

Taller de Métodos Numéricos

Guía de estudio

Marcos Prunello

Cecilia Rapelli

2020-05-26

Índice general

Acerca del Taller	4
Horario y lugar de cursado	4
Consultas	4
Material de estudio	4
Evaluación	4
Campus virtual	4
1 Introducción a los Métodos Numéricos	5
1.1 Cifras significativas	5
1.2 Notación Científica o Exponencial	6
1.3 Sistemas numéricos	6
1.4 ¿Por qué nos interesa el sistema binario?	7
1.5 Formato de coma flotante	7
1.6 Propiedades importantes del formato IEEE-754	8
1.7 Análisis de los errores	11
2 Solución Numérica de Ecuaciones No Lineales	13
2.1 Generalidades	13
2.2 Método de las Aproximaciones Sucesivas o del Punto Fijo	14
2.3 Método de Newton-Raphson	19
2.4 Método de von Mises	23
2.5 Método de Newton-Raphson de 2º Orden	23
3 Solución de Sistemas de Ecuaciones Lineales	25
3.1 Generalidades	25
3.2 Repaso	25
3.3 Notación	26
3.4 Métodos de Resolución de Sistemas de Ecuaciones	26
3.5 Métodos exactos	26
3.6 Métodos Aproximados o Iterativos	32
4 Aproximación Polinomial: interpolación y extrapolación	37
4.1 Generalidades	37
4.2 Diferencias finitas	37
4.3 Interpolación de Newton para incrementos constantes	39
4.4 Interpolación de Lagrange	44
5 Aproximación Polinomial: integración y derivación numérica	52
5.1 Integración numérica	52
5.2 Derivación numérica	59
6 Autovalores y Autovectores	60
6.1 Generalidades	60
6.2 Obtención los autovalores y autovectores	63
6.3 Método de Krylov	64
6.4 Método de Faddeev-LeVerrier	67

6.5 Método de Aproximaciones Sucesivas o de las Potencias	68
Anexo: Teoremas útiles	70
Teorema del Valor Intermedio o de Bolzano	70
Teorema del Valor Medio	70
Teorema de Taylor	70
Teorema del Punto Fijo	70
Teorema de Newton-Raphson	72
Deducción de la fórmula de recurrencia para el Método de Newton-Raphson de 2º Orden	73

Acerca del Taller

La presente guía resume algunos conceptos desarrollados en el Taller de Métodos Numéricos de la Licenciatura en Estadística en la Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario. La misma irá siendo revisada a lo largo del cuatrimestre y no está exenta de presentar errores o expresar ideas que puedan ser mejoradas. Contamos con la ayuda de ustedes para mejorarla entre todos.

Horario y lugar de cursado

Martes de 8:00 a 10:00 en el Laboratorio de Estadística (2do piso).

Consultas

A través del Foro de Consultas en el Campus Virtual o al finalizar las clases.

Material de estudio

El material de estudio está compuesto por esta guía y prácticas que, junto con cualquier otro material que necesitemos, estarán disponibles en nuestro espacio en el Campus Virtual de la UNR.

Todo el material fue creado utilizando el lenguaje de programación estadística **R** y las herramientas del entorno de desarrollo integrado **RStudio**.

Evaluación

En construcción.

Campus virtual

Además de alojar todo el material del curso, utilizaremos este espacio para la entrega de trabajos y para realizar consultas, que esperamos puedan ser debatidas y respondidas entre el estudiantado.

Unidad 1

Introducción a los Métodos Numéricos

- Son aquellos algoritmos que permiten resolver de forma aproximada problemas matemáticos que involucran el cálculo de determinados valores y no pueden ser abordados mediante técnicas analíticas, o cuya resolución analítica exacta tiene un costo o complejidad muy elevados.
- Por ejemplo, si tenemos que hallar el valor de x tal que $x + 2 = 5$, el cálculo es muy sencillo.
- Pero se complica si tenemos que calcular:

$$F(x_0) = \int_{-\infty}^{x_0} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

- Es decir, el Análisis Numérico trata de diseñar métodos para aproximar, de una manera eficiente, las soluciones de problemas expresados matemáticamente.
- Naturalmente, a los métodos numéricos les interesa controlar la diferencia entre la solución aproximada y el valor verdadero, que recibe el nombre de **error**.
- Están relacionados con muchas disciplinas, como la algorítmica y la programación, y se aplican en diversas áreas, como en la Estadística.
- El estudio y desarrollo de los métodos numéricos abarca los siguientes aspectos:
 - Análisis de las propiedades de convergencia de los métodos numéricos
 - Estudio de la programación de los métodos numéricos (implementación)
 - Aplicación de los métodos numéricos en un determinado campo
- Algunos de los problemas que toca el Análisis Numérico son los siguientes:
 - Resolución aproximada de ecuaciones algebraicas y sistemas de ecuaciones no lineales.
 - Problemas de optimización, en los que se maximiza o se minimiza una función.
 - Problemas de tipo matricial (hallar valores propios, invertir matrices, etc.)
 - Resolución de sistemas de ecuaciones lineales con gran número de ecuaciones e incógnitas que por su costo de cálculo son irresolubles por métodos clásicos como la regla de Cramer.
 - Resolución aproximada de ecuaciones diferenciales.
 - Problemas de interpolación.
 - Problemas de integración o derivación aproximada de funciones poco manejables.

Ejemplo 1

- No hay dudas que:

$$(a + b) - b = a$$

- ¿Pero qué sucede si hacemos $(10^{-9} + 10^9) - 10^9$ en la calculadora?
- En esta unidad nos dedicaremos a repasar y definir conceptos relacionados con la representación numérica de magnitudes, para intentar entender qué sucedió con la operación anterior en nuestra calculadora.

1.1. Cifras significativas

- Las cifras significativas de un número son las que aportan alguna información.
- Reglas:

- Cualquier dígito distinto de cero es significativo. Ej: 348 tiene 3 cifras significativas.
- Los ceros ubicados entre dos dígitos distintos de cero son significativos. Ej: 403 tiene 3; 10,609 tiene 5.
- Los ceros a la izquierda del primer dígito diferente de cero NO son significativos. Ej: 0,0042 tiene 2.
- En números que tienen coma decimal, ceros a la derecha son significativos. Ej: 0,050 tiene 2; 2.00 tiene 3 (estos ceros reflejan la precisión del número).
- En números que no tienen coma decimal, ceros a la derecha pueden ser o no significativos. Ej: 700 puede tener una (el 7), 2 (70) o 3 (700) cifras significativas, dependiendo de la precisión en la obtención del número.
- Los números exactos tienen infinitas cifras significativas pero no se reportan. Ej: si uno cuenta lápices y hay dos, el número de lápices es 2,0000...

1.2. Notación Científica o Exponencial

- Es un sistema que facilita la escritura de números muy grandes o muy pequeños.
- La representación en notación científica de un número real r es $r = \pm c \times b^e$, donde:
 - c : es el coeficiente (real).
 - b : es la base (10 en el sistema decimal, puede ser otra).
 - e : es el exponente u “orden de magnitud”, que eleva la base a una potencia.
 - \pm : es el signo del coeficiente, indica si el número es positivo o negativo.
- Si el coeficiente es un entero entre el 1 y el 9, seguido de una coma y de varios dígitos fraccionarios, se dice que el número está expresado con **notación científica estándar**.

1.2.1. Ejemplo 2:

- [illegible]

1.3. Sistemas numéricos

- Un sistema de numeración es un conjunto de símbolos y reglas que permiten construir todos los números válidos.
- Estamos acostumbrados a utilizar el sistema decimal, el cual está compuesto por 10 símbolos o cifras: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9.
- Es un sistema posicional, es decir, la posición ocupada por cada dígito tiene un significado exacto y determina la contribución de la cifra al valor numérico. Por ejemplo, 35 y 53 tienen las mismas cifras pero significados distintos.
- Cada cifra es multiplicada por una potencia de 10, con el exponente determinado por la posición de la cifra con respecto al punto decimal.
- Por ejemplo, el número decimal 1563 se puede escribir en forma desarrollada utilizando potencias con base 10 así:

$$1563 = 1000 + 500 + 60 + 3 = (1 \times 10^3) + (5 \times 10^2) + (6 \times 10^1) + (3 \times 10^0)$$

- Son los coeficientes 1, 5, 6 y 3 los que definen la representación de este número como 1563 en el sistema decimal.
- Pero si en vez de usar potencias con base 10, usamos potencias con base 2, el número 1563 se escribe como:

$$\begin{aligned} 1563 &= 1024 + 512 + 16 + 8 + 2 + 1 \\ &= (1 \times 2^{10}) + (1 \times 2^9) + (1 \times 2^4) + (1 \times 2^3) + (1 \times 2^1) + (1 \times 2^0) \\ &= (1 \times 2^{10}) + (1 \times 2^9) + (0 \times 2^8) + (0 \times 2^7) + (0 \times 2^6) + \\ &\quad (0 \times 2^5) + (1 \times 2^4) + (1 \times 2^3) + (0 \times 2^2) + (1 \times 2^1) + (1 \times 2^0) \end{aligned}$$

- Como en el sistema decimal, son los coeficientes los que definen la representación del número, por lo que 1563 en **sistema binario** o en **base dos** es igual a $11000011011_{(2)}$.

1.4. ¿Por qué nos interesa el sistema binario?

- Porque es el sistema de representación numérica que utilizan las computadoras.
- Este sistema es natural para las computadoras ya que su memoria consiste de un enorme número de dispositivos de registro electrónico, en los que cada elemento sólo tiene los estados de “encendido” y “apagado”.
- Estos elementos constituyen la unidad mínima de información, sólo pueden tomar dos valores posibles, 0 o 1, y reciben el nombre de **bit** (*binary digit*).
- Aunque nosotros no nos damos cuenta, toda operación numérica que le indicamos a la computadora en sistema decimal, es traducida y procesada internamente en binario.
- Por lo tanto es muy importante entender cómo opera la computadora, para entender qué sucede con las operaciones que queremos que realice. Por ejemplo...

1.4.1. Ejemplo 3

- Escribir un programa para realizar las siguientes operaciones, empleando estructuras iterativas para las sumatorias:
 - a) $10\,000 - \sum_{i=1}^{100\,000} 0,1$
 - b) $10\,000 - \sum_{i=1}^{80\,000} 0,125$
- ¿Cuál es el resultado exacto en estos cálculos? ¿Qué resultados arrojó la computadora?
- ¿Por qué ocurre esto?
 - Pensemos en el número decimal periódico $1/3 = 0, \overline{3}$.
 - Para aproximarlos, sólo podemos usar una cantidad finita de cifras, por ejemplo, 0,333 o 0,33333. Estas aproximaciones guardan cierto error, que depende de la cantidad de cifras empleadas.
 - Con los números binarios ocurre exactamente lo mismo.
 - $0,1_{(10)} = 0,0001100110011..._{(2)} = 0,0\overline{0011}_{(2)}$. Es decir, la representación de 0,1 en binario es periódica, la computadora necesariamente debe redondear o truncar para almacenar y operar.
 - Por esta razón, sumar 100 mil veces 0,1 no da exactamente 10000.
 - Por el contrario, 0,125 en binario no es periódico, la computadora lo puede representar exactamente y no se produjo error.

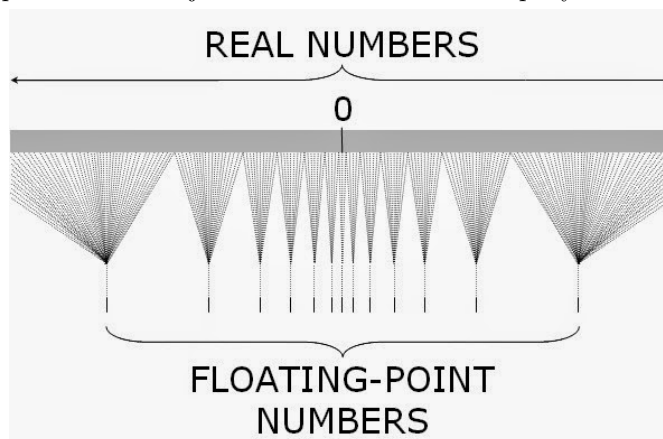
1.5. Formato de coma flotante

- La **representación de coma flotante** (en inglés, *floating point*) es una forma de notación científica usada en los computadores con la cual se pueden representar números reales extremadamente grandes y pequeños de una manera muy eficiente y compacta, y con la que se pueden realizar operaciones aritméticas.
- El estándar que establece las reglas para este tipo de representación se llama **IEEE-754**.
- En este formato, la base es 2 y tanto el exponente como el coeficiente pueden contener una cierta cantidad limitada de cifras binarias, lo cual implica que dada una cantidad x , la computadora sólo es capaz de almacenar una aproximación binaria a x :

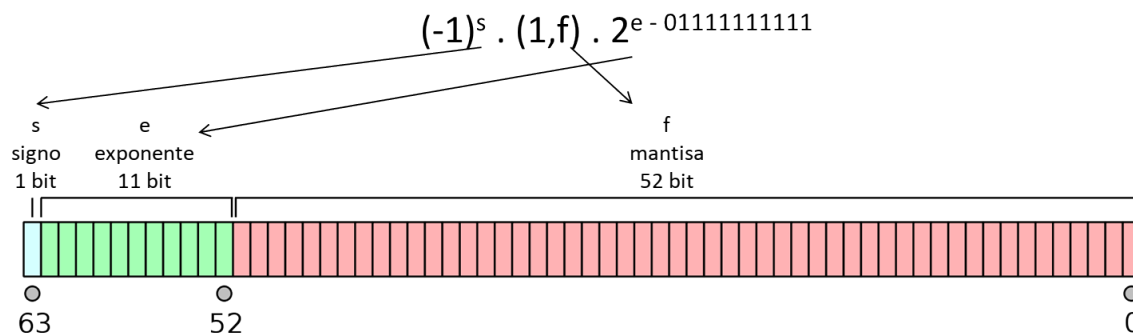
$$x \approx \pm m \times 2^n$$

donde el exponente n y coeficiente m (también llamado **mantisa**) son números binarios con una longitud establecida.

- A través de este sistema, las computadoras sólo pueden representar un subconjunto de los números racionales y no pueden representar números irracionales como π o $\sqrt{3}$ dado que tienen infinitos decimales no periódicos.
- Esto hace que en la representación surjan los errores de redondeo que ya vimos en el ejemplo 3.



- El **IEEE-754** define dos tipos de formatos: el de precisión simple (en el cual cada número ocupa 32 bit de memoria) y el de precisión doble (un número ocupa 64 bit).
- El formato de doble precisión en 64 bit, empleado actualmente en casi todas las computadoras, tiene la siguiente estructura.



- El bit que indica el signo, s , vale 0 si el número es positivo o 1 si es negativo.
- El exponente e ocupa 11 bit dando lugar a $2^{11} = 2048$ valores distintos. Se dice que es un exponente sesgado porque se le resta $01111111111_{(2)} = 1023_{(10)}$, por lo que puede variar entre -1023 y 1024.
- La mantisa ($1, f$) se registra *normalizada*, es decir con un solo dígito antes de la coma igual a 1, de manera que dicho 1 no se almacena, es implícito. Luego de la coma, la parte fraccionaria f ocupa 52 bit (2^{52} valores distintos).
- Por ejemplo, el siguiente conjunto de 64 bit representa al número decimal -74,5 (verificación opcional).

[illegible]

- En este link se puede encontrar una calculadora que convierte números entre sus representaciones en decimal y en coma flotante.

1.6. Propiedades importantes del formato IEEE-754

- Además de representar números, este sistema permite almacenar ciertos valores especiales. Uno de ellos es *NaN* (*Not a Number*), para resultados no definidos matemáticamente. Por ejemplo:


```
0/0
```

```
[1] NaN
```

```
sqrt(-1)
```

```
[1] NaN
```

Operaciones como $x + NaN$, $NaN \times x$, NaN/x , x/NaN resultan en NaN (excepto $NaN^0 = 1$).

Observación: en R **NaN** no es lo mismo que **NA** (*Not Available*). **NA** es una construcción de R, que no forma parte del IEEE-754, para representar valores desconocidos o faltantes.

- El infinito tiene su propia representación en este estándar y puede ser positivo o negativo. Un valor infinito puede darse como resultado de ciertas operaciones o por exceder los límites de almacenamiento. En R se representan con **Inf** y **-Inf**.

```
Inf
```

```
[1] Inf
```

```
Inf - Inf
```

```
[1] NaN
```

```
Inf / Inf
```

```
[1] NaN
```

- La división por cero resulta en infinito, aunque matemáticamente sea una indefinición. Y como contrapartida, dividir por infinito da cero.

```
3/0
```

```
[1] Inf
```

```
-3/0
```

```
[1] -Inf
```

```
3/Inf
```

```
[1] 0
```

- Infinito es el próximo valor al mayor valor que puede registrarse bajo este formato. En R, existe un objeto llamado **.Machine** que es una lista con información numérica de la máquina en la que se está trabajando, por ejemplo, cuál es el mayor número en coma flotante normalizada. Si a este valor lo multiplicamos, por ejemplo, por 2, tenemos infinito:

```
.Machine$double.xmax
```

```
[1] 1.797693e+308
```

```
.Machine$double.xmax * 2
```

```
[1] Inf
```

- El 0 tiene signo, puede ser $+0$ o -0 . -0 actúa como 0 y se muestra como 0, pero el signo negativo se propaga en las operaciones.

```
-0
```

```
[1] 0
```

```
3 * -0
```

```
[1] 0
```

```
1 / 0
```

```
[1] Inf
```

```
1 / -0
```

```
[1] -Inf
```

- Como ya dijimos, bajo este sistema no podemos representar todos los números reales. El **error de máquina** o **epsilon de máquina** es una medida del espacio que hay entre dos números consecutivos en la computadora. Se define como el menor número $\epsilon > 0$ tal que $1 + \epsilon > 1$.

Si bien $1 + x > 1$ debería verificarse para cualquier $x > 0$, en la computadora no es así, porque si x es muy pequeño, $1 + x$ será redondeado a 1. Por eso, el ϵ de máquina es aquella cantidad mínima necesaria para que $1 + \epsilon$ se distinga de 1.

En R, podemos averiguar el epsilon de máquina así:

```
.Machine$double.eps
```

```
[1] 2.220446e-16
```

Podemos usar la función `print` para elegir la cantidad de decimales que queremos ver y así experimentar con el ϵ de máquina:

```
print(1 + .Machine$double.eps, digits = 20)
```

```
[1] 1.000000000000000222
```

```
print(1 + .Machine$double.eps * 2, digits = 20)
```

```
[1] 1.0000000000000004441
```

```
print(1 + .Machine$double.eps / 2, digits = 20)
```

```
[1] 1
```

Aunque nos parezca despreciable, esto puede ser una limitación a considerar, por ejemplo, si trabajamos con medidas de partículas subatómicas.

Si hacemos $1000 + \epsilon$, encontraremos que el resultado es 1000. Es decir, $1000 + \epsilon$ no se distingue de 1000. Es porque ϵ está definido con respecto a 1. Como los números en coma flotante son cada vez más espaciados cuanto más nos alejamos de 0, a 1000 habrá que sumarle un x algo más grande para que $1000 + x$ se distinga de 1000.

```
print(1000 + .Machine$double.eps, digits = 20)
```

```
[1] 1000
```

```
print(1000 + 100 * .Machine$double.eps, digits = 20)
```

```
[1] 1000
```

```
print(1000 + 300 * .Machine$double.eps, digits = 20)
```

```
[1] 1000.000000000001137
```

- **Overflow (desbordamiento)**. Es el fenómeno que ocurre cuando un cálculo produce un resultado cuya magnitud es mayor que la que se puede almacenar o representar (en R, `.Machine$double.xmax`). Aunque este número probablemente es mucho mayor que cualquier número que utilicemos en la práctica, debería ser tenido en cuenta, por ejemplo, si hacemos cálculos en cosmología.

```
.Machine$double.xmax
```

```
[1] 1.797693e+308
```

```
.Machine$double.xmax * 2
```

```
[1] Inf
```

- **Underflow (desbordamiento).** Es el fenómeno que ocurre cuando cuando el resultado de una operación en coma flotante es menor en magnitud (más cercano a cero) que el menor número representable (en R, `.Machine$double.xmin`). Es difícil mostrar un ejemplo de esto porque, dependiendo de la máquina, la representación en coma flotante puede recurrir a números no normalizados para lograr números incluso por fuera de los límites de desbordamiento:

```
# Mínimo representable en coma flotante normalizada
.Machine$double.xmin
```

```
[1] 2.225074e-308
```

```
# Igual es posible conseguir números menores con representación no normalizada
.Machine$double.xmin / 2
```

```
[1] 1.112537e-308
```

Observación: el mínimo representable `.Machine$double.xmin` es menor que el ϵ de máquina ya que los números en coma flotante son más densos alrededor del 0 que del 1.

1.7. Análisis de los errores

- Como ya hemos dicho, los métodos numéricos proporcionan una solución aproximada de los problemas que tratan de resolver.
- Y aunque es cierto que algunos de los métodos que se estudiarán idealmente proporcionarían una solución exacta, esto tampoco ocurrirá por las aproximaciones que realiza la computadora en la representación numérica.
- Denominamos **error** a la diferencia entre la solución que los métodos, una vez programados, devuelven y la solución exacta del problema que se trata de resolver. Hay dos tipos fundamentales de errores.

