

Notas sobre Métodos Numéricos con R

Carlos E. Martínez-Rodríguez
Universidad Autónoma de la Ciudad de México
Academia de Matemáticas
`carlos.martinez@uacm.edu.mx`

Índice

1. Introducción	1
2. Revisión de temas importantes	2
2.1. Cálculo	2
2.1.1. Límites y continuidad	2
2.1.2. Continuidad y convergencia de sucesiones	2
2.1.3. Derivabilidad	3
2.1.4. Integración	4
2.1.5. Polinomios de Taylor	5
2.1.6. Teoremas adicionales	6
2.2. Álgebra Lineal	7
2.2.1. Matrices	7
2.2.2. Operaciones con matrices	7
2.2.3. Matrices especiales	8
2.2.4. Determinante de una matriz	8
2.2.5. Valores propios y vectores propios	9
2.2.6. Normas vectoriales y normas matriciales	9
3. Introducción al uso de R	10
3.1. Sesiones en RStudio	10
3.2. Uso de R	10
3.3. Definiciones de variables y asignación de valores	11

1. Introducción

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.

2. Revisión de temas importantes

2.1. Cálculo

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.

2.1.1. Límites y continuidad

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

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

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

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

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 $\lim_{n \rightarrow \infty} x_n = x$ o bien, $x_n \rightarrow x$ cuando $n \rightarrow \infty$, 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**.

2.1.2. Continuidad y convergencia de sucesiones

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 $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$.

Teorema 1 (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.*

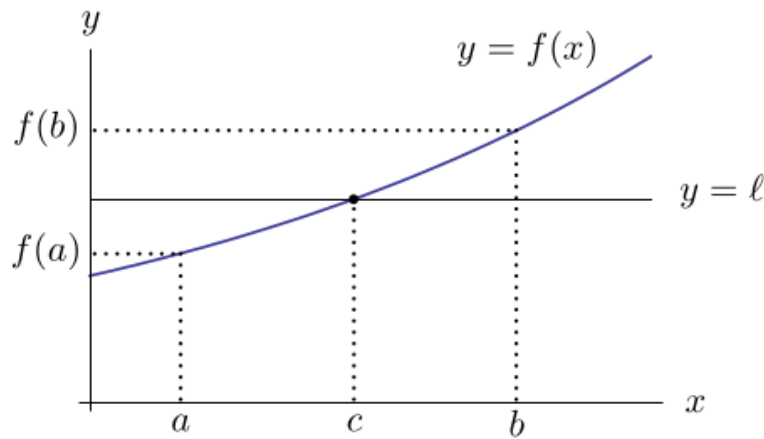


Figura 1: Teorema del valor intermedio o de Bolzano

Nota 1. *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.*

2.1.3. Derivabilidad

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

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

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.

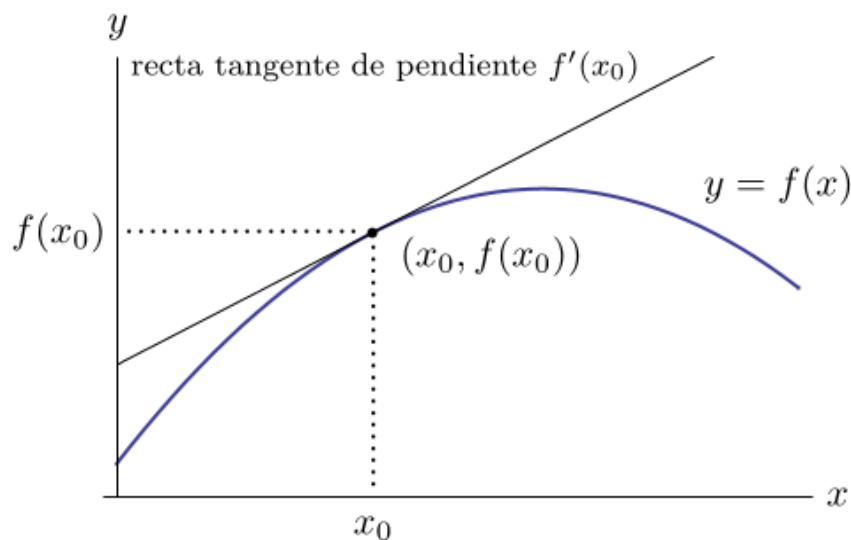


Figura 2: Derivada de una función en un punto

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 2 (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*

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

Geoméricamente 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 $(b, f(b))$.

