen las derivadas* $f'(x), f''(x), \ldots, f^{(n)}(x)$ *en* $(a; b)$ *y* $f(x_0) = f(x_1) = \cdots = f(x_n) = 0$ *(con* $x_0, x_1, \ldots, x_n \in [a; b]$*)* *entonces existe* ξ *perteneciente al* $(a; b)$ *tal que* $f^{(n)}(\xi) = 0$.

Teorema (Teorema de Lagrange)

Si f continua en $[a; b]$ y derivable en $(a; b)$ entonces existe ξ perteneciente al $(a; b)$ tal que $f(b) - f(a) = f'(\xi)(b - a)$.

Teorema (Teorema del valor medio ponderado)

Sea f continua en $[a, b]$ y g una función integrable Riemann en $[a, b]$. Si g no cambia de signo en $[a, b]$, entonces existe un número $c \in (a, b)$ tal que:

$$\int_a^b f(x)g(x)dx = f(c)\int_a^b g(x)dx$$

Una **matriz** es un arreglo multidimensional de escalares, llamados *elementos*, ordenados en filas y columnas. Una matriz de m filas y n columnas, o *matriz (de orden)* $m \times n$, es un conjunto de $m \cdot n$ elementos a_{ij}, con $i = 1, 2, \ldots, m$ y $j = 1, 2, \ldots, n$, que se representa de la siguiente forma:

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

Se puede abreviar la representación de la matriz anterior de la forma $A = (a_{ij})$ con $i = 1, 2, \ldots, m$; $j = 1, 2, \ldots, n$.

Hay una relación directa entre matrices y vectores puesto que podemos pensar una matriz como una composición de vectores fila o de vectores columna. Además, un vector es un caso especial de matriz: un *vector fila* es una matriz con una sola fila y varias columnas, y un *vector columna* es una matriz con varias filas y una sola columna.

- Si $A = (a_{ij})$ y $B = (b_{ij})$ son dos matrices que tienen el mismo orden, $m \times n$, decimos que A y B son **iguales** si $a_{ij} = b_{ij}$ para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$.
- Si $A = (a_{ij})$ y $B = (b_{ij})$ son dos matrices que tienen el mismo orden $m \times n$, la **suma** de A y B es una matriz $C = (c_{ij})$ del mismo orden con $c_{ij} = a_{ij} + b_{ij}$ para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$.
- Si $A = (a_{ij})$ es una matriz de orden $m \times n$, la **multiplicación de** A **por un escalar** λ es una matriz $C = (c_{ij})$ del mismo orden $m \times n$ con $c_{ij} = \lambda a_{ij}$ para todo $i = 1, \ldots, m$ y $j = 1, \ldots, n$.
- Si $A = (a_{ij})$ es una matriz de orden $m \times n$, la **matriz traspuesta** de A es la matriz que resulta de intercambiar sus filas por sus columnas, se denota por A^T y es de orden $n \times m$

- Si $A = (a_{ij})$ es una matriz de orden $m \times p$ y $B = (b_{ij})$ es una matriz de orden $p \times n$, el **producto de A por B** es una matriz $C = (c_{ij})$ de orden $m \times n$ con

$$c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \quad \text{para todo } i = 1, \ldots, m \text{ y } j = 1, \ldots, n.$$

Obsérvese que el producto de dos matrices solo está definido si el número de columnas de la primera matriz coincide con el número de filas de la segunda.

- Una matriz $A = (a_{ij})$ es **cuadrada** si tiene el mismo número de filas que de columnas, y de orden n si tiene n filas y n columnas. Se llama **diagonal principal** al conjunto de elementos $a_{11}, a_{22}, \ldots, a_{nn}$.
- Una **matriz diagonal** es una matriz cuadrada que tiene algún elemento distinto de cero en la diagonal principal y ceros en el resto de elementos.
- Una matriz cuadrada con ceros en todos los elementos por encima (debajo) de la diagonal principal se llama **matriz triangular inferior (superior)**.
- Una matriz diagonal con unos en la diagonal principal se denomina **matriz identidad** y se denota por I. Es la única matriz cuadrada tal que $AI = IA = A$ para cualquier matriz cuadrada A.