1.7.1. Error de truncamiento

- Dado que un método numérico propone un algoritmo para resolver de forma aproximada un problema que no se puede resolver mediante métodos analíticos, se llama **error de truncamiento** a la diferencia entre el valor aproximado propuesto por el método y la solución exacta del problema.
- Ocurre cuando un proceso que requiere un número infinito de pasos se detiene en un número finito de pasos.
- Por ejemplo, podemos recordar el desarrollo en serie de Taylor de la función $f(x) = e^{x^2}$:

$$e^{x^2} = 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{3!} + \dots + \frac{x^{2n}}{n!} + \dots$$

- Si nos quedamos sólo con los primeros 4 términos, estamos aproximando una suma que tiene infinita cantidad de sumandos sólo con los primeros 4, de manera que dicha aproximación presentará un error de truncamiento.
- Este tipo de error no depende directamente del sistema numérico que se emplee.

1.7.2. Error de redondeo

- Resulta de reemplazar un número por su forma en coma flotante, es decir, por su representación en la computadora mediante un número finito de bits.
- Recibe este nombre ya sea que la aproximación se realice con redondeo o poda (también conocida como *truncamiento*, pero no debe confundirse con el error de truncamiento).
- El error de redondeo está ligado fundamentalmente al tipo de precisión que se emplee (algo determinado por el procesador y el software usados).
- Sin embargo, el efecto final de los errores de redondeo depende también del algoritmo propuesto por el método numérico y por la forma de programarlo.
- Existen operaciones que son especialmente sensibles a los errores de redondeo o un algoritmo puede hacer que los mismos se amplifiquen.

1.7.3. Ejemplos de operaciones “delicadas”

Sustracción de números casi iguales en valor absoluto

- Da lugar a importantes pérdidas de precisión.
- El resultado tiene menos cifras significativas que los valores originales (**pérdida de cifras significativas o cancelación catastrófica**).
- Por ejemplo: sean $p = 3.1415926536$ y $q = 3.1415957341$.
- Tienen 11 cifras significativas cada uno.
- Sin embargo, $p - q = -0.0000030805$ tiene sólo 5 cifras significativas.
- Esto puede producir una reducción en la precisión final de la respuesta calculada.

División por cantidades pequeñas

- Un error mínimo en el dividendo se traduce en uno mucho mayor en el resultado, de modo que la falta de precisión podría ocasionar un *overflow* o pérdida de cifras significativas.
- Veremos un ejemplo de esto cuando resolvamos sistemas de ecuaciones lineales con la estrategia de pivoteo en el algoritmo de eliminación de Gauss.
- Dado que los números de punto flotante están más concentrados cerca del cero entonces al dividir por un número más grande es más probable conseguir una mejor aproximación.

Adición de un número grande y uno pequeño

- Puede hacer que el pequeño desaparezca.
- En ciertos casos esto no ocasiona un problema ya que, si tenemos un número de gran magnitud probablemente podamos considerar al más pequeño despreciable.
- Sin embargo debe tenerse mucho cuidado con el orden de las operaciones ya que si, por ejemplo, sumamos una gran cantidad de números pequeños entre ellos (que juntos tienen un peso considerable) y luego se lo sumamos a un número grande, todo funcionará correctamente.
- Pero si vamos sumando uno por uno los números pequeños al grande entonces en cada paso el número pequeño será considerado despreciable y llegaremos a un resultado erróneo.

1.7.4. Medida del error

Existen dos formas fundamentales de medir los errores que se cometen en la aproximación de la solución de un problema mediante un método numérico. Siendo x el valor exacto y \hat{x} su aproximación, definimos:

- **Error absoluto:** $E_a = |x - \hat{x}|$. Se mide en las mismas unidades de la variable que se trata de aproximar (por ejemplo, en amperios si se quiere aproximar la corriente que circula por un circuito eléctrico).
- **Error relativo:** $E_r = \frac{|x - \hat{x}|}{|x|}$, $x \neq 0$. Relaciona el error obtenido con la magnitud de la propia solución y se interpreta como porcentaje. Esta medida del error es invariante a los cambios de escala, ya que es independiente de las unidades de medida.

1.7.5. Error propagado

- Se define como **error propagado** al error que se tiene al final de una cadena de operaciones sucesivas por la existencia de diferentes errores en los pasos intermedios.
- Por ejemplo, si tenemos dos valores exactos p y q con valores aproximados \hat{p} y \hat{q} cuyos errores son ϵ_p y ϵ_q de modo que $\hat{p} = p + \epsilon_p$ y $\hat{q} = q + \epsilon_q$, al realizar la suma entre los valores aproximados:

$$\hat{p} + \hat{q} = (p + \epsilon_p) + (q + \epsilon_q) = (p + q) + (\epsilon_p + \epsilon_q)$$

- Si bien es normal que en una cadena los errores iniciales se propaguen, es deseable que un error pequeño en el comienzo produzca errores pequeños en el resultado final.
- Un algoritmo con esta cualidad se llama **estable** (el error se puede acotar), en caso contrario se dice **inestable**.
- Supongamos que ϵ representa un error inicial y que $\epsilon(n)$ representa el crecimiento de dicho error después de n operaciones:
 - Si $|\epsilon(n)| \propto n\epsilon$, el crecimiento es **lineal**.
 - Si $|\epsilon(n)| \propto k^n \epsilon$, el crecimiento es **exponencial**.
 - Si $k > 1$, el error crece cuando $n \rightarrow \infty$ sin que podamos acotarlo, el proceso es inestable.
 - Si $0 < k < 1$, el error decrece cuando $n \rightarrow \infty$, se puede acotar, el proceso es estable.

Unidad 2

Solución Numérica de Ecuaciones No Lineales

2.1. Generalidades

- En esta unidad estudiaremos uno de los problemas más básicos y antiguos de la aproximación numérica: la **solución de ecuaciones**.
- Consiste en obtener una solución de una ecuación $F(x) = 0$.
- Se presenta en una gran variedad de problemas.
- Las soluciones de una ecuación se llaman **raíces** o **ceros**.
- Una **ecuación lineal** es una igualdad que involucra una o más variables elevadas a la primera potencia y no contiene productos entre las variables (involucra solamente sumas y restas de las variables). Por ejemplo: $3x + 2 = 8$.
- Para este tipo de ecuaciones es posible hallar analíticamente una expresión para su solución.
- En una **ecuación no lineal** las incógnitas están elevadas a potencias distintas de 1, aparecen en denominadores o exponentes o están afectadas por funciones no lineales (como el logaritmo o las trigonométricas).
- Un tipo de ecuación no lineal es la **ecuación algebraica**, que se trata de un polinomio igualado a cero:

$$P_n(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n = 0$$

donde $a_0 \neq 0, n \in \mathbb{N}$ y a_0, \dots, a_n son constantes.

- Ejemplo: $x^3 - x^2 + 5x - 8 = 2x^5$.
- Sabemos que si, por ejemplo, $n = 2$, la solución de $ax^2 + bx + c = 0$ está dada por la resolvente:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

- Sin embargo, la solución analítica para este tipo de ecuaciones existe sólo para $n \leq 4$.
- Las restantes ecuaciones no lineales se dice que son **trascendentes**, por ejemplo:

$$\begin{aligned} x^3 - \ln x + \frac{3}{x} &= 2 \\ \operatorname{tg}(x + 45) &= 1 + \operatorname{sen}(2x) \\ xe^x &= 1 \\ 5^x &= 9^{x+1}3^x \end{aligned}$$

- En general, tampoco es posible hallar de manera analítica una solución exacta para estas ecuaciones.
- Excepto para algunos problemas, las ecuaciones no lineales carecen de solución exacta, por lo que requieren ser resueltas con métodos numéricos.
- Una técnica fundamental de los métodos numéricos es la **ITERACIÓN** (métodos iterativos).

- Se trata de **repetir un proceso** hasta que se obtiene un resultado para el problema.
- En la unidad se verán distintos métodos iterativos para encontrar las raíces, cada con sus propias ventajas y limitaciones.
- Requieren dos pasos generales:
 1. Determinación de un valor aproximado de la raíz que se busca.
 2. Mejoramiento de la solución hasta lograr un grado de precisión preestablecido.

2.2. Método de las Aproximaciones Sucesivas o del Punto Fijo

2.2.1. Punto fijo

Definición de Punto Fijo

- Un punto fijo de una función $f(x)$ es un número real P tal que $f(P) = P$.
- Ejemplos:
 - $f(x) = x^2 - 3x + 4$, 2 es un punto fijo de f porque $f(2) = 2$.
 - $f(x) = x^2$, 0 y 1 son puntos fijos de f porque $f(0) = 0$ y $f(1) = 1$.

¿Cómo encontrar un punto fijo de $f(x)$?

Sea f una función continua y $p_0, p_1, \dots, p_n, \dots$ una sucesión generada a partir de $p_n = f(p_{n-1})$ con un valor inicial p_0 , es decir:

$$\begin{aligned}
 & p_0 \\
 & p_1 = f(p_0) \\
 & p_2 = f(p_1) \\
 & \vdots \\
 & p_n = f(p_{n-1}) \\
 & \vdots
 \end{aligned}$$

Si $\lim_{n \rightarrow \infty} p_n = P$, entonces P es un punto fijo de $f(x)$.

2.2.2. Empleo del punto fijo para la resolución de ecuaciones

Siendo:

$$F(x) = 0 \tag{2.1}$$

la ecuación a resolver, el **Método de las Aproximaciones Sucesivas** propone reescribirla a través de la ecuación equivalente:

$$f(x) = x$$

de manera que la tarea de hallar un valor de x que satisface (2.1) es lo mismo que hallar un punto fijo de la función $f(x)$.

Entonces, el método para resolver $F(x) = 0$ consiste en:

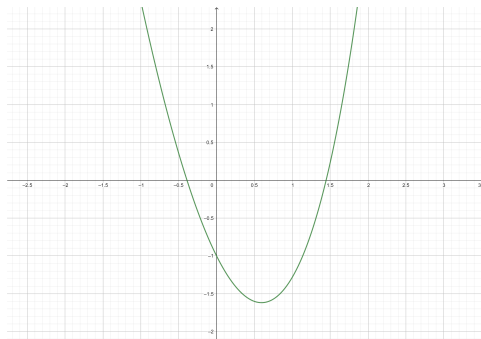
1. Expresar la ecuación en la forma $x = f(x)$.
2. Elegir un valor inicial adecuado x_0 .
3. Realizar el siguiente cálculo iterativo:

$$\begin{aligned}
 x_1 &= f(x_0) \\
 x_2 &= f(x_1) \\
 &\vdots \\
 x_n &= f(x_{n-1}) \\
 &\vdots
 \end{aligned}$$

- Si a medida que n crece los x_n se aproximan a un valor fijo, se dice que el método converge y la iteración se detiene cuando la diferencia entre dos valores consecutivos x_{n-1} y x_n sea tan pequeña como se desee.
- El valor x_n será una raíz aproximada de $F(x)$.

2.2.2.0.1. Ejemplo

- Hallar las raíces de la ecuación no lineal: $F(x) = x^2 - 3x + e^x - 2 = 0$
- Graficamos y vemos que las raíces están cercanas a -0.4 y 1.4.



- Nota: para hacer representaciones gráficas de funciones con rapidez y sencillez pueden utilizar herramientas disponibles online. Por ejemplo, nos gusta mucho usar Geogebra.
- Reescribimos $F(x) = 0$ como $f(x) = x$
- Por ejemplo:

$$F(x) = x^2 - 3x + e^x - 2 = 0$$

$$\Rightarrow \underbrace{\frac{x^2 + e^x - 2}{3}}_{f(x)} = x$$

$$\Rightarrow f(x) = \frac{x^2 + e^x - 2}{3}$$

- Para $x_0 = -1.5$, el proceso converge al valor -0.390271 que consideraremos como la aproximación para la raíz buscada.

Iteracion	x
0	-1.5000000
1	0.1577101
2	-0.2681003
3	-0.3877637
4	-0.3903555
5	-0.3902688
6	-0.3902718
7	-0.3902717

- Verificar estos resultados, pueden hacerlo rápidamente en una planilla de Excel.

2.2.3. Criterios para detener el proceso iterativo

- **Criterios para convergencia:**

1. Error absoluto: $|x_{j+1} - x_j| < \epsilon$
2. Error relativo: $\left| \frac{x_{j+1} - x_j}{x_j} \right| < \epsilon$
3. Error relativo respecto al valor inicial: $\left| \frac{x_{j+1} - x_j}{x_0} \right| < \epsilon$
4. $|F(x_j)| < \epsilon$

- **Criterios para divergencia:**

1. $j > r$, r número máximo de iteraciones
2. $|x_j - x_1| > k$
3. $|F(x_j)| > k$
4. $|x_{j+1} - x_j| > k$
5. $\left| \frac{x_j}{x_1} \right| > k$

Ejemplo

- En cada paso calculamos el error relativo y nos detuvimos cuando el mismo fue menor a 1E-6.

Iteracion	x	Error
0	-1.5000000	NA
1	0.1577101	1.105140e+00
2	-0.2681003	2.699957e+00
3	-0.3877637	4.463383e-01
4	-0.3903555	6.684034e-03
5	-0.3902688	2.222629e-04
6	-0.3902718	7.690434e-06
7	-0.3902717	2.657452e-07

2.2.4. Teorema del Punto Fijo

- Pero esto no funciona siempre, para cualquier f o cualquier x_0 ...
- ¿Cuándo funciona? Cuando se cumplen las condiciones del **Teorema del Punto Fijo**.
- A saber:

Dadas las siguientes condiciones:

- a. f es una función continua en el intervalo $[a, b]$
- b. $f(x) \in [a, b] \quad \forall x \in [a, b]$
- c. f' existe en (a, b) con $|f'(x)| \leq m < 1 \quad \forall x \in (a, b)$

Si x_0 es cualquier número en $[a, b]$, entonces la sucesión definida por

$$x_n = f(x_{n-1}), \quad n \geq 1,$$

converge al único punto fijo que f posee en $[a, b]$.

Ver demostración en el anexo.

Ejemplo

- En el ejemplo anterior, dada la ecuación $F(x) = x^2 - 3x + e^x - 2 = 0$, la reexpresamos como:

$$x = \frac{x^2 + e^x - 2}{3} \Rightarrow f(x) = \frac{x^2 + e^x - 2}{3}$$

$$\Rightarrow f'(x) = \frac{1}{3}(2x + e^x)$$

- Verificar condiciones del Teorema: para cada región donde se encuentra la raíz, tomar un intervalo $[a, b]$ que la cubra y graficar f y su derivada para poder observar el cumplimiento o no de las condiciones.

- Si no se cumplen las condiciones, podemos probar con otra expresión para $f(x)$.

2.2.5. Interpretación gráfica

- Dado que el método plantea encontrar el valor de x que satisface $x = f(x)$, resolver la ecuación original es equivalente a resolver el sistema:

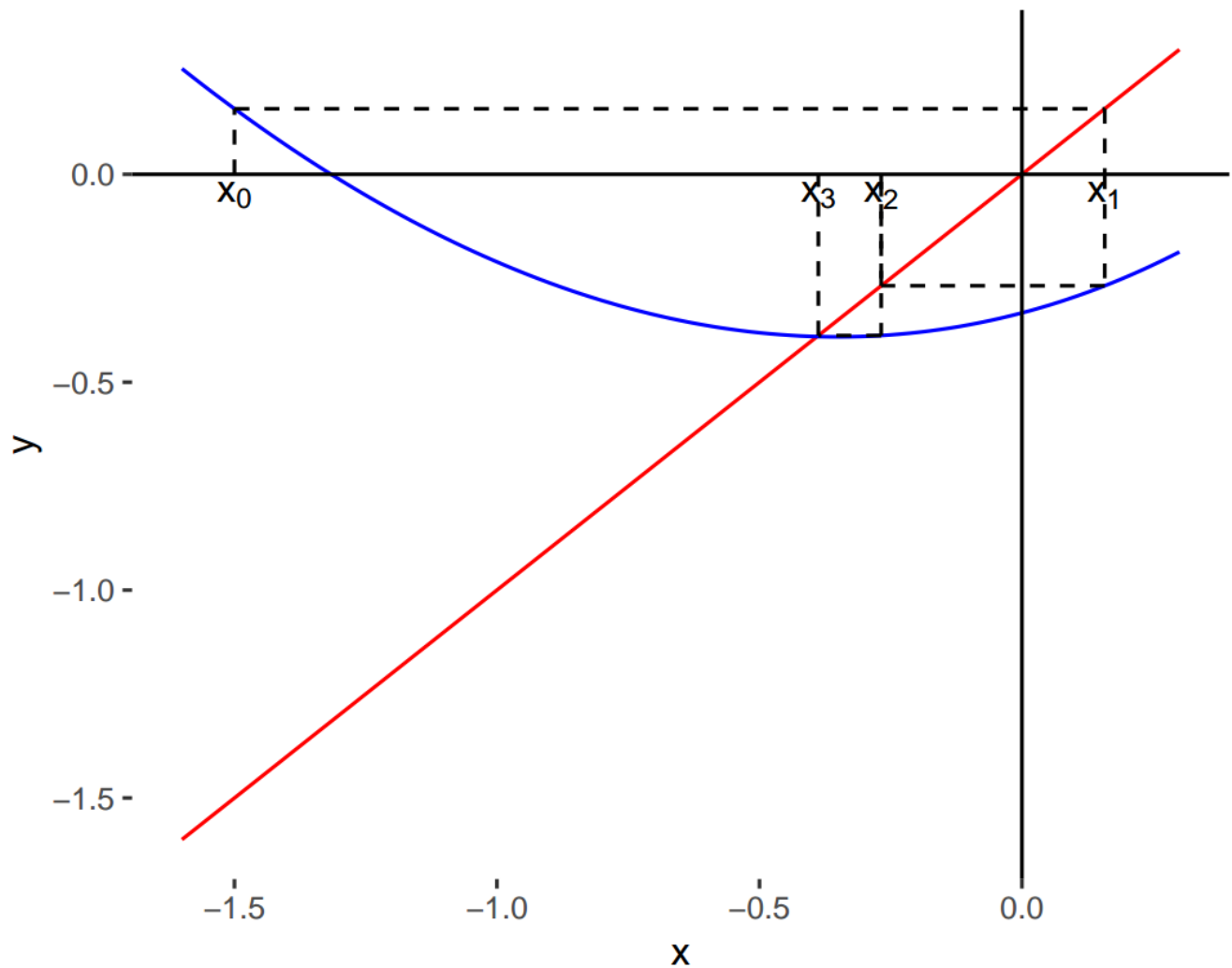
$$\begin{cases} y = f(x) \\ y = x \end{cases} \quad (2.2)$$

- Es decir, que geométricamente el valor buscado es el punto de intersección de la curva $y = f(x)$ con la recta $y = x$.

Ejemplo

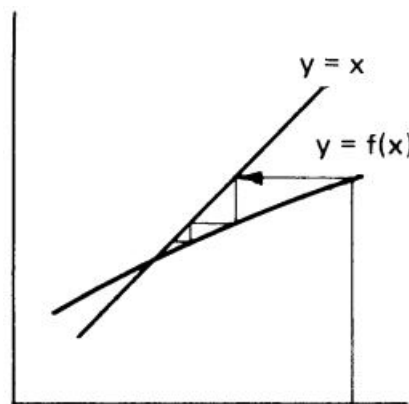
- Para $x_0 = -1.5$, el proceso converge en 7 iteraciones a la raíz -0.390271 con un error relativo menor a -1E+6.