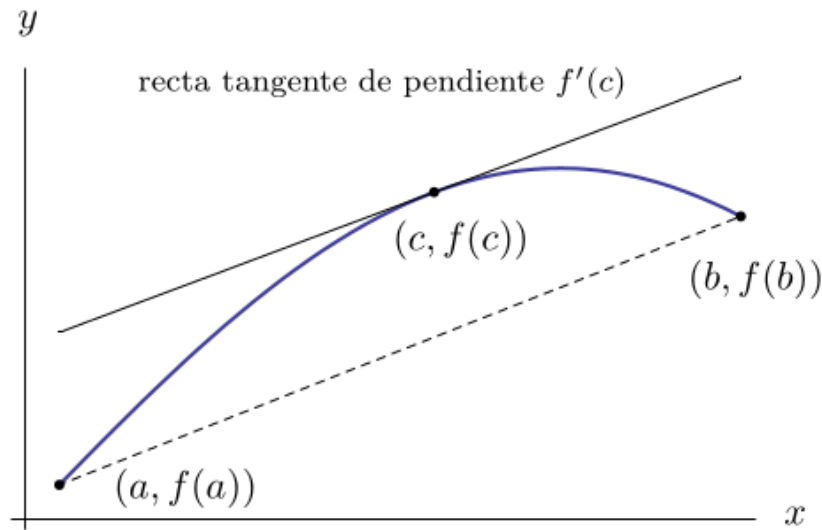


Figura 3: Teorema del valor medio o de Lagrange

Teorema 3 (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.*

2.1.4. Integración

Teorema 4 (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 5 (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 6 (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 2. 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

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

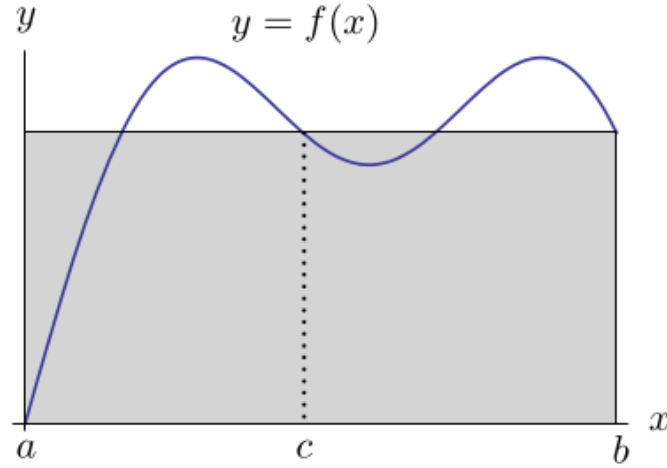


Figura 4: Teorema del valor medio para integrales

2.1.5. Polinomios de Taylor

Teorema 7 (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

$$P_n(x) = f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2!}h^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}h^n = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}h^k, \quad (1)$$

$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}h^{n+1}, \quad y \quad h = x - x_0. \quad (2)$$

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 .

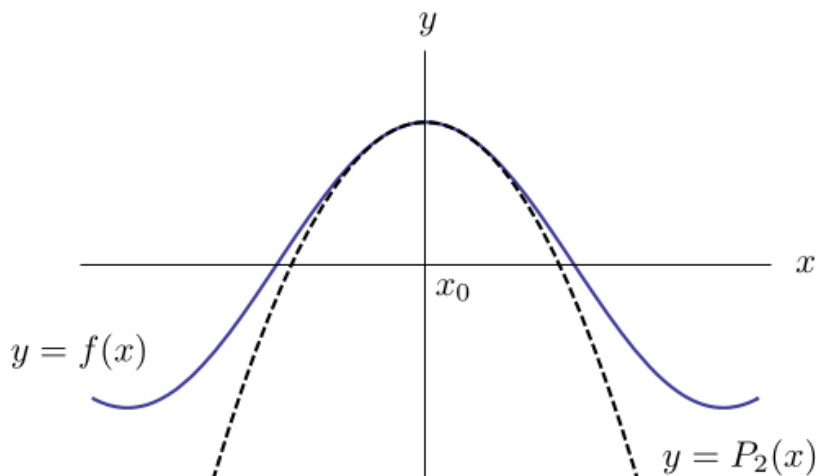


Figura 5: Gráficas de $y = f(x)$ y de su polinomio de Taylor $y = P_2(x)$ alrededor de x_0 .

Nota 3. La serie infinita que resulta al tomar límite en la expresión de $P_n(x)$ cuando $n \rightarrow \infty$ 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.

2.1.6. Teoremas adicionales

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

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

Definición 2. f es de clase C^n en el intervalo $[a; b]$ si $f^{(n)}$ es continua en $[a; b]$.

Definición 3. f es de clase C^∞ en el intervalo I si f es infinitas veces derivable y continua en I .

Teorema 8. (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 9 (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 10 (Teorema de acotabilidad). Si $f : [a; b] \mapsto \mathbb{R}$ es continua en $[a; b]$, entonces f está acotada en $[a; b]$.