- Una **matriz simétrica** es una matriz cuadrada A tal que $A = A^T$.
- La **matriz cero** es una matriz con todos sus elementos iguales a cero.
- Decimos que una matriz cuadrada A es **invertible** (o *regular* o *no singular*) si existe una matriz cuadrada B tal que $AB = BA = I$. Se dice entonces que B es la **matriz inversa** de A y se denota por A^{-1}. (Una matriz que no es invertible se dice *singular*.)
- Si una matriz A es invertible, su inversa también lo es y $A^{-1})^{-1} = A$.
- Si A y B son dos matrices invertibles, su producto también lo es y $(AB)^{-1} = B^{-1}A^{-1}$.

El **determinante** de una matriz solo está definido para matrices cuadradas y su valor es un escalar. El determinante de una matriz A cuadrada de orden n se denota por $|A|$ o $\det(A)$, y se define como

$$\det(A) = \sum_j (-1)^{i+j} a_{ij} \cdot \det(A_{ij}),$$

donde la suma se toma para todas las $n!$ permutaciones de grado n y s es el número de intercambios necesarios para poner el segundo subíndice en el orden $1, 2, \ldots, n$.

Algunas propiedades de los determinantes son:

- $\det(A^T) = \det(A)$
- $\det(AB) = \det(A)\det(B)$
- $\det(A^{-1}) = \frac{1}{\det(A)}$
- Si dos filas o dos columnas de una matriz coinciden, el determinante de esta matriz es cero.

- Cuando se intercambian dos filas o dos columnas de una matriz, su determinante cambia de signo.
- El determinante de una matriz diagonal es el producto de los elementos de la diagonal.
- Si denotamos por A_{ij} la matriz de orden $(n-1)$ que se obtiene de eliminar la fila i y la columna j de la matriz A, llamamos **menor complementario** asociado al elemento a_{ij} de la matriz A al $\det(A_{ij})$.
- Se llama k-**ésimo menor principal** de la matriz A al determinante de la submatriz principal de orden k.
- Definimos el **cofactor** del elemento a_{ij} de la matriz A por $\Delta_{ij} = (-1)^{i+j}\det(A_{ij})$.

- Si A es una matriz invertible de orden n, entonces

$$A^{-1} = \frac{1}{\det(A)} C,$$

donde C es la matriz de elementos Δ_{ij} para todo $i, j = 1, 2, \ldots, n$. Obsérvese entonces que una matriz cuadrada es invertible si y sólo si su determinante es distinto de cero.

- Si A es una matriz cuadrada de orden n, un número λ es un **valor propio** de A si existe un vector no nulo v tal que $Av = \lambda v$. Al vector v se le llama **vector propio** asociado al valor propio λ.

- El valor propio λ es solución de la **ecuación característica**

$$\det(A - \lambda I) = 0,$$

  donde $\det(A - \lambda I)$ se llama **polinomio característico**. Este polinomio es de grado n en λ y tiene n valores propios (no necesariamente distintos).

Para medir la **longitud** de los vectores y el **tamaño** de las matrices se suele utilizar el concepto de **norma**, que es una función que toma valores reales. Un ejemplo simple en el espacio euclidiano tridimensional es un vector $v = (v_1, v_2, v_3)$, donde v_1, v_2 y v_3 son las distancias a lo largo de los ejes x, y, z respectivamente.

La **longitud del vector** v (es decir, la distancia del punto $(0,0,0)$ al punto (v_1, v_2, v_3)) se calcula como

$$\|v\| = \sqrt{v_1^2 + v_2^2 + v_3^2},$$

donde la notación $\|v\|$ indica que esta longitud se refiere a la *norma euclidiana* del vector v. De forma similar, para un vector v de dimensión n, $v = (v_1, v_2, \ldots, v_n)$, la norma euclidiana se calcula como

$$\|v\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

UACM Universidad Autónoma de la Ciudad de México Nada humano me es ajeno

Este concepto puede extenderse a una matriz $m \times n$, $A = (a_{ij})$, de la siguiente manera:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2},$$

que recibe el nombre de **norma de Frobenius**.
Hay otras alternativas a las normas euclidiana y de Frobenius. Dos normas usuales son la **norma 1** y la **norma infinito**:

- La **norma 1** de un vector $v = (v_1, v_2, \ldots, v_n)$ se define como $\|v\|_1 = \sum_{i=1}^{n} |v_i|$. De forma similar, la norma 1 de una matriz $m \times n$, $A = (a_{ij})$, se define como

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{m} |a_{ij}|.$$

- La **norma infinito** de un vector $v = (v_1, v_2, \ldots, v_n)$ se define como $\|v\|_\infty = \max_i |v_i|$. La norma infinito de una matriz $m \times n$, $A = (a_{ij})$, se define como

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|.$$

- Todas las normas son equivalentes en un espacio vectorial de dimensión finita.

El sistema numérico que se utiliza frecuentemente es el sistema decimal, en la que la base de expresión es el 10. Sin embargo las computadoras utilizan el sistema binario, sistema de base 2, es decir solamente $\{0, 1\}$.

Proposición

Para cualquier número natural N, existen $a_0, a_1, a_2, \ldots, a_K$, con $a_i \in \mathbb{R}$ tales que

$$N = a_K \times 2^K + a_{K-1} \times 2^{K-1} + a_{K-2} \times 2^{K-2} + \cdots + a_1 \times 2 + a_0 \times 2^0 \quad (3)$$

Para ver lo anterior lo que tenemos que hacer es calcular $\frac{N}{2}$, es decir, $\frac{N}{2} = P_0 + \frac{a_0}{2}$, donde $P_0 = a_K \times 2^{K-1} + a_{K-1} \times 2^{K-2} + a_{K-2} \times 2^{K-3} + \cdots + a_1 \times 2^0$, es decir a_0 es el resto de dividir N entre 2. Ahora hagamos lo mismo para P_0:

$$\frac{P_0}{2} = a_K \times 2^{K-2} + a_{K-1} \times 2^{K-3} + a_{K-2} \times 2^{K-4} + \cdots + a_2 \times 2^0 + \frac{a_1}{2},$$

por lo tanto

$$\frac{P_0}{2} = P_1 + \frac{a_1}{2},$$

donde

$$P_1 = a_K \times 2^{K-2} + a_{K-1} \times 2^{K-3} + a_{K-2} \times 2^{K-4} + \cdots + a_2 \times 2^0,$$

es decir P_1 es el resto de dividir P_0 entre 2. Siguiendo este procedimiento de manera análoga hasta que encontremos un valor K tal que $P_K = 0$. Por lo tanto tenemos el siguiente algoritmo: Para un valor N natural, los términos a_k en la ecuación 3 se encuentran

$$\begin{aligned}
N &= 2P_0 + a_0, \\
P_0 &= 2P_1 + a_1, \\
&\vdots \\
P_{K-2} &= 2P_{K-1} + a_{K-1}, \\
P_{K-1} &= 2P_K + a_K, \\
P_K &= 0.
\end{aligned} \tag{4}$$

UACM Universidad Autónoma de la Ciudad de México Nada humano me es ajeno

Ejemplo

Convertir 24563

$$
\begin{aligned}
24563 &= 12281 \times 2 + 1, & a_0 &= 1 \\
12281 &= 6140 \times 2 + 1, & a_1 &= 1 \\
6140 &= 3070 \times 2 + 0, & a_2 &= 0 \\
3070 &= 1535 \times 2 + 0, & a_3 &= 0 \\
1535 &= 767 \times 2 + 1, & a_4 &= 1 \\
767 &= 383 \times 2 + 1, & a_5 &= 1 \\
383 &= 191 \times 2 + 1, & a_6 &= 1 \\
191 &= 95 \times 2 + 1, & a_7 &= 1 \\
95 &= 47 \times 2 + 1, & a_8 &= 1
\end{aligned}
$$

$$
\begin{aligned}
47 &= 23 \times 2 + 1, & a_9 &= 1 \\
23 &= 11 \times 2 + 1, & a_{10} &= 1 \\
11 &= 5 \times 2 + 1, & a_{11} &= 1 \\
5 &= 2 \times 2 + 1, & a_{12} &= 1 \\
2 &= 1 \times 2 + 0, & a_{13} &= 0 \\
1 &= 0 \times 2 + 1, & a_{14} &= 1
\end{aligned}
$$

Por lo tanto el número binario es:
$(24563)_{10} = (101111111110011)_2$.