Iteracion	x	Error
0	-1.500000	NA
1	0.1577101	1.105140e+00
2	-0.2681003	2.699957e+00
3	-0.3877637	4.463383e-01
4	-0.3903555	6.684034e-03
5	-0.3902688	2.222629e-04
6	-0.3902718	7.690434e-06
7	-0.3902717	2.657452e-07

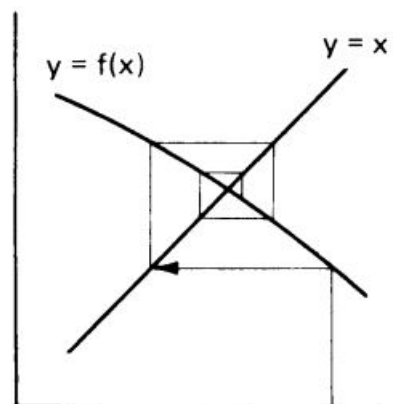


2.2.6. Algunos diagramas

Ejemplos de convergencia:

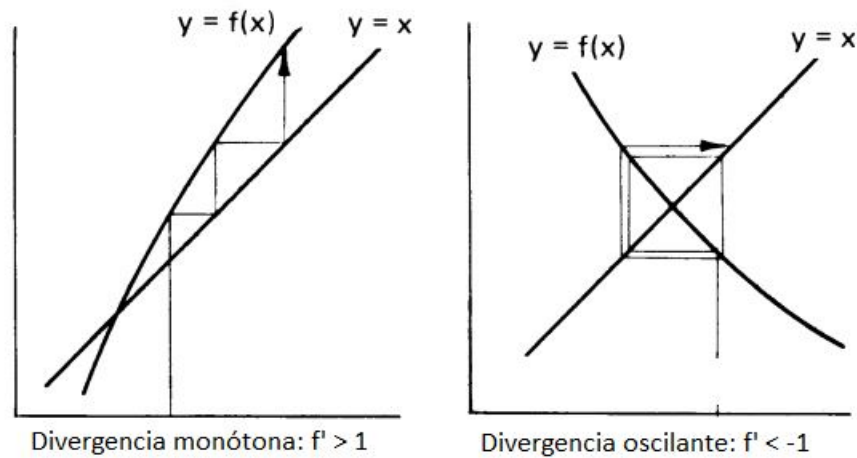


Convergencia monótona: $0 < f' < 1$



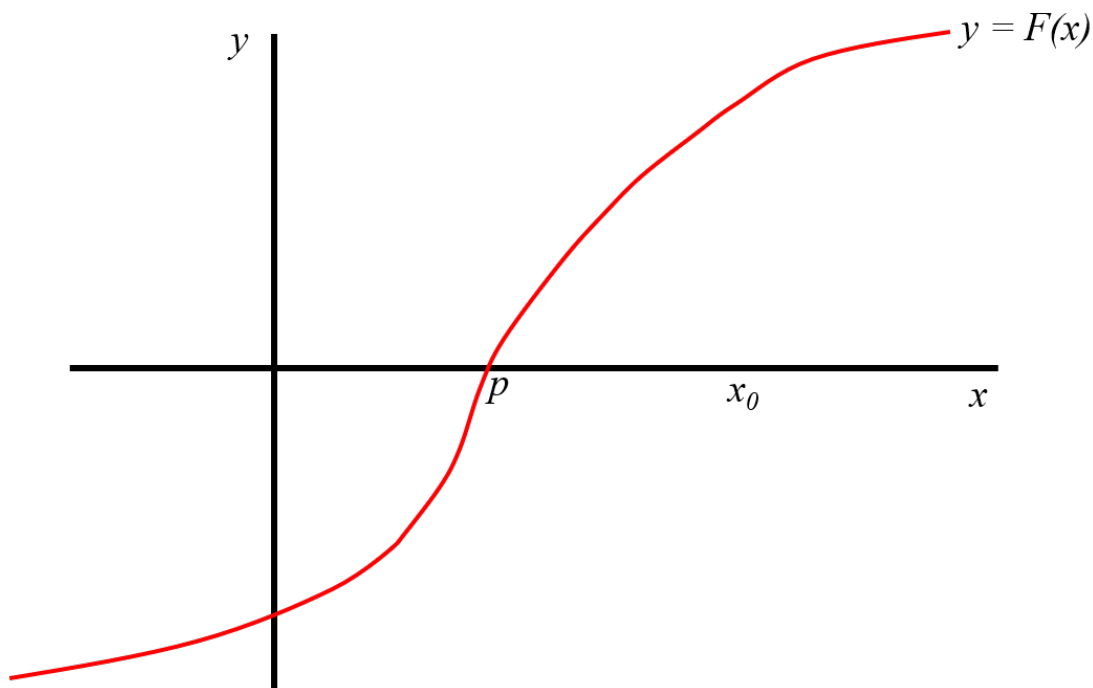
Convergencia oscilante: $-1 < f' < 0$

Ejemplos de divergencia:

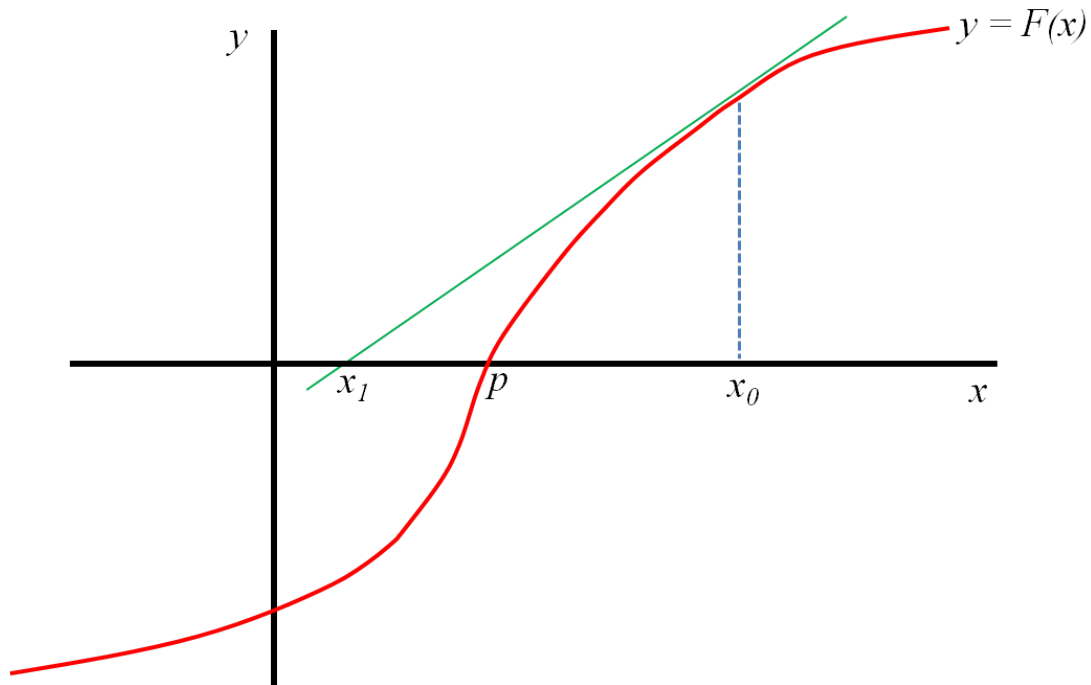


2.3. Método de Newton-Raphson

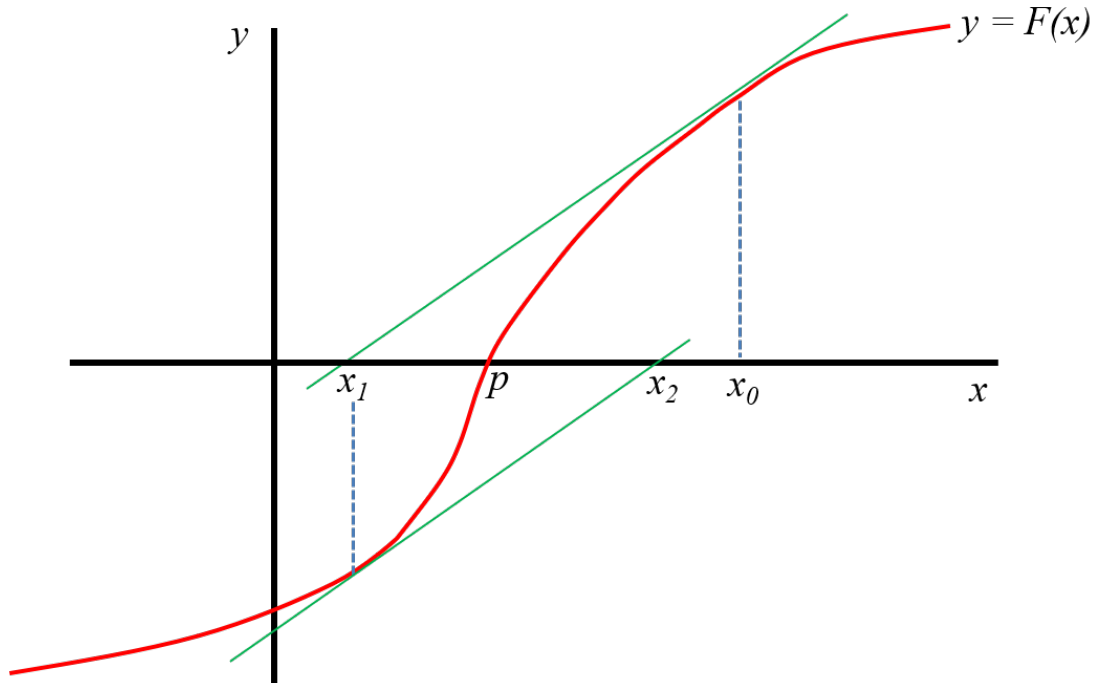
- Si la función F y sus derivadas F' y F'' son continuas cerca de una raíz p , se pueden usar estas características de F para desarrollar algoritmos que produzcan sucesiones $\{x_k\}$ que converjan a p más rápidamente.
- El método de **Newton-Raphson** es uno de los más útiles y conocidos.
- Vamos a introducir este método a partir de su interpretación geométrica y su representación gráfica.
- **Recordar:** La *tangente* a una curva en un punto es una recta que toca a la curva sólo en dicho punto.
- Veamos el siguiente ejemplo donde el objetivo es hallar la raíz de la función F , es decir, el valor p tal que $F(p) = 0$.
- (Sí, tenemos que hacer más lindos estos esquemas, ya los vamos a reemplazar por alguna animación...)
- Supongamos que contamos con una aproximación inicial x_0 cercana a la raíz p .



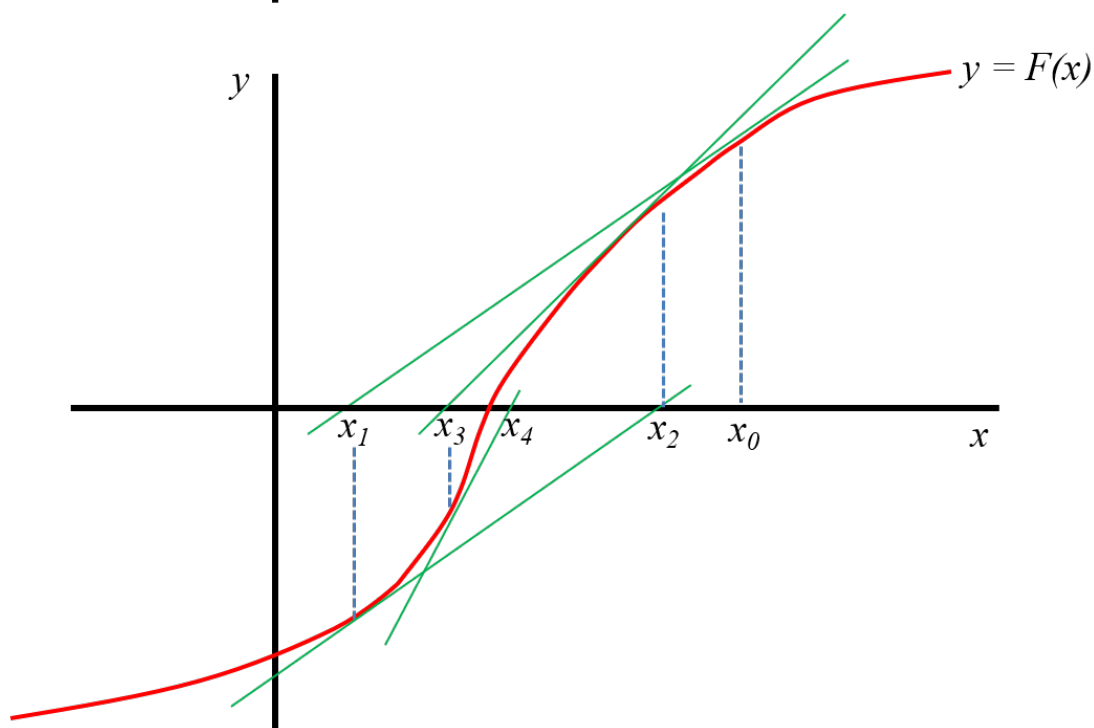
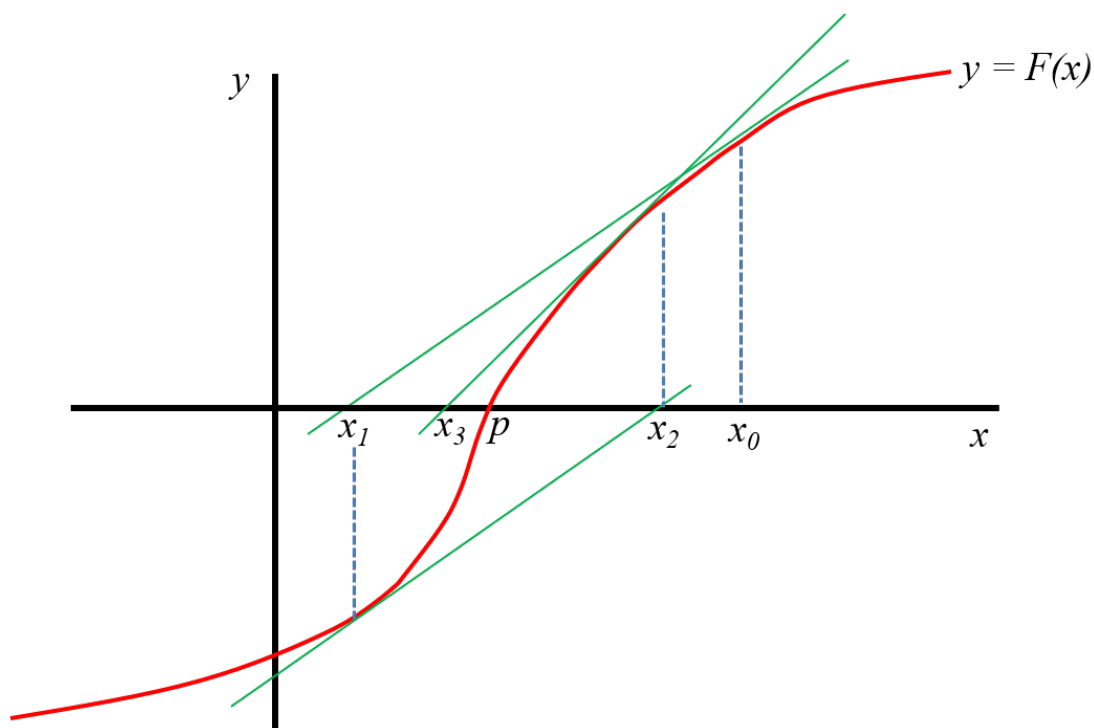
- Definimos a x_1 como el punto de intersección del eje de las abscisas con la recta tangente a la curva F en x_0 .
- En el caso que muestra la figura, se puede observar que x_1 está más cerca de p que x_0 .



- Ahora definimos a x_2 como el punto de intersección del eje de las abscisas con la recta tangente a la curva F en x_1 .



- Nuevamente, para el caso del ejemplo, podemos ver cómo x_2 está aún más cerca de p .
- Si continuamos repitiendo este proceso, esperamos encontrar un x_n que sea una buena aproximación para p .



- ¿Podemos expresar esto que observamos gráficamente a través de una fórmula?
- Es decir, a partir de x_0 , ¿podemos encontrar una fórmula para x_1 ?
- Sí, para eso hay prestarle atención a la pendiente m de la recta tangente en x_0 .
- Por un lado, sabemos que la pendiente de la recta tangente a la curva en un punto es igual a la derivada de la función en dicho punto:

$$m = F'(x_0) \quad (2.3)$$

- Pero además sabemos que para cualquier recta, la pendiente es igual a:

$$m = \frac{y_1 - y_0}{x_1 - x_0} \quad (2.4)$$

siendo (x_0, y_0) y (x_1, y_1) dos puntos distintos que pertenecen a la misma.

- Para expresar la pendiente de la recta tangente en x_0 , podemos tomar los puntos $(x_0, F(x_0))$ y $(x_1, 0)$ (el punto donde la tangente intersecta al eje x), de manera que a partir de la fórmula anterior:

$$m = \frac{0 - F(x_0)}{x_1 - x_0} = -\frac{F(x_0)}{x_1 - x_0} \quad (2.5)$$

- Igualando (2.3) y (2.5) y despejando x_1 nos queda:

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)} \quad (2.6)$$

- Si repetimos este pensamiento empezando desde x_1 con la recta tangente a F en el punto x_1 , vamos a encontrar que:

$$x_2 = x_1 - \frac{F(x_1)}{F'(x_1)} \quad (2.7)$$

- De esta manera hemos deducido una fórmula recursiva que nos permitirá hallar una aproximación para el verdadero valor de la raíz de F .
- Las ideas anteriores se formalizan analíticamente a través del siguiente teorema, en el cual se deduce la fórmula recursiva a partir del desarrollo en serie de Taylor de la función F .

2.3.1. Teorema de Newton-Raphson

Supongamos que la función F es continua, con derivada segunda continua en el intervalo $[a; b]$, y que existe un número $p \in [a; b]$ tal que $F(p) = 0$. Si $F'(p) \neq 0$, entonces existe $\delta > 0$ tal que la sucesión $\{x_k\}_{k=0}^{\infty}$ definida por el proceso iterativo

$$x_k = x_{k-1} - \frac{F(x_{k-1})}{F'(x_{k-1})} \quad k = 1, 2, \dots$$

converge a p cualquiera sea la aproximación inicial $x_0 \in [p - \delta; p + \delta]$

Convergencia

Observación: para garantizar la convergencia, δ debe ser elegido tal que:

$$\frac{|F(x)F''(x)|}{[F'(x)]^2} < 1 \quad \forall x \in [p - \delta, p + \delta]$$

Esto significa que:

- x_0 debe estar suficientemente cerca a la raíz de $F(x) = 0$.
- $F''(x)$ no debe ser excesivamente grande.
- $F'(x)$ no debe estar muy próxima a cero.

Ejemplo

- Evaluar si Newton-Raphson permite hallar la raíz positiva de $F(x) = x^2 - 3x + e^x - 2$, que no pudo ser hallada con Aproximaciones Sucesivas.

2.3.2. Ventajas y desventajas

Ventajas

- Aparece la expresión original F en lugar de tener que buscar una f .
- Converge más rápido que el método de las aproximaciones sucesivas.

- En algunos casos en que aproximaciones sucesivas diverge, N-R converge.
- Se puede adaptar para hallar raíces complejas.

Limitaciones

- Si x_0 está demasiado lejos de la raíz deseada, la sucesión $\{x_k\}$ puede converger a otra raíz (la pendiente $F'(x_0)$ es muy pequeña).
- Obtener la derivada primera de la función F puede ser difícil o imposible. En ese caso se podría aproximar $F'(x_{k-1})$ con:

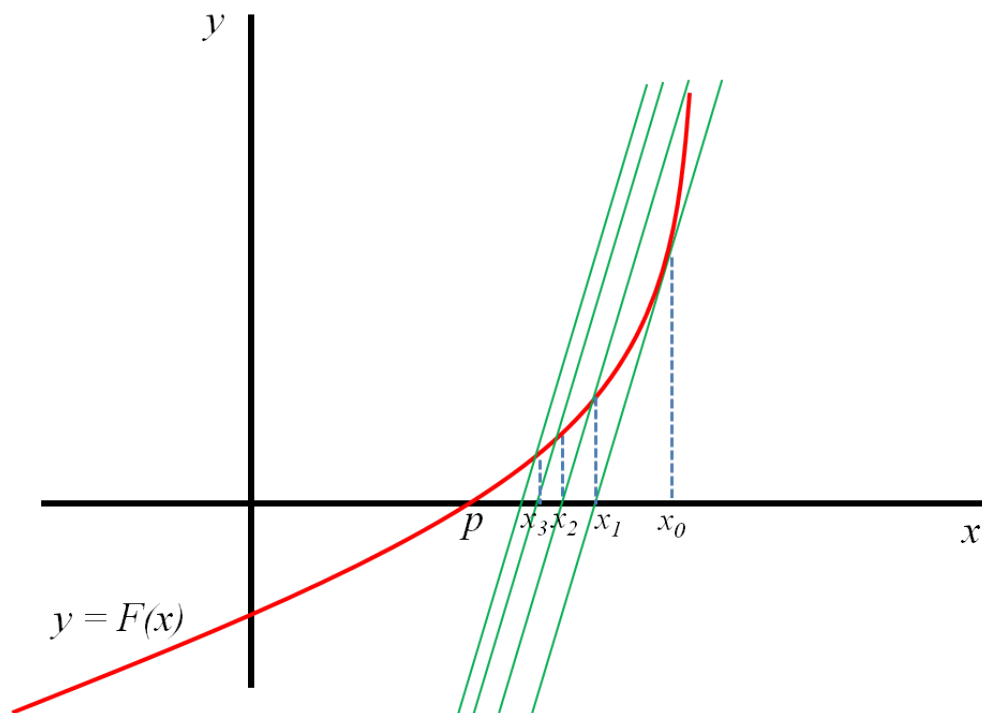
$$F'(x_{k-1}) \approx \frac{F(x_{k-1} + h) - F(x_{k-1})}{h}$$

donde h es un valor pequeño, por ejemplo, $h = 0,001$.