Teorema 11 (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 12 (Teorema Generalizado de Rolle). Si f continua en $[a; b]$, y existen las derivadas $f'(x), f''(x), \dots, f^{(n)}(x)$ en $(a; b)$ y $f(x_0) = f(x_1) = \dots = f(x_n) = 0$ (con $x_0, x_1, \dots, x_n \in [a; b]$) entonces existe ξ perteneciente al $(a; b)$ tal que $f^{(n)}(\xi) = 0$.

Teorema 13 (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 14 (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:

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

2.2. Álgebra Lineal

2.2.1. Matrices

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 $i = 1, 2, \dots, m$ y $j = 1, 2, \dots, n$, 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 $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$.

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.

2.2.2. Operaciones con matrices

- 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 $i = 1, \dots, m$ y $j = 1, \dots, n$.
- 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 $i = 1, \dots, m$ y $j = 1, \dots, n$.
- 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 $i = 1, \dots, m$ y $j = 1, \dots, n$.
- 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

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

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.

2.2.3. Matrices especiales

- 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 $a_{11}, a_{22}, \dots, a_{nn}$.
- 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 $(AB)^{-1} = B^{-1}A^{-1}$.

2.2.4. Determinante de una matriz

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 $1, 2, \dots, n$.

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 $i, j = 1, 2, \dots, n$. Obsérvese entonces que una matriz cuadrada es invertible si y sólo si su determinante es distinto de cero.

2.2.5. Valores propios y vectores propios

- 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).

2.2.6. Normas vectoriales y normas matriciales

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 , $v = (v_1, v_2, \dots, v_n)$, la norma euclidiana se calcula como

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

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 $v = (v_1, v_2, \dots, v_n)$ 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 $v = (v_1, v_2, \dots, v_n)$ 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.

3. Introducción al uso de R

3.1. Sesiones en RStudio

Al utilizar R, existen varios entornos que facilitan la gestión y ejecución de rutinas, *archivos con extensión .R*, el más popular es *RStudio* o bien directamente desde la terminal o ejecutando simplemente *R*, la ventaja de *RStudio* es que permite que en una pantalla podamos visualizar: Consola (lugar donde se ejecutan los comandos directamente), History (el histórico de las variables y funciones definidas mismo que puede guardarse para ser invocado posteriormente), Plots (ventana en la que se muestran los gráficos generados), Help (la ayuda sobre comandos, funciones, sintaxis en R), Files (lugar donde se manejan los archivos, es decir leer, guardar, mover o renombrar archivos), y Packages (espacio para instalar o cargar paquetes de manera gráfica), todo esto para facilitar el manejo y ejecución de rutinas compatibles con R. Por otra parte **Workspace** es un entorno en el que se incluyen todos los objetos definidos, al final de una sesión de R, este entorno puede guardarse una imagen del mismo para ser cargada posteriormente.

Durante el uso de R en ocasiones se requiere limpiar la consola, para esto al presionar **Ctrl+L**.

3.2. Uso de R

- **Constantes:** π , $\exp(1)$
- Las constantes pueden ser de tipo *integer*, *double* o *complex*, el tipo de constante se puede consultar con la función **typeof()**

```
> typeof('mi constante')
[1] "character"
```

- Operadores: $<$, $>$, $>=$, $<=$, $!=$, (Not), \parallel (OR), $\&$ (And), $==$ (comparar)
- Operadores aritméticos: $+$, $-$, $*$, $^$ potencia, $\%\%$ resto de la división entera, $\%/\%$ división entera.
- Logaritmos y exponenciales: **log** logaritmo natural, **log(x,b)** ($\log_b x$) y **exp(x)** (e^x).

- Funciones trigonométricas `cos(x)`, `sin(x)`, `tan(x)`, `acos(x)`, `asin(x)`, `atan(x)`.
- Funciones misceláneas `abs(x)`, `sqrt(x)`, `floor(x)`, `ceiling(x)`, `max(x)`, `sign(x)`.
- Comando `options(digits=k)`:

```
> 1/3.0
[1] 0.3333333
> options(digits=3)
> 1/3
[1] 0.333
> 1/17.0
[1] 0.0588
> options(digits=3)
> 1/17.0
[1] 0.0588
> options(digits=5)
> 1/17.0
[1] 0.058824
> options(digits=9)
> 1/17.0
[1] 0.0588235294
```

- Comando `round(x,n)` redondea x a n decimales, el valor por defecto es $n = 6$.
- Comando `cat('caracter1','caracter2')` concatena dos cadenas o valores y el resultado lo convierte a un objeto tipo *string*.

3.3. Funciones

```
nombrefun = function(a1,a2,...,an) {
# código ...
instruccion-1
instruccion-2
# ...
return( ... ) #valor que retorna (o también la última instrucción, si ésta retorna algo)
}
```

3.4. Definiciones de variables y asignación de valores