Proposición

Sea $Q \in \mathbb{R}$, *tal que* $0 < Q < 1$, *entonces existen términos* $b_1, b_2, \ldots, b_k$ *tales que* $Q = 0.b_1 b_2 b_3 \cdots b_k$, *y por tanto*

$$Q = b_1 \times 2^{-1} + b_2 \times 2^{-2} + b_3 \times 2^{-3} + \cdots + b_k \times 2^{-k} + \cdots \quad (5)$$

Si multiplicamos Q por 2, se tiene que

$$2Q = b_1 + b_2 \times 2^{-1} + b_3 \times 2^{-2} + b_4 \times 2^{-3} + \cdots + b_k \times 2^{-k+1} + \cdots$$

Si $F_1 = frac(2Q)$, con $frac(x)$ la parte fraccionaria de x, y $b_1 = [[2Q]]$, donde $[[x]]$ es la parte entera de x, entonces

$$F_1 = b_2 \times 2^{-1} + b_3 \times 2^{-2} + b_4 \times 2^{-3} + \cdots + b_k \times 2^{-k+1} + \cdots,$$

de donde

$$2F_1 = b_2 \times 2^0 + b_3 \times 2^{-1} + b_4 \times 2^{-2} + \cdots + b_k \times 2^{-k+2} + \cdots = b_2 + F_2,$$

donde $F_2 = frac(2F_1)$, y $b_2 = [[2F_1]]$. Procediendo de manera análoga para el resto de los términos se tienen las suceciones $\{b_k\}$ y $\{F_k\}$, dadas por $b_k = [[2F_{k-1}]]$ y $F_k = frac(2F_{k-1})$, con $b_1 = [[2Q]]$ y $F_1 = frac(2Q)$. Por lo tanto se tiene la representación binaria de Q dada por

$$Q = \sum_{i=1}^{\infty} b_i 2^{-i} \tag{6}$$

Ejemplo

Convertir el número 3,5786*. Sea* $Q = 0{,}5786$*, entonces*

$$
\begin{aligned}
2Q &= 1{,}1572,\ b_1 = [[1{,}1572]] = 1,\ F_1 = frac(1{,}1572) = 0{,}1572 \\
2F_1 &= 0{,}3144,\ b_2 = [[0{,}3144]] = 0,\ F_2 = frac(0{,}3144) = 0{,}3144 \\
2F_2 &= 0{,}6288,\ b_3 = [[0{,}6288]] = 0,\ F_3 = frac(0{,}6288) = 0{,}6288 \\
2F_3 &= 1{,}2576,\ b_4 = [[1{,}2576]] = 1 F_4 = frac(1{,}2576) = 0{,}2576 \\
2F_4 &= 0{,}5152,\ b_5 = [[0{,}5152]] = 0,\ F_5 = frac(0{,}5152) = 0{,}5152 \\
2F_5 &= 1{,}0304,\ b_6 = [[1{,}0304]] = 1 F_6 = frac(1{,}0304) = 0{,}0304 \\
2F_6 &= 0{,}0608,\ b_7 = [[0{,}0608]] = 0,\ F_7 = frac(0{,}0608) = 0{,}0608 \\
2F_7 &= 0{,}1216,\ b_8 = [[0{,}1216]] = 0,\ F_8 = frac(0{,}1216) = 0{,}1216
\end{aligned}
$$

De lo anterior se tiene que:

$$0{,}5786 = (0{,}10010100\ldots)_2$$

Ejemplo

1. *Convertir los siguientes números de base 10 a base 2.*
    1. 324
    2. 27
    3. 1423
    4. 235,25
    5. 41,596