2.4. Método de von Mises

- En el método de N-R, el denominador $F'(x_k)$ hace que geométricamente se pase de una aproximación a la siguiente por la tangente de la curva $y = F(x)$ en el punto correspondiente a la aproximación presente x_k .
- Esto puede producir problemas cuando se esté en puntos alejados de raíces y cerca de puntos donde el valor de $F'(x)$ sea cercano a 0 (tangentes cercanas a la horizontal).
- Para resolver este problema, von Mises sugirió sustituir $F'(x_k)$ en el denominador por $F'(x_0)$.
- Es decir, obtener geométricamente las siguientes aproximaciones por medio de rectas paralelas siempre a la primera tangente.
- La fórmula de recurrencia resultante es:

$$x_k = x_{k-1} - \frac{F(x_{k-1})}{F'(x_0)} \quad k = 1, 2, \dots$$



2.5. Método de Newton-Raphson de 2º Orden

- Otra modificación al método de N-R se deriva a partir de la utilización de un término más en el desarrollo por serie de Taylor de la función $F(x)$.
- Dada la existencia de las correspondientes derivadas, la fórmula de recurrencia resultante es:

$$x_k = x_{k-1} + \frac{F(x_{k-1})F'(x_{k-1})}{0.5F(x_{k-1})F''(x_{k-1}) - [F'(x_{k-1})]^2} \quad k = 1, 2, \dots$$

- El método de N-R de 2º orden llega más rápidamente a la raíz, aunque la fórmula es más difícil de obtener.

Unidad 3

Solución de Sistemas de Ecuaciones Lineales

3.1. Generalidades

- **Objetivo:** examinar los aspectos numéricos que se presentan al resolver sistemas de ecuaciones lineales de la forma:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

- n ecuaciones, n incógnitas: sistema de orden $n \times n$.
- Los coeficientes a_{ij} y los términos independientes b_i son reales fijos.

3.2. Repaso

- Representación matricial: $\mathbf{Ax} = \mathbf{b}$, de dimensión $n \times n$, $n \times 1$ y $n \times 1$, respectivamente.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- Llamamos **matriz ampliada** o **aumentada** a:

$$[\mathbf{A} \quad \mathbf{b}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix}$$

- Un sistema de ecuaciones lineales se clasifica en:
 - **Compatible determinado:** tiene una única solución
 - **Compatible indeterminado:** tiene infinitas soluciones
 - **Incompatible:** no existe solución
- **Teorema.** Las siguientes condiciones son equivalentes:
 - El sistema $\mathbf{Ax} = \mathbf{b}$ tiene solución única (es compatible determinado).
 - La matriz \mathbf{A} es invertible (existe \mathbf{A}^{-1}).
 - La matriz \mathbf{A} es no singular ($\det \mathbf{A} \neq 0$).
 - El sistema $\mathbf{Ax} = \mathbf{0}$ tiene como única solución $\mathbf{x} = \mathbf{0}$.

- Dos sistemas de orden $n \times n$ son **equivalentes** si tienen el mismo conjunto de soluciones.
- Existen ciertas transformaciones sobre las ecuaciones de un sistema que no cambian el conjunto de soluciones (producen un sistema equivalente):
 - **Intercambio:** el orden de las ecuaciones puede cambiarse.
 - **Escalado:** multiplicar una ecuación por una constante no nula.
 - **Sustitución:** una ecuación puede ser reemplazada por una combinación lineal de las ecuaciones del sistema (*teorema fundamental de la equivalencia*)
- Realizar estas transformaciones sobre las ecuaciones es equivalente a realizar las mismas operaciones sobre las filas de la matriz aumentada.

3.3. Notación

- En esta sección presentamos la notación utilizaremos para expresar algoritmos con operaciones matriciales, facilitando su traducción a IML o R.
- Dada una matriz \mathbf{Z} de dimensión $n \times m$, anotamos:
 - $z_{ij} = \mathbf{Z}[i, j]$: elemento en la fila i y columna j de la matriz \mathbf{Z}
 - $\mathbf{Z}[i,]$: vector fila de dimensión $1 \times m$ constituido por la i -ésima fila de la matriz \mathbf{Z}
 - $\mathbf{Z}[, j]$: vector columna $n \times 1$ constituido por la j -ésima columna de la matriz \mathbf{Z}
 - $\mathbf{Z}[i, k : l]$: matriz de dimensión $1 \times (l - k + 1)$ constituida con los elementos $z_{i,k}, z_{i,k+1}, \dots, z_{i,l}$ de la matriz \mathbf{Z} , $l \geq k$.
 - $\mathbf{Z}[c : d, k : l]$: matriz de dimensión $(d - c + 1) \times (l - k + 1)$ constituida por la submatriz que contiene las filas de \mathbf{Z} desde la c hasta la d y las columnas de \mathbf{Z} desde la k hasta la l , $d \geq c$, $l \geq k$.
- Dado un vector \mathbf{Z} de largo n , anotamos:
 - $\mathbf{Z}[i]$: elemento i -ésimo del vector \mathbf{Z}
 - $\mathbf{Z}[k : l]$: vector de largo $(l - k + 1)$ constituido con los elementos z_k, z_{k+1}, \dots, z_l del vector \mathbf{Z} , $l \geq k$.

3.4. Métodos de Resolución de Sistemas de Ecuaciones

- **Métodos exactos:** permiten obtener la solución del sistema de manera directa.
 - Método de Eliminación de Gauss
 - Método de Gauss-Jordan
- **Métodos aproximados:** utilizan algoritmos iterativos que calculan la solución del sistema por aproximaciones sucesivas.
 - Método de Jacobi
 - Método de Gauss-Seidel
- En muchas ocasiones los métodos aproximados permiten obtener un grado de exactitud superior al que se puede obtener empleando los denominados métodos exactos, debido fundamentalmente a los errores de truncamiento que se producen en el proceso.

3.5. Métodos exactos

En esta sección repasaremos los métodos que ya conocen y que han aplicado “a mano” muchas veces para resolver sistemas de ecuaciones, con la particularidad de que pensaremos cómo expresar sus algoritmos para poder programarlos en la computadora. Entre los aspectos que tendremos que considerar, se encuentra la necesidad de aplicar estrategias de pivoteo (para evitar un *pivote* igual a cero cuando se resuelve el sistema) y la posibilidad de realizar reordenamientos de la matriz de coeficientes para disminuir los errores producidos en los cálculos computacionales.

3.5.1. Sistemas fáciles de resolver

Caso 1: La matriz \mathbf{A} es diagonal

$$\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- El sistema se reduce a n ecuaciones simples y la solución es:

$$\mathbf{x} = \begin{bmatrix} b_1/a_{11} \\ b_2/a_{22} \\ \vdots \\ b_n/a_{nn} \end{bmatrix}$$

Caso 2: La matrix \mathbf{A} es triangular superior:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix} \times \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

- La solución de x_n es inmediata y a partir de ella se encuentran las restantes en orden inverso x_{n-1}, \dots, x_1 , aplicando el algoritmo de **sustitución regresiva**:

$$x_n = \frac{b_n}{a_{nn}} \text{ y } x_k = \frac{b_k - \sum_{j=k+1}^n a_{kj}x_j}{a_{kk}} \quad k = n-1, n-2, \dots, 1$$

- Empleando la notación matricial vista, la suma en el algoritmo puede reescribirse con productos matriciales:

$$\mathbf{x}[n] = \frac{\mathbf{b}[n]}{\mathbf{A}[n, n]} \quad \mathbf{x}[k] = \frac{\mathbf{b}[k] - \mathbf{A}[k, k+1:n] \times \mathbf{x}[k+1:n]}{\mathbf{A}[k, k]} \\ k = n-1, n-2, \dots, 1$$

- Iniciando el vector solución con ceros, $\mathbf{x} = [0 \ 0 \ \cdots \ 0]$, la expresión anterior se simplifica como:

$$\mathbf{x}[n] = \frac{\mathbf{b}[n]}{\mathbf{A}[n, n]} \quad \mathbf{x}[k] = \frac{\mathbf{b}[k] - \mathbf{A}[k, :] \times \mathbf{x}}{\mathbf{A}[k, k]} \quad k = n-1, n-2, \dots, 1$$

Tarea: ver el algoritmo en documento adjunto y programarlo.

Caso 3: La matrix \mathbf{A} es triangular inferior:

- Obtención de la solución análoga al caso anterior.

Observaciones

- En los casos anteriores asumimos $a_{kk} \neq 0 \quad \forall \quad k$. De lo contrario el sistema no tiene solución o tiene infinitas.
- Recordar que un sistema lineal $\mathbf{Ax} = \mathbf{b}$ tiene solución única si y sólo si $\det \mathbf{A} \neq 0$ y que si un elemento de la diagonal principal de una matriz triangular es cero, entonces $\det \mathbf{A} = 0$.

3.5.2. Eliminación gaussiana

- Es un método para resolver un sistema lineal general $\mathbf{Ax} = \mathbf{b}$ de n ecuaciones con n incógnitas.
- El objetivo es construir un sistema equivalente donde la matriz de coeficientes sea triangular superior para obtener las soluciones con el algoritmo de sustitución regresiva.
- El método consiste en ir eliminando incógnitas en las ecuaciones de manera sucesiva.
- Ejemplo: resolver el siguiente sistema

$$\begin{cases} x + 2y - z + 3t = -8 \\ 2x + 2z - t = 13 \\ -x + y + z - t = 8 \\ 3x + 3y - z + 2t = -1 \end{cases} \quad \mathbf{A} = \begin{bmatrix} 1 & 2 & -1 & 3 \\ 2 & 0 & 2 & -1 \\ -1 & 1 & 1 & -1 \\ 3 & 3 & -1 & 2 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -8 \\ 13 \\ 8 \\ -1 \end{bmatrix}$$

- Matriz aumentada:

$$[\mathbf{A} \quad \mathbf{b}] = \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 2 & 0 & 2 & -1 & 13 \\ -1 & 1 & 1 & -1 & 8 \\ 3 & 3 & -1 & 2 & -1 \end{array} \right]$$

- En el primer paso eliminamos la incógnita x en las ecuaciones 2, 3 y 4, buscando que queden ceros en toda la primera columna excepto en el elemento diagonal.
- La eliminación se hace empleando las transformaciones que producen un sistema equivalente.
- Observando con detenimiento, esto se puede lograr realizando los siguientes reemplazos:

$$\begin{aligned} \text{Fila 2} &= \text{Fila 2} - 2 \times \text{Fila 1} \\ \text{Fila 3} &= \text{Fila 3} - (-1) \times \text{Fila 1} \\ \text{Fila 4} &= \text{Fila 4} - 3 \times \text{Fila 1} \end{aligned}$$

- A la Fila 1 se le dice **fila pivote** y al elemento $a_{11} = 1$, **pivote**. A los valores 2, -1 y 3 que multiplican a la fila pivote en los reemplazos se les dice **multiplicadores**.
- De manera general, los multiplicadores se simbolizan y obtienen con:

$$m_{r1} = a_{r1}/a_{11}, \quad r = 2, 3, \dots, n$$

- De esta forma, los reemplazos realizados se generalizan como:

$$\text{Fila } r = \text{Fila } r - m_{r1} \times \text{Fila 1}$$

- El resultado de los reemplazos anteriores es:

$$\begin{array}{l} \text{pivote} \rightarrow \\ m_{21} = 2 \\ m_{31} = -1 \\ m_{41} = 3 \end{array} \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 2 & 0 & 2 & -1 & 13 \\ -1 & 1 & 1 & -1 & 8 \\ 3 & 3 & -1 & 2 & -1 \end{array} \right] \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 3 & 0 & 2 & 0 \\ 0 & -3 & 2 & -7 & 23 \end{array} \right]$$

- En el segundo paso, eliminamos la incógnita y en las ecuaciones 3 y 4. La fila pivote pasa a ser la segunda y el pivote es $a_{22} = -4$. Los multiplicadores son $m_{r2} = a_{r2}/a_{22}$, $r = 3, 4$, dando lugar a los reemplazos:

$$\begin{aligned} \text{Fila 3} &= \text{Fila 3} - (-3/4) \times \text{Fila 2} \\ \text{Fila 4} &= \text{Fila 4} - 3/4 \times \text{Fila 2} \end{aligned}$$

- El resultado es:

$$\begin{array}{l} \text{pivote} \rightarrow \\ m_{32} = -3/4 \\ m_{42} = 3/4 \end{array} \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 3 & 0 & 2 & 0 \\ 0 & -3 & 2 & -7 & 23 \end{array} \right] \Rightarrow$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 0 & -12 & 13 & -87 \\ 0 & 0 & 4 & 7 & -5 \end{array} \right]$$

- Finalmente, eliminamos la incógnita z en la última ecuación. La fila pivote es la 3ª y el pivote es $a_{33} = -12$. El multiplicador es $m_{43} = a_{43}/a_{33} = -4/12$.
- El reemplazo a realizar es:

$$\text{Fila 4} = \text{Fila 4} - (-4/12) \times \text{Fila 3}$$

- El resultado es:

$$\text{pivote} \rightarrow m_{43} = -4/12 \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 0 & -12 & 13 & -87 \\ 0 & 0 & 4 & 7 & -5 \end{array} \right] \Rightarrow$$

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 0 & -12 & 13 & -87 \\ 0 & 0 & 0 & -136 & 408 \end{array} \right]$$

- Hemos llegado a un sistema equivalente cuya matriz de coeficientes es triangular superior, en el que aplicamos el algoritmo de sustitución regresiva:

$$\begin{cases} x + 2y - z - 3t = -8 \\ -4y + 4z - 7t = 29 \\ -12z + 13t = -87 \\ -136t = 408 \end{cases} \Rightarrow \begin{cases} x = 1 \\ y = 2 \\ z = 4 \\ t = -3 \end{cases}$$

Tarea: revisar el algoritmo y la función ya programada.

Eliminación gaussiana con pivoteo trivial

- En el algoritmo anterior es necesario que los pivotes $a_{qq} \neq 0 \forall q$.
- Si en uno de los pasos encontramos un $a_{qq} = 0$, debemos intercambiar la q -ésima fila por una cualquiera de las siguientes, por ejemplo la fila k , en la que $a_{kq} \neq 0, k > q$.
- Esta estrategia para hallar un pivote no nulo se llama **pivoteo trivial**.
- **Tarea:** verificar “a mano” que con el siguiente sistema, si seguimos el algoritmo anterior, incurrimos en este problema. En cambio, si seguimos el algoritmo con pivoteo trivial (ver en documento adjunto y su programa) no tenemos dicho problema.

$$\begin{cases} x - 2y + z = -4 \\ -2x + 4y - 3z = 3 \\ x - 3y - 4z = -1 \end{cases}$$

3.5.3. Estrategias de pivoteo para reducir los errores

- Como ya sabemos, dado que las computadoras usan una aritmética cuya precisión está fijada de antemano, es posible que cada vez que se realice una operación aritmética se introduzca un pequeño error.
- En la resolución de ecuaciones por eliminación gaussiana, un adecuado reordenamiento de las filas en el momento indicado puede reducir notablemente los errores cometidos.
- Por ejemplo, se puede mostrar cómo buscar un multiplicador de menor magnitud (es decir, un pivote de mayor magnitud) mejora la precisión de los resultados.
- Por eso, existen estrategias de pivoteo que no solamente hacen intercambio de filas cuando se tiene un pivote nulo, si no también cuando alguna de las filas posteriores tiene un pivote de mayor valor absoluto.

Pivoteo parcial

- Para reducir la propagación de los errores de redondeo, antes de comenzar una nueva ronda de reemplazos con el pivote a_{qq} se evalúa si debajo en la misma columna hay algún elemento con mayor valor absoluto y en ese caso se intercambian las respectivas filas.
- Es decir, se busca si existe r tal que $|a_{rq}| > |a_{qq}|$, $r > q$ para luego intercambiar las filas q y r .
- Este proceso suele conservar las magnitudes relativas de los elementos de la matriz triangular superior en el mismo orden que las de los coeficientes de la matriz original.
- **Tarea:** leer el algoritmo del pivoteo parcial en el documento adjunto y programarlo.

Pivoteo parcial escalado

- Reduce aún más los efectos de la propagación de los errores.
- Se elige el elemento de la columna q -ésima, en o por debajo de la diagonal principal, que tiene mayor tamaño relativo con respecto al resto de los elementos de su fila.
- **Paso 1:** buscar el máximo valor absoluto en cada fila:

$$s_r = \max\{|a_{rq}|, |a_{r,q+1}|, \dots, |a_{rn}|\} \quad r = q, q+1, \dots, n$$

- **Paso 2:** elegir como fila pivote a la que tenga el mayor valor de $\frac{|a_{rq}|}{s_r}$, $r = q, q+1, \dots, n$.
- **Paso 3:** intercambiar la fila q con la fila hallada en el paso 2.
- **Tarea:** leer el algoritmo del pivoteo parcial en el documento adjunto y programarlo.

3.5.4. Método de eliminación de Gauss-Jordan

- Ya vimos que el método de Gauss transforma la matriz de coeficientes en una matriz triangular superior.
- El método de Gauss-Jordan continúa el proceso de transformación hasta obtener la matriz identidad.
- En el ejemplo de la sección 3.6, cuando aplicamos eliminación de Gauss habíamos llegado a la siguiente matriz triangular:

$$\left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & -4 & 4 & -7 & 29 \\ 0 & 0 & -12 & 13 & -87 \\ 0 & 0 & 0 & -136 & 408 \end{array} \right]$$

- Luego, con el método de sustitución regresiva, llegamos a encontrar $x = 1$, $y = 2$, $z = 4$ y $t = -3$.
- En lugar de aplicar sustitución regresiva, seguimos operando por fila hasta lograr una matriz diagonal:

$$\begin{array}{lcl} \text{Fila 2} & = & \text{Fila 2}/(-4) \\ \text{Fila 3} & = & \text{Fila 3}/(-12) \\ \text{Fila 4} & = & \text{Fila 4}/(-136) \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & -1 & 3 & -8 \\ 0 & 1 & -1 & 7/4 & -29/4 \\ 0 & 0 & 1 & -13/12 & 87/12 \\ 0 & 0 & 0 & 1 & -3 \end{array} \right]$$

$$\begin{array}{lcl} \text{Fila 1} & = & \text{Fila 1} - 3 \text{ Fila 4} \\ \text{Fila 2} & = & \text{Fila 2} - 7/4 \text{ Fila 4} \\ \text{Fila 3} & = & \text{Fila 3} + 13/12 \text{ Fila 4} \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & -2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{array} \right]$$

$$\begin{array}{lcl} \text{Fila 1} & = & \text{Fila 1} + \text{Fila 3} \\ \text{Fila 2} & = & \text{Fila 2} + \text{Fila 3} \end{array} \Rightarrow \left[\begin{array}{cccc|c} 1 & 2 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{array} \right]$$

$$\text{Fila 1} = \text{Fila 1} - 2 \text{ Fila 2} \Rightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & -3 \end{array} \right]$$

$$\Rightarrow \begin{cases} x = 1 \\ y = 2 \\ z = 4 \\ t = -3 \end{cases}$$

- La ventaja de este método es que permite también hallar fácilmente la matriz inversa de **A**.
- Para esto hay que concatenar a la derecha de la matriz aumentada una matriz identidad de orden n . Cuando en la submatriz izquierda se llega a la matriz identidad, en el centro habrá quedado el vector solución y a su derecha la matriz inversa.
- **Ejemplo.** Resolver el siguiente sistema y hallar la inversa de la matriz de coeficientes:

$$\begin{cases} x - y + z = -4 \\ 5x - 4y + 3z = -12 \\ 2x + y + z = 11 \end{cases}$$

$$\left[\begin{array}{ccc|ccc} 1 & -1 & 1 & -4 & 1 & 0 & 0 \\ 5 & -4 & 3 & -12 & 0 & 1 & 0 \\ 2 & 1 & 1 & 11 & 0 & 0 & 1 \end{array} \right]$$

$$\begin{aligned} F2 &= F2 - 5F1 \\ F3 &= F3 - 2F1 \\ F1 &= F1 + F2 \\ F3 &= F3 - 3F2 \\ F1 &= F1 + 1/5F3 \\ F2 &= F2 + 2/5F3 \\ F3 &= F3/5 \end{aligned}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 3 & -7/5 & 2/5 & 1/5 \\ 0 & 1 & 0 & 6 & 1/5 & -1/5 & 2/5 \\ 0 & 0 & 1 & -1 & 13/5 & -3/2 & 1/5 \end{array} \right]$$

- ¿Por qué este procedimiento nos devuelva la inversa de la matriz de coeficientes?
- Recordar que la inversa de \mathbf{A} es aquella matriz \mathbf{A}^{-1} que verifica $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$
- Entonces si queremos hallar la matriz \mathbf{A}^{-1} :

$$\mathbf{A}^{-1} = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}$$

podemos plantear el siguiente producto matricial:

$$\left[\begin{array}{ccc} 1 & -1 & 1 \\ 5 & -4 & 3 \\ 2 & 1 & 1 \end{array} \right] \times \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array} \right]$$

que da lugar a tres sistemas de ecuaciones con la misma matriz de coeficientes:

$$\Rightarrow \begin{bmatrix} x_{11} - x_{21} + x_{31} & x_{12} - x_{22} + x_{32} & x_{13} - x_{23} + x_{33} \\ 5x_{11} - 4x_{21} + 3x_{31} & 5x_{12} - 4x_{22} + 3x_{32} & 5x_{13} - 4x_{23} + 3x_{33} \\ 2x_{11} + x_{21} + x_{31} & 2x_{12} - x_{22} + x_{32} & 2x_{13} - x_{23} + x_{33} \end{bmatrix} =$$

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

$$\begin{cases} x_{11} - x_{21} + x_{31} = 1 \\ 5x_{11} - 4x_{21} + 3x_{31} = 0 \\ 2x_{11} + x_{21} + x_{31} = 0 \end{cases}$$

$$\begin{cases} x_{12} - x_{22} + x_{32} = 0 \\ 5x_{12} - 4x_{22} + 3x_{32} = 1 \\ 2x_{12} + x_{22} + x_{32} = 0 \end{cases}$$

$$\begin{cases} x_{13} - x_{23} + x_{33} = 0 \\ 5x_{13} - 4x_{23} + 3x_{33} = 0 \\ 2x_{13} + x_{23} + x_{33} = 1 \end{cases}$$

- Los resolvemos simultáneamente al agregarle a la matriz aumentada del sistema original, estos tres “nuevos vectores de términos independientes” que forman una matriz identidad de orden n . En cierto modo, estamos resolviendo cuatro sistemas de ecuaciones a la vez, el sistema original y los tres que permiten hallar cada columna de \mathbf{A}^{-1} .
- **Tarea.** Revisar el algoritmo y programarlo.

3.6. Métodos Aproximados o Iterativos

- **Objetivo:** extender a espacios de dimensión mayor que 1 las ideas de los métodos iterativos vistos en la Unidad 2 para resolver sistemas de ecuaciones lineales.
- Veremos:
 - Método de Jacobi
 - Método de Gauss-Seidel

3.6.1. Método de Jacobi

- Consideremos el sistema:

$$\begin{cases} 4x - y + z = 7 \\ 4x - 8y + z = -21 \\ -2x + y + 5z = 15 \end{cases} \implies \begin{cases} x = (7 + y - z)/4 \\ y = (21 + 4x + z)/8 \\ z = (15 + 2x - y)/5 \end{cases}$$

- Despejar una incógnita en cada ecuación provee expresiones que sugieren la idea de un proceso iterativo.
- Dado un vector de valores iniciales $\mathbf{x}_0 = (x_0, y_0, z_0)$, operar con la siguiente fórmula de recurrencia hasta la convergencia:

$$\begin{cases} x_{k+1} = (7 + y_k - z_k)/4 \\ y_{k+1} = (21 + 4x_k + z_k)/8 \\ z_{k+1} = (15 + 2x_k - y_k)/5 \end{cases}$$

- Por ejemplo, tomando el valor inicial $(1, 2, 2)$, el proceso converge hacia la solución exacta del sistema $(2, 4, 3)$.
- Usando 4 posiciones decimales con redondeo luego de la coma:

k	x_k	y_k	z_k
0	1	2	2
1	1.7500	3.3750	3.0000
2	1.8438	3.8750	3.0250
3	1.9625	3.9250	2.9625
4	1.9906	3.9766	3.0000
5	1.9942	3.9953	3.0009
6	1.9986	3.9972	2.9986
7	1.9997	3.9991	3.0000
8	1.9998	3.9999	3.0001
9	2.0000	3.9999	2.9999
10	2.0000	4.0000	3.0000

- *Observación:* no siempre este método converge. Es sensible al ordenamiento de las ecuaciones dentro del sistema.
- Ejemplo: tomamos el mismo sistema de antes pero intercambiamos las filas 1 y 3:

$$\begin{cases} 4x - y + z = 7 \\ 4x - 8y + z = -21 \\ -2x + y + 5z = 15 \end{cases} \implies \begin{cases} -2x + y + 5z = 15 \\ 4x - 8y + z = -21 \\ 4x - y + z = 7 \end{cases}$$

$$\Rightarrow \begin{cases} x = (-15 + y + 5z)/2 \\ y = (21 + 4x + z)/8 \\ z = 7 - 4x + y \end{cases} \Rightarrow \begin{cases} x_{k+1} = (-15 + y_k + 5z_k)/2 \\ y_{k+1} = (21 + 4x_k + z_k)/8 \\ z_{k+1} = 7 - 4x_k + y_k \end{cases}$$

- Tomando el mismo valor inicial $(1, 2, 2)$, esta vez el proceso diverge:

k	x_k	y_k	z_k
0	1	2	2
1	-1.5000	3.3750	5.0000
2	6.6875	2.5000	16.3750
3	34.6875	8.0156	-17.2500
4	-46.6172	17.8125	-123.7344
5	-307.9298	-36.1504	211.2813
6	502.6281	-124.9297	1202.5688
7	2936.4572	404.2602	-2128.4421
8	-5126.4752	1204.7983	-11334.5686
9	-27741.5224	-3977.4337	21717.6991
10	52298.0309	-11153.4238	106995.6559

3.6.2. Método de Gauss-Seidel

- Toma la misma idea que Jacobi, pero con una pequeña modificación para acelerar la convergencia. Retomando el ejemplo inicial:

$$\begin{cases} 4x - y + z = 7 \\ 4x - 8y + z = -21 \\ -2x + y + 5z = 15 \end{cases} \Rightarrow \begin{cases} x = (7 + y - z)/4 \\ y = (21 + 4x + z)/8 \\ z = (15 + 2x - y)/5 \end{cases}$$

Jacobi:

$$\begin{cases} x_{k+1} = (7 + y_k - z_k)/4 \\ y_{k+1} = (21 + 4x_k + z_k)/8 \\ z_{k+1} = (15 + 2x_k - y_k)/5 \end{cases}$$

Gauss-Seidel:

$$\begin{cases} x_{k+1} = (7 + y_k - z_k)/4 \\ y_{k+1} = (21 + 4x_{k+1} + z_k)/8 \\ z_{k+1} = (15 + 2x_{k+1} - y_{k+1})/5 \end{cases}$$

- La diferencia está en que apenas calcula un nuevo valor de las incógnitas, Gauss-Seidel lo usa inmediatamente en el cálculo de las restantes, en lugar de esperar a la próxima ronda.
- Tomando otra vez el valor inicial $(1, 2, 2)$ y operando con 4 posiciones decimales con redondeo luego de la coma:

Jacobi				Gauss-Seidel			
k	x_k	y_k	z_k	k	x_k	y_k	z_k
0	1	2	2	0	1	2	2
1	1.7500	3.3750	3.0000	1	1.7500	3.7500	2.9500
2	1.8438	3.8750	3.0250	2	1.9500	3.9688	2.9862
3	1.9625	3.9250	2.9625	3	1.9957	3.9961	2.9991
4	1.9906	3.9766	3.0000	4	1.9993	3.9995	2.9998
5	1.9942	3.9953	3.0009	5	1.9999	3.9999	3.0000
6	1.9986	3.9972	2.9986	6	2.0000	4.0000	3.0000

Jacobi		Gauss-Seidel	
7	1.9997	3.9991	3.0000
8	1.9998	3.9999	3.0001
9	2.0000	3.9999	2.9999
10	2.0000	4.0000	3.0000

3.6.3. Generalización

- Ahora que sabemos qué hace cada método y en qué se diferencian, vamos a generalizar las fórmulas recursivas.
- Pensemos en un sistema de n ecuaciones lineales con n incógnitas y despejamos una variable en cada ecuación:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{j1}x_1 + a_{j2}x_2 + \cdots + a_{jn}x_n = b_j \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

$$\Rightarrow \begin{cases} x_1 = \frac{b_1 - a_{12}x_2 - \cdots - a_{1n}x_n}{a_{11}} \\ x_2 = \frac{b_2 - a_{21}x_1 - a_{23}x_3 - \cdots - a_{2n}x_n}{a_{22}} \\ \vdots \\ x_j = \frac{b_j - a_{j1}x_1 - \cdots - a_{j(j-1)}x_{j-1} - a_{j(j+1)}x_{j+1} - \cdots - a_{jn}x_n}{a_{jj}} \\ \vdots \\ x_n = \frac{b_n - a_{n1}x_1 - \cdots - a_{n(n-1)}x_{n-1}}{a_{nn}} \end{cases}$$

Generalización del Método de Jacobi

Con los supra índices indicando el número de iteración, la fórmula de recurrencia es:

$$\begin{cases} x_1^{(k+1)} = \frac{b_1 - a_{12}x_2^{(k)} - \cdots - a_{1n}x_n^{(k)}}{a_{11}} \\ x_2^{(k+1)} = \frac{b_2 - a_{21}x_1^{(k)} - a_{23}x_3^{(k)} - \cdots - a_{2n}x_n^{(k)}}{a_{22}} \\ \vdots \\ x_j^{(k+1)} = \frac{b_j - a_{j1}x_1^{(k)} - \cdots - a_{j(j-1)}x_{j-1}^{(k)} - a_{j(j+1)}x_{j+1}^{(k)} - \cdots - a_{jn}x_n^{(k)}}{a_{jj}} \\ \vdots \\ x_n^{(k+1)} = \frac{b_n - a_{n1}x_1^{(k)} - \cdots - a_{n(n-1)}x_{n-1}^{(k)}}{a_{nn}} \end{cases}$$

$$\Rightarrow x_j^{(k+1)} = \frac{b_j - a_{j1}x_1^{(k)} - \cdots - a_{j(j-1)}x_{j-1}^{(k)} - a_{j(j+1)}x_{j+1}^{(k)} - \cdots - a_{jn}x_n^{(k)}}{a_{jj}}$$

$$k = 1, 2, \dots, n$$

Generalización del Método de Gauss-Seidel

Con los supra índices indicando el número de iteración, la fórmula de recurrencia es:

$$\begin{cases} x_1^{(k+1)} = \frac{b_1 - a_{12}x_2^{(k)} - \dots - a_{1n}x_n^{(k)}}{a_{11}} \\ x_2^{(k+1)} = \frac{b_2 - a_{21}x_1^{(k+1)} - a_{23}x_3^{(k)} - \dots - a_{2n}x_n^{(k)}}{a_{22}} \\ \vdots \\ x_j^{(k+1)} = \frac{b_j - a_{j1}x_1^{(k+1)} - \dots - a_{j(j-1)}x_{j-1}^{(k+1)} - a_{j(j+1)}x_{j+1}^{(k)} - \dots - a_{jn}x_n^{(k)}}{a_{jj}} \\ \vdots \\ x_n^{(k+1)} = \frac{b_n - a_{n1}x_1^{(k+1)} - \dots - a_{n(n-1)}x_{n-1}^{(k+1)}}{a_{nn}} \end{cases}$$

$$\Rightarrow x_j^{(k+1)} = \frac{b_j - a_{j1}x_1^{(k+1)} - \dots - a_{j(j-1)}x_{j-1}^{(k+1)} - a_{j(j+1)}x_{j+1}^{(k)} - \dots - a_{jn}x_n^{(k)}}{a_{jj}}$$

$$k = 1, 2, \dots, n$$

3.6.4. Expresión matricial

- Ahora vamos a ver que es posible escribir las fórmulas recursivas anteriores de manera matricial, de modo que sea más fácil programar estos métodos.
- **Tarea:** corroborar con un sistema pequeño la coincidencia entre los procesos iterativos mencionados en la sección anterior y la iteración con matrices como se muestra a continuación. Ver ambos algoritmos y programar el de Gauss-Seidel.

Expresión matricial para el Método de Jacobi

- Si descomponemos a la matriz \mathbf{A} como $\mathbf{A} = \mathbf{D} + \mathbf{R}$, donde \mathbf{D} es la matriz diagonal formada con la diagonal de \mathbf{A} y \mathbf{R} es igual a \mathbf{A} excepto en la diagonal donde posee todos ceros, tenemos:

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ (\mathbf{D} + \mathbf{R})\mathbf{x} &= \mathbf{b} \\ \mathbf{Dx} &= \mathbf{b} - \mathbf{Rx} \\ \mathbf{x} &= \mathbf{D}^{-1}(\mathbf{b} - \mathbf{Rx}) \end{aligned}$$

- Esto da lugar a la siguiente fórmula de recurrencia, que es equivalente a las vistas anteriormente y facilita la programación del método:

$$\mathbf{x}^{(k+1)} = \mathbf{D}^{-1}(\mathbf{b} - \mathbf{Rx}^{(k)})$$

- Requisito: ningún elemento en la diagonal de \mathbf{A} es cero.

Expresión matricial para el Método de Gauss-Seidel

- Si descomponemos a la matriz \mathbf{A} como $\mathbf{A} = \mathbf{L} + \mathbf{U}$, donde \mathbf{L} es una matriz triangular inferior (incluyendo la diagonal de \mathbf{A}) y \mathbf{U} es una matriz triangular superior (con ceros en la diagonal), tenemos:

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ (\mathbf{L} + \mathbf{U})\mathbf{x} &= \mathbf{b} \\ \mathbf{Lx} &= \mathbf{b} - \mathbf{Ux} \\ \mathbf{x} &= \mathbf{L}^{-1}(\mathbf{b} - \mathbf{Ux}) \end{aligned}$$

- Esto da lugar a la siguiente fórmula de recurrencia, que es equivalente a las vistas anteriormente, y facilita la programación del método:

$$\mathbf{x}^{(k+1)} = \mathbf{L}^{-1}(\mathbf{b} - \mathbf{U}\mathbf{x}^{(k)})$$

- Requisito: ningún elemento en la diagonal de \mathbf{A} es cero.

3.6.5. Convergencia

Definición

Se dice que una matriz \mathbf{A} de orden $n \times n$ es **diagonal estrictamente dominante** cuando cada elemento diagonal es mayor a la suma del resto de los elementos de su fila en valor absoluto:

$$|a_{kk}| > \sum_{\substack{j=0 \\ j \neq k}}^n |a_{kj}| \quad k = 1, 2, \dots, n$$

Teorema

Si la matriz \mathbf{A} es diagonal estrictamente dominante, entonces el sistema lineal $\mathbf{Ax} = \mathbf{b}$ tiene solución única y los procesos iterativos de Jacobi y de Gauss-Seidel convergen hacia la misma solución cualquiera sea el vector de partida \mathbf{x}_0 .

- Es una condición suficiente pero no necesaria.
- En la práctica, ante un sistema buscamos reordenar las columnas y filas para intentar obtener una matriz con estas características. Si no es posible, igualmente buscamos colocar en la diagonal los elementos de mayor valor absoluto, puesto que esto puede favorecer a la convergencia.

Criterios para la convergencia

- **Norma L1 (distancia Manhattan):**

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_1 = \sum_{j=1}^n |x_j^{(k+1)} - x_j^{(k)}| < \epsilon$$

- **Norma L2 (norma euclídea):**

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_2 = \sqrt{(\mathbf{x}_{k+1} - \mathbf{x}_k)'(\mathbf{x}_{k+1} - \mathbf{x}_k)} = \sqrt{\sum_{j=1}^n (x_j^{(k+1)} - x_j^{(k)})^2} < \epsilon$$

- **Norma L_∞ (máxima diferencia):**

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\|_\infty = \max_j |x_j^{(k+1)} - x_j^{(k)}| < \epsilon$$

- Establecer un número máximo de iteraciones

Ventajas de Gauss-Seidel

- Converge más rápidamente.
- Puede converger cuando Jacobi no lo hace (aunque \mathbf{A} no sea diagonal estrictamente dominante, por ejemplo si \mathbf{A} es semidefinida positiva).
- Pero puede haber casos en los que Jacobi converge y Gauss-Seidel no.

Unidad 4

Aproximación Polinomial: interpolación y extrapolación

4.1. Generalidades

- En matemática estudiamos funciones de la forma $y = f(x)$, donde se conoce la expresión matemática que define a $f(x)$, para determinar sus derivadas y calcular integrales definidas.
- Sin embargo, muchas veces en aplicaciones prácticas se requiere trabajar con funciones cuya derivación e integración presenta dificultades porque:
 - no son funciones elementales (polinomios, expresiones racionales, funciones trigonométricas, exponenciales o combinaciones sencillas de estas);
 - directamente se desconoce su expresión analítica (ejemplo: datos obtenidos experimentalmente).
- En estos casos, se suele contar con una tabla de valores compuestas por puntos (x_i, y_i) a partir de la cual se desea:
 - aproximar $f(x)$ en abscisas que no están tabuladas (valores de x no dados), es decir, interpolar
 - integrar $f(x)$ en un intervalo determinado
 - derivar $f(x)$
- En esta unidad estudiaremos métodos numéricos que cumplen con estos objetivos al sustituir la función complicada o que está determinada tabularmente, por una función polinomial que se aproxime a los puntos disponibles.

4.2. Diferencias finitas

- Antes de ver el primer método veremos cómo crear una tabla de diferencias finitas que utilizaremos más adelante.
- Dada una función $y = f(x)$ definida en forma tabular con saltos equiespaciados en x :

x_k	$y_k = f(x_k)$
x_0	y_0
$x_1 = x_0 + h$	y_1
$x_2 = x_0 + 2h$	y_2
\vdots	\vdots
$x_{n-1} = x_0 + (n-1)h$	y_{n-1}
$x_n = x_0 + nh$	y_n

- Las diferencias finitas se expresan como:

$$\Delta^j y_k = \Delta^{j-1} y_{k+1} - \Delta^{j-1} y_k$$

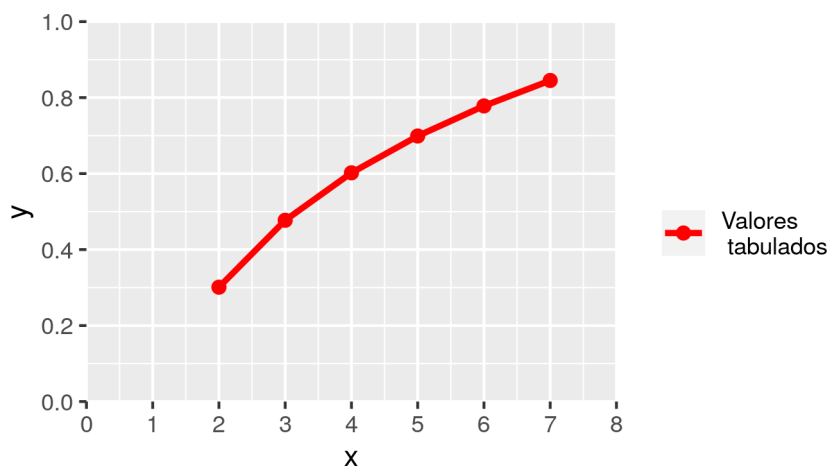
$$j = 1, \dots, n \quad k = 0, \dots, j-1 \quad \Delta^0 y_k = y_k$$

- Suelen ser presentadas en una *tabla de diferencias finitas*:

x_k	$y_k = f(x_k)$	$\Delta y_k = y_{k+1} - y_k$	$\Delta^2 y_k = \Delta y_{k+1} - \Delta y_k$	$\Delta^3 y_k = \Delta^2 y_{k+1} - \Delta^2 y_k$	\dots	$\Delta^{n-1} y_k = \Delta^{n-2} y_{k+1} - \Delta^{n-2} y_k$	$\Delta^n y_k = \Delta^{n-1} y_{k+1} - \Delta^{n-1} y_k$
x_0	y_0	$\Delta y_0 = y_1 - y_0$	$\Delta^2 y_0 = \Delta y_1 - \Delta y_0$	$\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0$	\dots	$\Delta^{n-1} y_0 = \Delta^{n-2} y_1 - \Delta^{n-2} y_0$	$\Delta^n y_0 = \Delta^{n-1} y_1 - \Delta^{n-1} y_0$
$x_1 = x_0 + h$	y_1	$\Delta y_1 = y_2 - y_1$	$\Delta^2 y_1 = \Delta y_2 - \Delta y_1$	$\Delta^3 y_1 = \Delta^2 y_2 - \Delta^2 y_1$	\dots	$\Delta^{n-1} y_1 = \Delta^{n-2} y_2 - \Delta^{n-2} y_1$	-
$x_2 = x_0 + 2h$	y_2	$\Delta y_2 = y_3 - y_2$	$\Delta^2 y_2 = \Delta y_3 - \Delta y_2$	$\Delta^3 y_2 = \Delta^2 y_3 - \Delta^2 y_2$	\dots	-	-
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
$x_{n-1} = x_0 + (n-1)h$	y_{n-1}	$\Delta y_{n-1} = y_n - y_{n-1}$	-	-	\dots	-	-
$x_n = x_0 + nh$	y_n	-	-	-	\dots	-	-

Ejemplo 1. Sea $y = f(x)$ una función desconocida de la cual se tienen los valores tabulados (x_k, y_k) que se presentan a continuación, junto con una representación gráfica de los mismos:

k	x_k	y_k
0	2	0,3010
1	3	0,4771
2	4	0,6021
3	5	0,6990
4	6	0,7781
5	7	0,8451



La correspondiente tabla de diferencias finitas es:

k	x_k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$	$\Delta^5 y_k$
0	2	0,3010	0,1761	-0,0511	0,0230	-0,0127	0,0081
1	3	0,4771	0,1250	-0,0281	0,0103	-0,0046	-
2	4	0,6021	0,0969	-0,0178	0,0057	-	-
3	5	0,6990	0,0791	-0,0121	-	-	-
4	6	0,7781	0,0670	-	-	-	-
5	7	0,8451	-	-	-	-	-

Observaciones:

- Si tenemos $n + 1$ puntos podemos calcular n columnas de diferencias hacia adelante.
- Si la función $f(x)$ que dio lugar a la tabla es un polinomio de orden q , entonces la columna para la diferencia de orden q es constante y las siguientes columnas son todas nulas.
- Por lo tanto, si en el proceso de obtención de las diferencias sucesivas de una función, las diferencias de orden q se vuelven constantes (o aproximadamente constantes), sabemos que los datos provienen exactamente (o muy aproximadamente) de un polinomio de orden q .
- Errores de redondeo podrían hacer que a pesar de que los datos provengan de un polinomio, no encontremos

diferencias constantes.

4.3. Interpolación de Newton para incrementos constantes

- Ahora intentaremos resolver el problema de aproximar la función $f(x)$ para valores de x que no forman parte de la tabla.
- Operando con los valores de la tabla de diferencias, se comprueba que los valores tabulados y_k se pueden escribir como:

$$\begin{aligned} y_k &= \sum_{j=0}^k \binom{k}{j} \Delta^j y_0 \\ &= y_0 + k\Delta y_0 + \frac{k(k-1)}{2!} \Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!} \Delta^3 y_0 + \dots + \Delta^k y_0 \end{aligned}$$

donde $\Delta^0 y_0 = y_0$.

- Por ejemplo:

$$y_2 = 0.3010 + 2 \cdot 0.1761 + \frac{2 \cdot 1}{2!} (-0.0511) = 0.6021$$

- Esta fórmula se emplea para proporcionar valores aproximados de $y = f(x)$ para cualquier x :

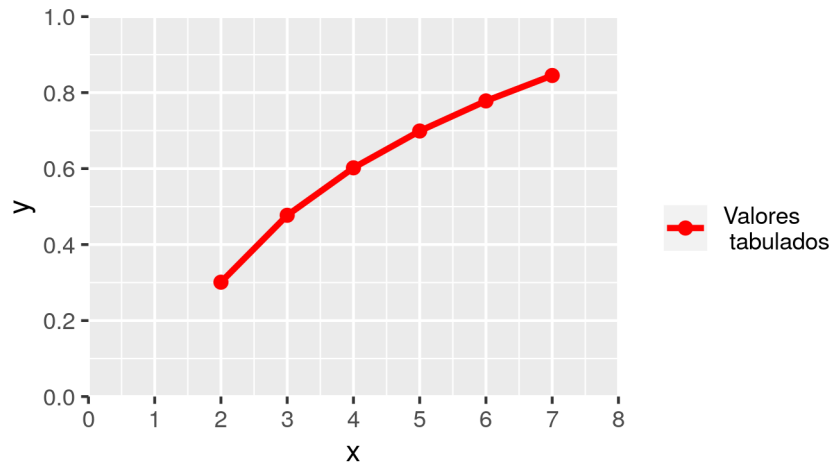
$$\begin{aligned} y = f(x) &\cong \sum_{j=0}^g \binom{k}{j} \Delta^j y_0 = \\ &y_0 + k\Delta y_0 + \frac{k(k-1)}{2!} \Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!} \Delta^3 y_0 + \dots \end{aligned}$$

donde g es el orden de la aproximación, $\Delta^0 y_0 = y_0$ y $k = \frac{x-x_0}{h}$.

- Si $x_0 < x < x_n$, este proceso se llama **interpolación** y la fórmula anterior, **Fórmula de interpolación de Newton para incrementos constantes**.
- Si para hallar una aproximación sólo empleamos la diferencia de 1º orden, tenemos una **interpolación lineal**: $f(x) \approx y_0 + k\Delta y_0$. El polinomio interpolador es de grado 1: una recta que pasa por los puntos (x_0, y_0) y (x_1, y_1) .
- Si empleamos hasta la diferencia de 2º orden, tenemos una **interpolación cuadrática**: $f(x) \approx y_0 + k\Delta y_0 + \frac{k(k-1)}{2!} \Delta^2 y_0$. El polinomio interpolador es de grado 2: una parábola que pasa por los puntos (x_0, y_0) , (x_1, y_1) y (x_2, y_2) .
- Cuantas más diferencias empleemos, el polinomio interpolador es de mayor orden y puede brindar mejores aproximaciones.
- Si empleamos las n diferencias, el polinomio interpolador es el único polinomio de grado n que pasa exactamente por los $n+1$ puntos tabulados.

Retomando el Ejemplo 1: vamos a aproximar el valor de $f(2, 3)$. Recordamos la tabla de diferencias y el gráfico:

x_k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$	$\Delta^5 y_k$
2	0,3010	0,1761	-0,0511	0,0230	-0,0127	0,0081
3	0,4771	0,1250	-0,0281	0,0103	-0,0046	-
4	0,6021	0,0969	-0,0178	0,0057	-	-
5	0,6990	0,0791	-0,0121	-	-	-
6	0,7781	0,0670	-	-	-	-
7	0,8451	-	-	-	-	-



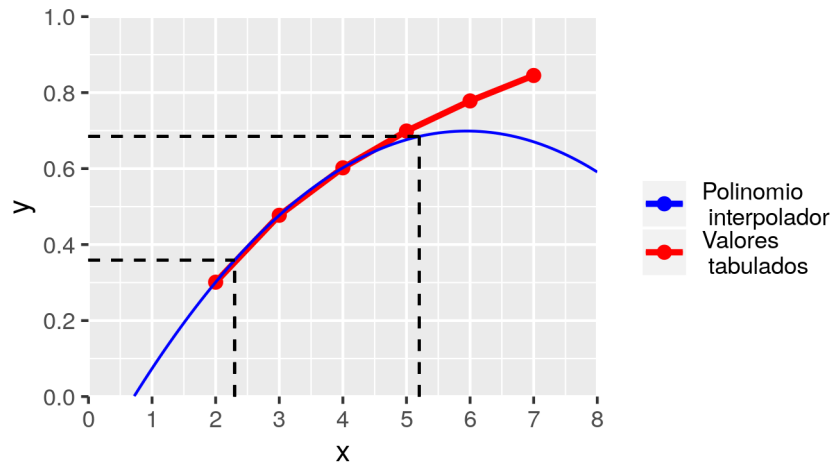
- ¿Hasta qué orden de diferencias deberíamos incluir? Es decir, ¿qué grado elegimos para el polinomio interpolador?
- Viendo el gráfico podemos pensar que una función cuadrática puede proveer un buen ajuste.
- Por eso, proponemos un polinomio interpolador de grado 2, usando hasta la diferencia de segundo orden:
 - $x = 2, 3$
 - $x_0 = 2$
 - $h = 1$
 - $k = \frac{x-x_0}{h} = \frac{2,3-2}{1} = 0,3$
 - $y_0 = 0,3010$; $\Delta y_0 = 0,1761$; $\Delta^2 y_0 = -0,0511$

$$\begin{aligned}
 y = f(2,3) &\approx y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 \\
 &= 0,3010 + k \cdot 0,1761 + \frac{k(k-1)}{2!} \cdot (-0,0511) \\
 &= 0,3010 + 0,3 \cdot 0,1761 + \frac{0,3(0,3-1)}{2}(-0,0511) \\
 &= 0,3592
 \end{aligned}$$

- ¿Podemos encontrar la expresión del polinomio interpolador que acabamos de usar?
- Sí, sólo debemos reemplazar k en la fórmula anterior por $k = \frac{x-x_0}{h} = \frac{x-2}{1} = x-2$:

$$\begin{aligned}
 y = f(x) &\approx 0,3010 + k \cdot 0,1761 + \frac{k(k-1)}{2!} \cdot (-0,0511) \\
 &= 0,3010 + (x-2) \cdot 0,1761 + \frac{(x-2)(x-2-1)}{2!} \cdot (-0,0511) \\
 &= -0,02555x^2 + 0,30385x - 0,2045
 \end{aligned}$$

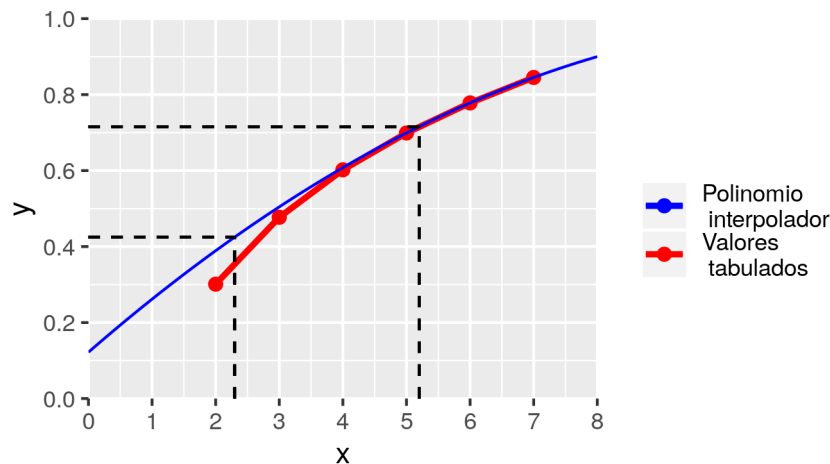
- Al representar gráficamente el polinomio interpolador, observamos que ajusta exactamente a los 3 primeros puntos tabulados y parece dar una aproximación razonable para $x = 2, 3$.



- Sin embargo, si queremos interpolar para $x = 5,2$, este polinomio no parece ser muy útil, porque 5,2 se aleja demasiado de los puntos ajustados.
- Por eso, siempre “corremos” el valor inicial x_0 hasta el inmediato inferior del que queremos aproximar.
- Entonces, para $x = 5,2$ tomamos:
 - $x_0 = 5$
 - $h = 1$
 - $k = \frac{x-x_0}{h} = \frac{5,2-5}{1} = 0,2$
 - $y_0 = 0,6990$; $\Delta y_0 = 0,0791$; $\Delta^2 y_0 = -0,0121$

$$\begin{aligned}
 y = f(5,2) &\approx y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 \\
 &= 0,6990 + 0,2 \cdot 0,0791 + \frac{0,2(0,2-1)}{2}(-0,0121) \\
 &= 0,7158
 \end{aligned}$$

- Así, con $x_0 = 5$, la expresión del polinomio interpolador queda $f(x) \approx -0,00605x^2 + 0,14565x - 0,122$ y ajusta exactamente a los tres últimos puntos tabulados:



- Retomemos la interpolación cuadrática para $x = 2,3$ que nos dio $f(x) \approx 0,3592$.
- ¿Cómo cambiará la aproximación si usamos polinomios de mayor orden, incluyendo diferencias superiores?
- Incrementar un grado en el orden del polinomio interpolador es muy sencillo, porque solamente tenemos que sumarle un término al polinomio de grado inferior.
- En el ejemplo con $x = 2,3$; $x_0 = 2$; $h = 1$; $k = 0,3$:

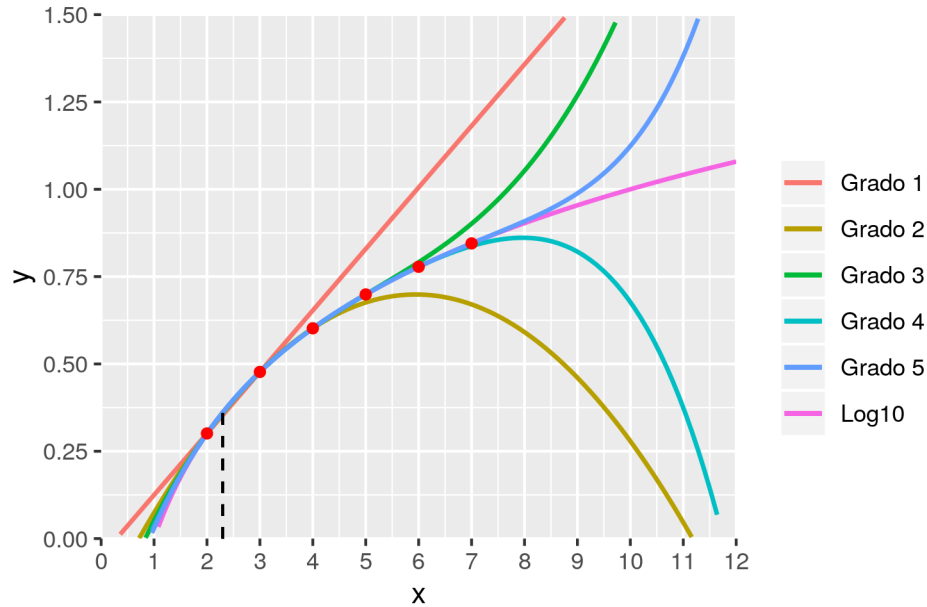
Grado	$f(2,3) \approx$
1	$y_0 + k\Delta y_0 = 0.3538$

Grado	$f(2.3) \approx$
2	$0.3538 + \frac{k(k-1)}{2!} \Delta^2 y_0 = 0.3592$
3	$0.3592 + \frac{k(k-1)(k-2)}{3!} \Delta^3 y_0 = 0.3606$
4	$0.3606 + \frac{k(k-1)(k-2)(k-3)}{4!} \Delta^4 y_0 = 0.3611$
5	$0.3611 + \frac{k(k-1)(k-2)(k-3)(k-4)}{5!} \Delta^5 y_0 = 0.3613$

- La función que generó la tabla de este ejemplo es el logaritmo en base 10, es decir, que el valor exacto es $f(2, 3) = \log(2, 3) = 0,3617$.
- Con esta información, podemos calcular el Error Relativo que resulta de hacer las aproximaciones anteriores:

Grado	$f(2.3) \approx$	ER
1	0.3538	0.0218
2	0.3592	0.0070
3	0.3606	0.0032
4	0.3611	0.0018
5	0.3613	0.0012

Polinomios interpoladores con $x_0 = 2$:



- Ahora vamos a realizar una interpolación cúbica para $x = 6.4$.
- Si tomamos $x_0 = 6$, nos faltan las diferencias de 2º y 3º orden y no podemos obtener el polinomio deseado.
- Modificación:** tomar como x_0 al inmediato superior y emplear las diferencias que se encuentran en la diagonal ascendente que comienza en dicho punto.
- Se debe redefinir k y alternar signos positivos y negativos en la fórmula:

x_k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$	$\Delta^5 y_k$
2	0,3010	0,1761	-0,0511	0,0230	-0,0127	0,0081
3	0,4771	0,1250	-0,0281	0,0103	-0,0046	-
4	0,6021	0,0969	-0,0178	0,0057	-	-
5	0,6990	0,0791	-0,0121	-	-	-
6	0,7781	0,0670	-	-	-	-
7	0,8451	-	-	-	-	-

$$x_0 = 7$$

$$k = \frac{x_0 - x}{h} = \frac{7 - 6.4}{1} = 0.6$$

$$y_0 = 0,8451; \Delta y_0 = 0,0670; \Delta^2 y_0 = -0.0121; \Delta^3 y_0 = 0.0057$$

$$\begin{aligned} y = f(6.4) &\approx y_0 - k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 - \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 \\ &= 0.8451 - 0,6 \cdot 0.0670 + \frac{0.6(0.6-1)}{2}(-0.0121) - \frac{0.6(0.6-1)(0.6-2)}{6}0.0057 \\ &= 0.8063 \end{aligned}$$

Ejemplo 2. Dada la siguiente función tabulada, interpolar para $x = 3.2$.

x	$y = f(x)$
0	2
2	8
4	62
6	212
8	506
10	992

La tabla de diferencias es:

x_k	$y_k = f(x_k)$	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$
0	2	6	48	48	0
2	8	54	96	48	0
4	62	150	144	48	-
6	212	294	192	-	-
8	506	486	-	-	-
10	992	-	-	-	-

- Las diferencias de orden 3 son constantes. Esto significa que la función $f(x)$ es exactamente un polinomio de grado 3.
- Es decir, si usamos hasta la diferencia de orden 3, con la fórmula de Newton podemos hallar exactamente el valor de $f(x)$ para cualquier x (y no importa cuál valor tomamos como x_0).

Continuando con el objetivo de hallar $f(3.2)$:

- $x = 3.2$
- $x_0 = 2$
- $h = 2$
- $k = \frac{x - x_0}{h} = \frac{3.2 - 2}{2} = 0.6$
- $y_0 = 8; \Delta y_0 = 54; \Delta^2 y_0 = 96; \Delta^3 y_0 = 48$

$$\begin{aligned} y = f(3.2) &= y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 \\ &= 8 + 0.6 \cdot 54 + \frac{0.6(0.6-1)}{2}96 + \frac{0.6(0.6-1)(0.6-2)}{6}48 \\ &= 31.568 \end{aligned}$$

- Ahora vamos a interpolar para $x = 9$.
- Si tomamos como x_0 al inmediato inferior, es decir, $x_0 = 8$, nos faltan dos diferencias para completar el cálculo.
- Pero como tenemos diferencias constantes y la verdadera función es un polinomio, cualquier punto de partida nos dará el valor exacto de $f(9)$.
- Así que podemos tomar, por ejemplo, el primer renglón de la tabla: $x_0 = 0$:
- $x = 9$
- $x_0 = 0$

- $h = 2$
- $k = \frac{x-x_0}{h} = \frac{9-0}{2} = 4.5$
- $y_0 = 2; \Delta y_0 = 6; \Delta^2 y_0 = 48; \Delta^3 y_0 = 48$

$$\begin{aligned}y = f(9) &= y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 \\&= 2 + 4.5 \cdot 6 + \frac{4.5(4.5-1)}{2}48 + \frac{4.5(4.5-1)(4.5-2)}{6}48 \\&= 722\end{aligned}$$

- En los ejemplos anteriores obtuvimos el valor de y para x comprendida entre dos valores x_k de la tabla (*interpolación*).
- Ahora vamos a calcular el valor de y para una x fuera del rango de los valores de x_k en la tabla.
- Esto se llama **extrapolación**.

Ejemplo 3. Extrapolar para $x = -1$.

- $x = -1$
- $x_0 = 0$
- $h = 2$
- $k = \frac{x-x_0}{h} = \frac{-1-0}{2} = -0.5$
- $y_0 = 2; \Delta y_0 = 6; \Delta^2 y_0 = 48; \Delta^3 y_0 = 48$

$$\begin{aligned}y = f(-1) &= y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 \\&= 2 - 0.5 \cdot 6 + \frac{-0.5(-0.5-1)}{2}48 + \frac{-0.5(-0.5-1)(-0.5-2)}{6}48 \\&= 2\end{aligned}$$

Ejemplo 4. Extrapolar para $x = 11$.

- $x = 11$
- $x_0 = 0$
- $h = 2$
- $k = \frac{x-x_0}{h} = \frac{11-0}{2} = 5.5$
- $y_0 = 2; \Delta y_0 = 6; \Delta^2 y_0 = 48; \Delta^3 y_0 = 48$

$$y = f(11) = \dots = 1322$$

4.4. Interpolación de Lagrange

- Como ya mencionamos, en la interpolación lineal se utiliza un segmento rectilíneo que pasa por dos puntos que se conocen.
- De nuestros conocimientos de Geometría sabemos que la ecuación de la recta que pasa por dos puntos (x_0, y_0) y (x_1, y_1) es:

$$y = f(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0)$$

- Es fácil ver que la recta anterior pasa por los puntos dados, ya que $f(x_0) = y_0$ y $f(x_1) = y_1$.
- Lagrange propuso reescribir la recta anterior como:

$$f(x) = \frac{x - x_1}{x_0 - x_1}y_0 + \frac{x - x_0}{x_1 - x_0}y_1$$

- Esta expresión se puede extender para obtener polinomios de grados superiores que pasen por más puntos.
- Por ejemplo, para hallar el polinomio que pasa por los puntos (x_0, y_0) , (x_1, y_1) y (x_2, y_2) :

$$f(x) = \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)}y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)}y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)}y_2$$

- Se puede ver fácilmente que este polinomio pasa exactamente por los tres puntos dados.
- De la misma forma, el polinomio de grado 3 que pasa por cuatro puntos (x_0, y_0) , (x_1, y_1) , (x_2, y_2) y (x_3, y_3) es:

$$f(x) = \frac{(x-x_1)(x-x_2)(x-x_3)}{(x_0-x_1)(x_0-x_2)(x_0-x_3)}y_0 + \frac{(x-x_0)(x-x_2)(x-x_3)}{(x_1-x_0)(x_1-x_2)(x_1-x_3)}y_1 \\ + \frac{(x-x_0)(x-x_1)(x-x_3)}{(x_2-x_0)(x_2-x_1)(x_2-x_3)}y_2 + \frac{(x-x_0)(x-x_1)(x-x_2)}{(x_3-x_0)(x_3-x_1)(x_3-x_2)}y_3$$

- Se puede ver fácilmente que este polinomio pasa exactamente por los cuatro puntos dados.
- Generalizando, la fórmula de interpolación de Lagrange para construir un polinomio de grado n que pase por $n+1$ puntos $(x_0, y_0), (x_1, y_1), \dots, (x_n, y_n)$ es:

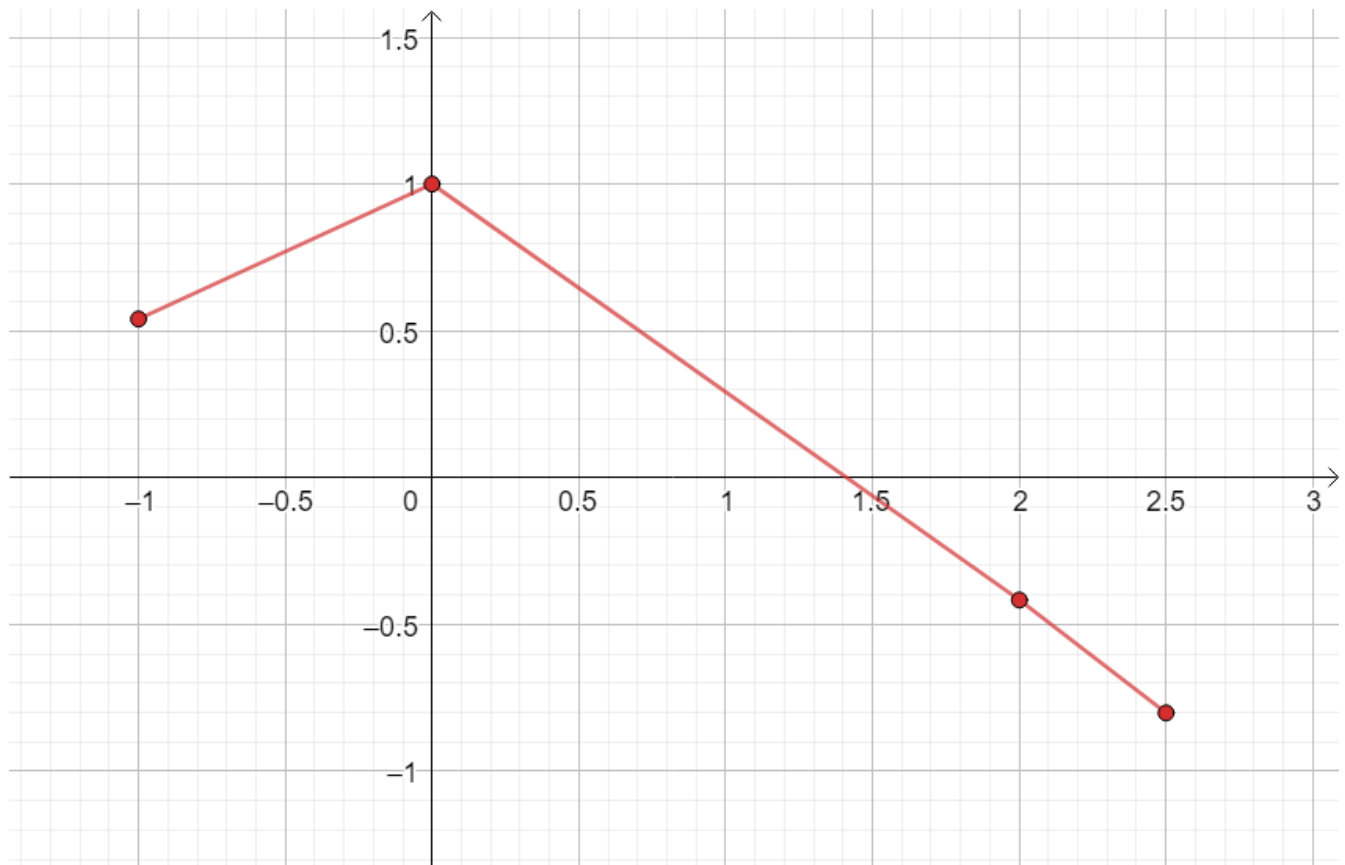
$$y = f(x) = \sum_{i=0}^n \frac{\prod_{\substack{j=0 \\ j \neq i}}^n (x - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^n (x_i - x_j)} y_i$$

Comparación con el método de Newton

- **Ventaja:** no requiere que los valores tabulados de x estén equiespaciados. Si lo están, ambos métodos coinciden.
- **Desventaja:**
 - En ocasiones es útil considerar varios polinomios interpoladores de distinto grado para luego elegir el más adecuado.
 - Con el método de Newton esto es muy sencillo, ya que como vimos se pueden construir recursivamente (por ejemplo, el de grado 3 es igual al de grado 2 más un término extra).
 - En cambio, con el método de Lagrange no hay relación entre la construcción de $P_{n-1}(x)$ y la de $P_n(x)$; cada polinomio debe construirse individualmente realizando todos los cálculos otra vez.
 - Esto implica que este método sea menos eficiente al tener que recalcular todo el polinomio si varía el número de puntos.

Ejemplo 5. Emplear el método de Lagrange para interpolar el valor de la función f en $x = 2.25$, con el polinomio interpolador de grado 3 que pasa por los cuatro puntos tabulados:

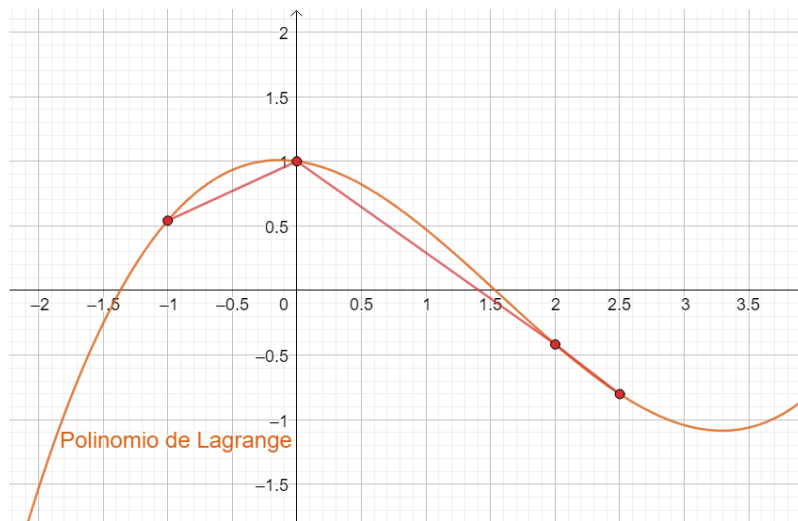
x_k	y_k
-1	0,5403
0	1,0000
2	-0,4162
2.5	-0,8011



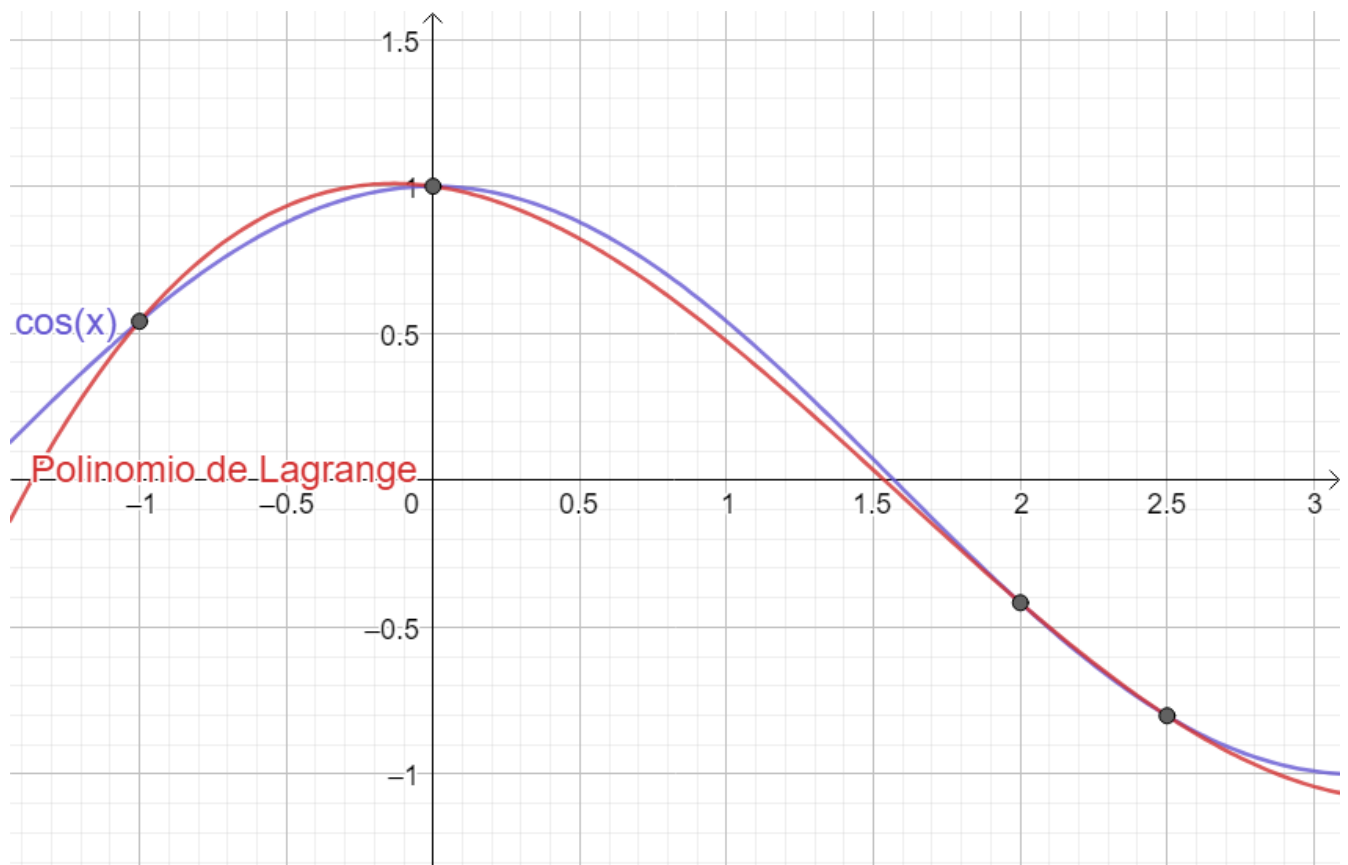
$$\begin{aligned}
 f(x) &= \sum_{i=0}^3 \frac{\prod_{\substack{j=0 \\ j \neq i}}^3 (x - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^3 (x_i - x_j)} y_i = \\
 &\frac{(x - x_1)(x - x_2)(x - x_3)}{(x_0 - x_1)(x_0 - x_2)(x_0 - x_3)} y_0 + \\
 &\frac{(x - x_0)(x - x_2)(x - x_3)}{(x_1 - x_0)(x_1 - x_2)(x_1 - x_3)} y_1 + \\
 &\frac{(x - x_0)(x - x_1)(x - x_3)}{(x_2 - x_0)(x_2 - x_1)(x_2 - x_3)} y_2 + \\
 &\frac{(x - x_0)(x - x_1)(x - x_2)}{(x_3 - x_0)(x_3 - x_1)(x_3 - x_2)} y_3 \\
 \Rightarrow y = f(2.25) &\approx \frac{(2.25 - 0)(2.25 - 2)(2.25 - 2.5)}{(-1 - 0)(-1 - 2)(-1 - 2.5)} 0.5403 + \\
 &\frac{(2.25 + 1)(2.25 - 2)(2.25 - 2.5)}{(0 + 1)(0 - 2)(0 - 2.5)} 1.0000 + \\
 &\frac{(2.25 + 1)(2.25 - 0)(2.25 - 2.5)}{(2 + 1)(2 - 0)(2 - 2.5)} (-0.4162) + \\
 &\frac{(2.25 + 1)(2.25 - 0)(2.25 - 2)}{(2.5 + 1)(2.5 - 0)(2.5 - 2)} (-0.8011) = \\
 &- 0.6218 \\
 \therefore f(2.25) &\approx -0.6218
 \end{aligned}$$

Si en la fórmula anterior en lugar de reemplazar x por un valor particular (en este caso, 2.5) operamos y reacomodamos los términos, podemos hallar la expresión del polinomio interpolador:

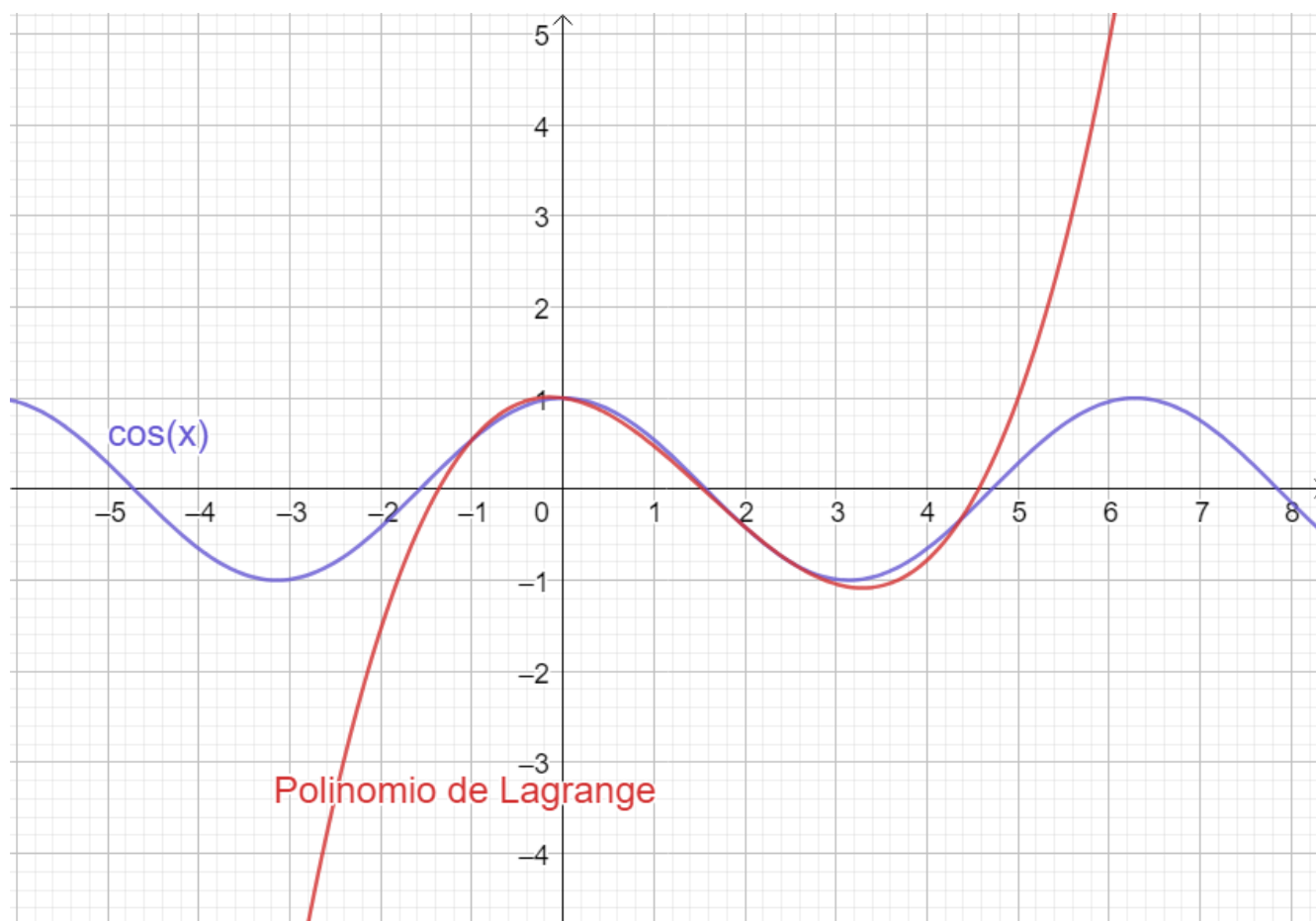
$$f(x) = 0.1042x^3 - 0.4934x^2 - 0.1379x + 1$$



La función que generó los valores tabulados es $\cos(x)$:
En el rango tabulado:



Un poco más lejos:



Siendo $\cos(2.25) = -0.6282$, el error relativo de la aproximación fue $|-0.6282 - 0.6218|/|-0.6285| = 1.02\%$

Interpolación inversa

- El problema de interpolación consiste en determinar el valor de la función desconocida $f(x)$ a partir de un valor conocido de x .
- En la *interpolación inversa* se resuelve el problema contrario: determinar el valor de x conociendo el de $f(x)$.
- Podemos implementar interpolación inversa con el método de Lagrange intercambiando el rol de las columnas tabuladas x e y , **pero sólo para rangos tabulados donde la función f es monótona**.

Ejemplo 6. Con los datos del Ejemplo 1, ¿cuál es el valor de x tal que $f(x) = -0.75$?

x_k	y_k
-1	0,5403
0	1,0000
2	-0,4162
2.5	-0,8011

↓

x_k	y_k
1,0000	0
-0,4162	2
-0,8011	2.5

$$f(x) = \sum_{i=0}^2 \frac{\prod_{\substack{j=0 \\ j \neq i}}^2 (x - x_j)}{\prod_{\substack{j=0 \\ j \neq i}}^2 (x_i - x_j)} y_i$$

$$= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} y_0 + \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} y_1 + \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} y_2$$

$$\Rightarrow f(-0.75) \approx \frac{(-0.75 + 0.4162)(-0.75 + 0.8011)}{(1 + 0.4162)(1 + 0.8011)} 0$$

$$+ \frac{(-0.75 - 1)(-0.75 + 0.8011)}{(-0.4162 - 1)(-0.4162 + 0.8011)} 2$$

$$+ \frac{(-0.75 - 1)(-0.75 + 0.4162)}{(-0.8011 - 1)(-0.8011 + 0.4162)} 2.5$$

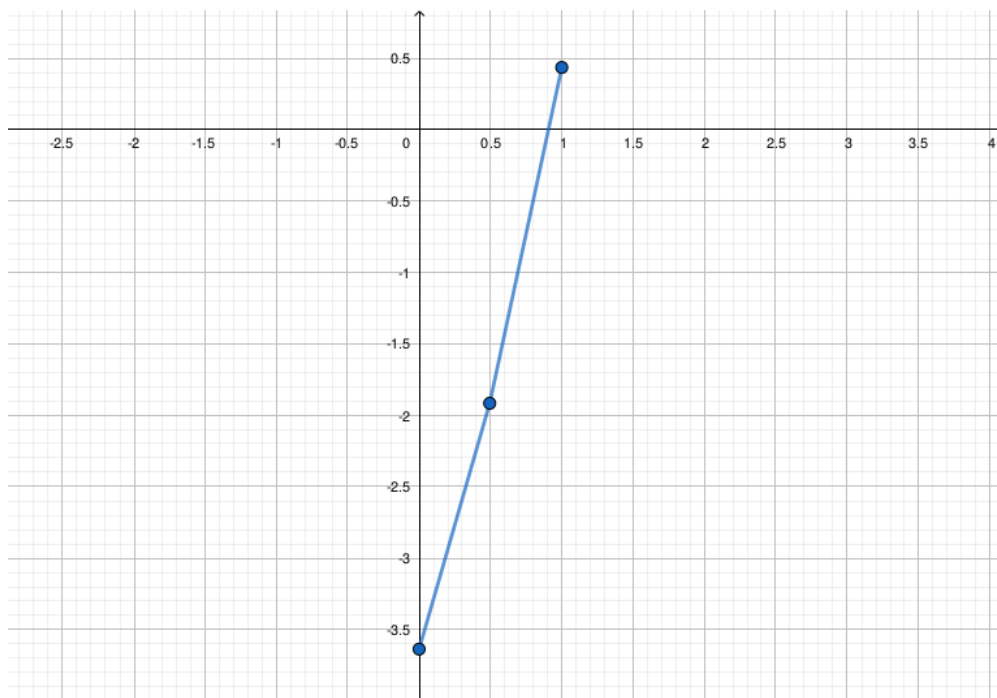
$$= 2.4347$$

\therefore el valor de x tal que $f(x) = -0.75$ es $x \approx 2.4347$.

Interpolación inversa. Podemos usar la interpolación inversa para resolver ecuaciones.

Ejemplo 7: hallar una raíz positiva para la ecuación $f(x) = x - 4 \sin(x + 2) = 0$. Tabulamos algunos valores y graficamos:

x_k	y_k
0	-3.6372
0.5	-1.8939
1	0.4355



Podemos usar los dos últimos puntos donde vemos que la función cortó a las abscisas y hacer una interpolación lineal, o usar los tres puntos y tener una interpolación cuadrática, o incluso agregar más puntos. Al intercambiar roles:

x_k	y_k
-3.6372	0
-1.8939	0.5
0.4355	1

$$y = f(x) = \frac{x - x_1}{x_0 - x_1} y_0 + \frac{x - x_0}{x_1 - x_0} y_1$$

$$\Rightarrow f(0) \approx \frac{0 - 0.4355}{-1.8939 - 0.4355} 0.5 + \frac{0 + 1.8939}{0.4355 + 1.8939} 1$$

$$= 0.9065$$

\therefore La solución positiva de la ecuación es $x \approx 0.9065$ ya que $f(0.9065) \approx 0$.

4.4.1. Observaciones finales

- Un polinomio de grado n ajustado a $n + 1$ puntos es único.
- El polinomio de interpolación se puede expresar en varias formas distintas, pero todas son equivalentes por el punto anterior.
- Si una función se aproxima mediante un polinomio de interpolación, no hay garantía de que dicho polinomio converja a la función exacta al aumentar el número de datos. En general, la interpolación mediante un polinomio de orden grande debe evitarse o utilizarse con precauciones extremas.
- Aunque no existe un criterio para determinar el orden óptimo del polinomio de interpolación, generalmente se recomienda utilizar uno con orden relativamente bajo en un pequeño rango de x .

Unidad 5

Aproximación Polinomial: integración y derivación numérica

5.1. Integración numérica

- Dada la función $y = f(x)$ definida en forma tabular con a través de $n + 1$ valores de x equiespaciados $x_0, x_1 = x_0 + h, \dots, x_n = x_0 + nh$, se desea hallar una aproximación de la integral definida:

$$\int_{x_0}^{x_n} f(x) dx \quad (5.1)$$

- Para esto, aproximaremos a $f(x)$ con el polinomio de Newton:

$$f(x) \cong y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 + \dots \quad (5.2)$$

$$k = \frac{x - x_0}{h}$$

- En (5.1) la variable es x , mientras que en (5.2) la variable está expresada como $k = (x - x_0)/h$, por lo tanto para poder reemplazar (5.2) en (5.1) se debe realizar un cambio de variables:

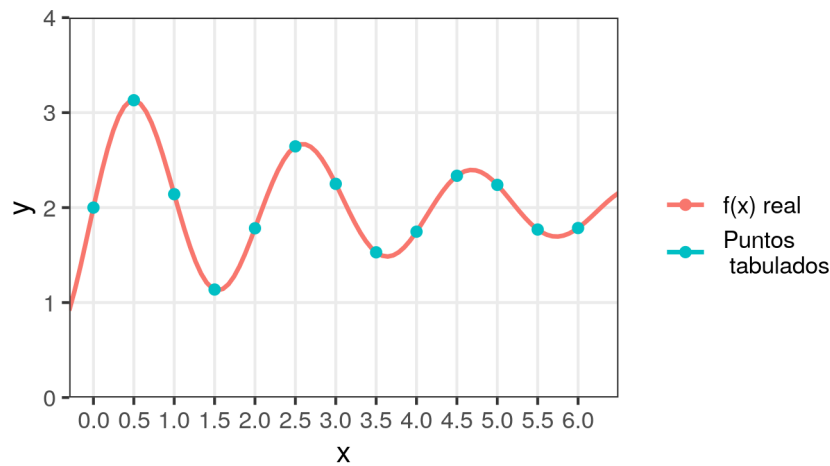
$$k = \frac{x - x_0}{h} \Rightarrow \begin{cases} x = x_0 + hk \\ dx = h dk \\ x = x_0 \Rightarrow k = \frac{x_0 - x_0}{h} = 0 \\ x = x_n \Rightarrow k = \frac{x_n - x_0}{h} = \frac{x_0 + nh - x_0}{h} = n \end{cases}$$

- Luego:

$$\begin{aligned} & \int_{x_0}^{x_n} f(x) dx \\ & \cong \int_0^n \left(y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 + \dots \right) h dk \\ & = h \int_0^n \left[y_0 + k\Delta y_0 + \left(\frac{k^2}{2} - \frac{k}{2} \right) \Delta^2 y_0 + \left(\frac{k^3}{6} - \frac{k^2}{2} + \frac{k}{3} \right) \Delta^3 y_0 + \dots \right] dk \\ & = h \left[y_0 k + \frac{k^2}{2} \Delta y_0 + \left(\frac{k^3}{6} - \frac{k^2}{4} \right) \Delta^2 y_0 + \left(\frac{k^4}{24} - \frac{k^3}{6} + \frac{k^2}{6} \right) \Delta^3 y_0 + \dots \right] \Big|_0^n \\ & = h \left[y_0 n + \frac{n^2}{2} \Delta y_0 + \left(\frac{n^3}{6} - \frac{n^2}{4} \right) \Delta^2 y_0 + \left(\frac{n^4}{24} - \frac{n^3}{6} + \frac{n^2}{6} \right) \Delta^3 y_0 + \dots \right] \end{aligned}$$

Ejemplo. Se tienen los siguientes valores tabulados de $f(x)$ y se desea hallar su integral entre 0 y 6.

x	$y = f(x)$
0.0	2.00
0.5	3.13
1.0	2.14
1.5	1.14
2.0	1.78
2.5	2.64
3.0	2.25
3.5	1.53
4.0	1.75
4.5	2.34
5.0	2.24
5.5	1.77
6.0	1.78



La curva roja es la verdadera función $f(x)$ que originó la tabla, la cual suponemos desconocida o difícil de integrar.

5.1.1. Fórmula trapezoidal

- Si la interpolación se limita al primer orden y la integral sólo se calcula entre los dos primeros valores de x , se obtiene:

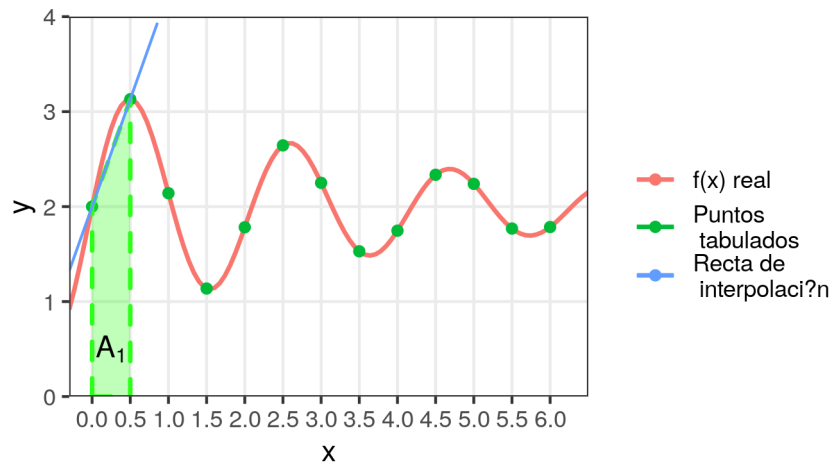
$$\begin{aligned} \int_{x_0}^{x_1} f(x) dx &\cong \int_0^1 (y_0 + k\Delta y_0) h dk = h \left(y_0 k + \frac{k^2}{2} \Delta y_0 \right) \Big|_0^1 \\ &= h \left(y_0 + \frac{\Delta y_0}{2} \right) = h \left(y_0 + \frac{y_1 - y_0}{2} \right) = \frac{h}{2} (y_0 + y_1) \end{aligned}$$

- En el ejemplo:

$$\int_0^{0.5} f(x) dx \cong \frac{0.5}{2} (3.13 + 2) = 1.2825$$

- Geométricamente, esto equivale al área A_1 del trapecio formado por la recta de interpolación y el eje de las abscisas, entre x_0 y x_1 :

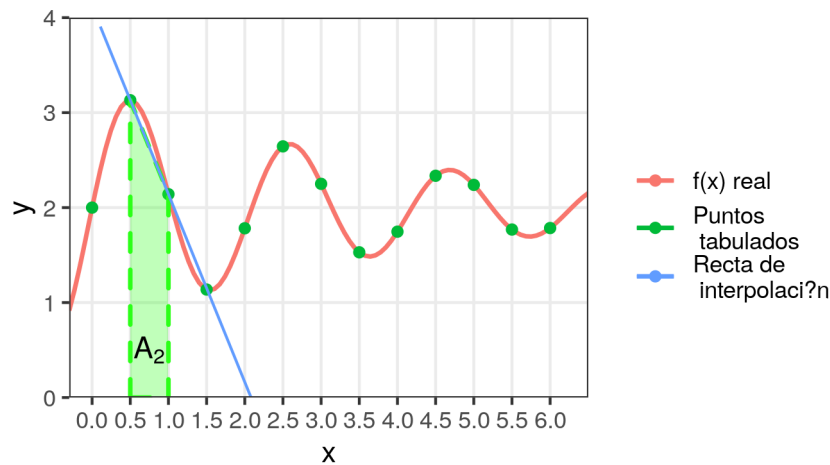
$$A_2 = 1.2825$$



- De manera semejante, se puede emplear la interpolación lineal de Newton para obtener una aproximación de la integral entre x_1 y x_2 :

$$\int_{x_1}^{x_2} f(x) dx \cong A_2 = \frac{h}{2}(y_1 + y_2)$$

$$A_2 = 1.3175$$



- Y sucesivamente para todos los intervalos:

$$\int_{x_{i-1}}^{x_i} f(x) dx \cong A_i = \frac{h}{2}(y_{i-1} + y_i) \quad i = 1, \dots, n$$

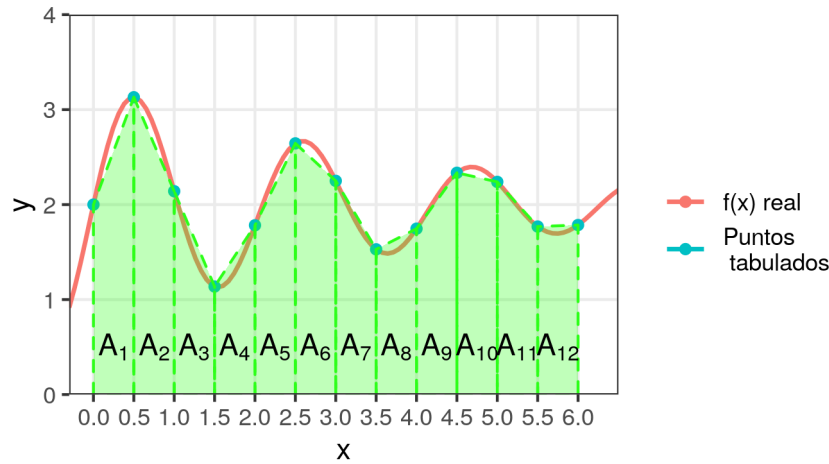
- De modo que la suma de las áreas de los trapecios A_i resulta ser la aproximación para la integral entre x_0 y x_n :

$$\int_{x_0}^{x_n} f(x) dx \cong \sum_{i=1}^n A_i = \sum_{i=1}^n \frac{h}{2}(y_{i-1} + y_i) = \frac{h}{2} \left(y_0 + y_n + 2 \sum_{i=1}^{n-1} y_i \right)$$

- La fórmula hallada se conoce como **fórmula trapecial** y se la simboliza con:

$$A_{1/2} = \frac{h}{2} \left(y_0 + y_n + 2 \sum_{i=1}^{n-1} y_i \right)$$

- Cuanto menor sea el ancho de los intervalos h y más se acerque $f(x)$ a una recta dentro de dichos intervalos, mejor será la aproximación así obtenida.
- Gráficamente:



- En el ejemplo: $A_{1/2} = 12.3000$.
- El valor exacto es: $\int_0^6 f(x)dx = 12.2935$, con lo cual el error relativo de la aproximación con la fórmula trapezoidal fue: 0.05%.

5.1.2. Fórmula de Simpson de 1/3

- Si la interpolación es de segundo orden y la integral sólo se calcula entre los tres primeros valores de x , se obtiene:

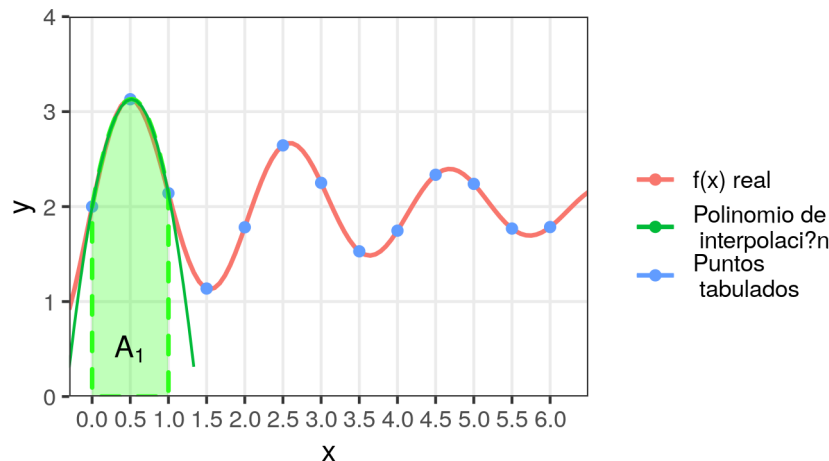
$$\begin{aligned} \int_{x_0}^{x_2} f(x)dx &\cong \int_0^2 \left[y_0 + k\Delta y_0 + \left(\frac{k^2}{2} - \frac{k}{2} \right) \Delta^2 y_0 \right] h dk \\ &= h \left[y_0 k + \frac{k^2}{2} \Delta y_0 + \left(\frac{k^3}{6} - \frac{k^2}{4} \right) \Delta^2 y_0 \right] \Big|_0^2 \\ &= h \left[2y_0 + 2\Delta y_0 + \frac{1}{3} \Delta^2 y_0 \right] \end{aligned}$$

- Dado que $\Delta y_0 = y_1 - y_0$ y $\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 + y_0$, nos queda:

$$\begin{aligned} \int_{x_0}^{x_2} f(x)dx &\cong h \left[2y_0 + 2(y_1 - y_0) + \frac{1}{3}(y_2 - 2y_1 + y_0) \right] \\ &= \frac{h}{3}(y_0 + 4y_1 + y_2) \end{aligned}$$

- Geométricamente, esto equivale al área A_1 encerrada entre el eje de las abscisas, x_0 y x_2 y el polinomio integrador que pasa por (x_0, y_0) , (x_1, y_1) y (x_2, y_2) :

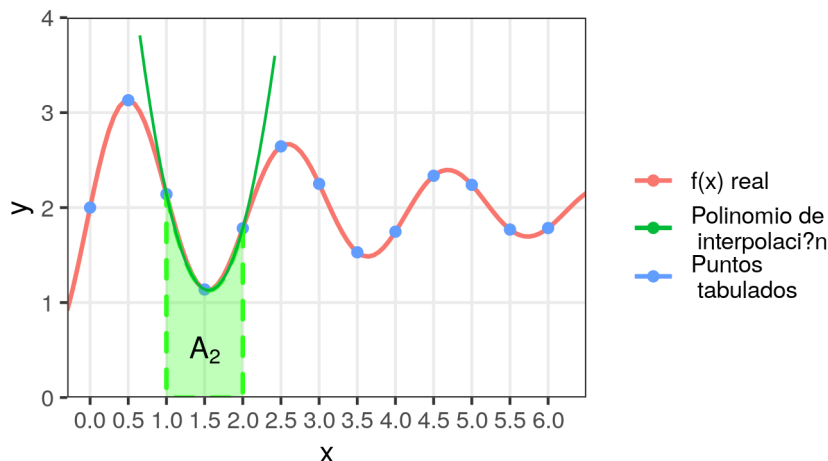
$$A_1 = 2.7766$$



- De manera semejante, se puede emplear la interpolación cuadrática de Newton para obtener una aproximación de la integral entre x_2 y x_4 :

$$\int_{x_2}^{x_4} f(x)dx \cong A_2 = \frac{h}{3}(y_2 + 4y_3 + y_4)$$

$$A_2 = 1.4133$$



- Y sucesivamente para todos los intervalos:

$$\int_{x_{i-1}}^{x_{i+1}} f(x)dx \cong \frac{h}{3}(y_{i-1} + 4y_i + y_{i+1}) \quad i = 1, 3, 5, \dots, n-1$$

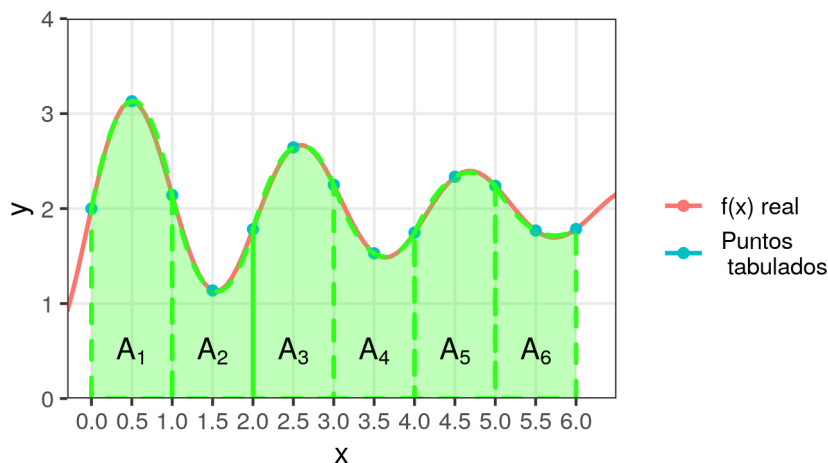
- De modo que la suma de estas áreas resulta ser la aproximación para la integral entre x_0 y x_n :

$$\int_{x_0}^{x_n} f(x)dx \cong \sum_{\substack{i=1 \\ i \text{ impar}}}^{n-1} \frac{h}{3}(y_{i-1} + 4y_i + y_{i+1}) = \frac{h}{3} \left(y_0 + y_n + 2 \sum_{\substack{i=2 \\ i \text{ par}}}^{n-2} y_i + 4 \sum_{\substack{i=1 \\ i \text{ impar}}}^{n-1} y_i \right)$$

- La fórmula hallada se conoce como **fórmula de Simpson de 1/3** y se la simboliza con:

$$A_{1/3} = \frac{h}{3} \left(y_0 + y_n + 2 \sum_{\substack{i=2 \\ i \text{ par}}}^{n-2} y_i + 4 \sum_{\substack{i=1 \\ i \text{ impar}}}^{n-1} y_i \right)$$

- Para poder aplicarla, es necesario que la cantidad de puntos tabulados sea impar, es decir que la tabla tenga una cantidad par de intervalos.
- Gráficamente:



- En el ejemplo: $A_{1/3} = 12.3833$.
- El valor exacto es: $\int_0^6 f(x)dx = 12.2935$, con lo cual el error relativo de la aproximación con la fórmula trapecial fue: 7.3 %.

5.1.3. Fórmula de Simpson de 3/8

- Si la interpolación es de tercer orden y la integral sólo se calcula entre los 4 primeros valores de x , se obtiene:

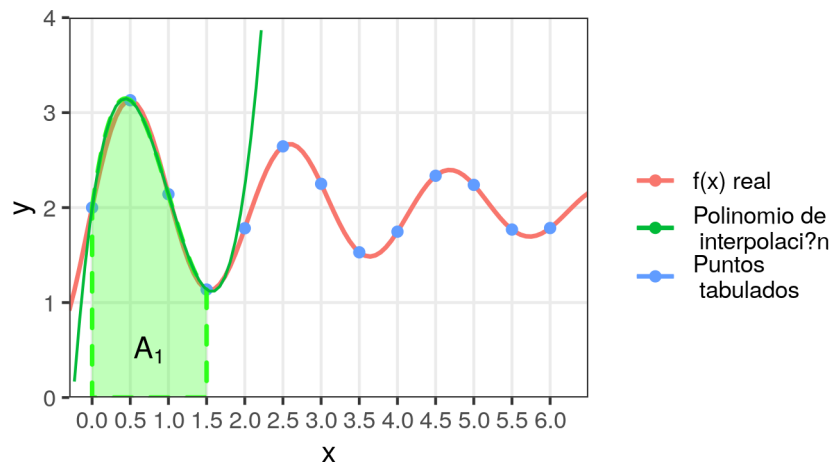
$$\begin{aligned} \int_{x_0}^{x_3} f(x)dx &\cong \int_0^3 \left[y_0 + k\Delta y_0 + \left(\frac{k^2}{2} - \frac{k}{2} \right) \Delta^2 y_0 + \left(\frac{k^3}{6} - \frac{k^2}{2} + \frac{k}{3} \right) \Delta^3 y_0 \right] h dk \\ &= h \left[y_0 k + \frac{k^2}{2} \Delta y_0 + \left(\frac{k^3}{6} - \frac{k^2}{4} \right) \Delta^2 y_0 + \left(\frac{k^4}{24} - \frac{k^3}{6} + \frac{k^2}{6} \right) \Delta^3 y_0 \right] \Big|_0^3 \\ &= h \left[3y_0 + \frac{9}{2} \Delta y_0 + \frac{9}{4} \Delta^2 y_0 + \frac{3}{8} \Delta^3 y_0 \right] \end{aligned}$$

- Dado que $\Delta y_0 = y_1 - y_0$, $\Delta^2 y_0 = \Delta y_1 - \Delta y_0 = y_2 - 2y_1 + y_0$, y $\Delta^3 y_0 = \Delta^2 y_1 - \Delta^2 y_0 = y_3 - 3y_2 - 3y_1 + y_0$ nos queda:

$$\int_{x_0}^{x_3} f(x)dx \cong \frac{3}{8}h(y_0 + 3y_1 + 3y_2 + y_3)$$

- Geométricamente, esto equivale al área A_1 encerrada entre el eje de las abscisas, x_0 y x_3 y el polinomio integrador que pasa por (x_0, y_0) , (x_1, y_1) , (x_2, y_2) y (x_3, y_3) :

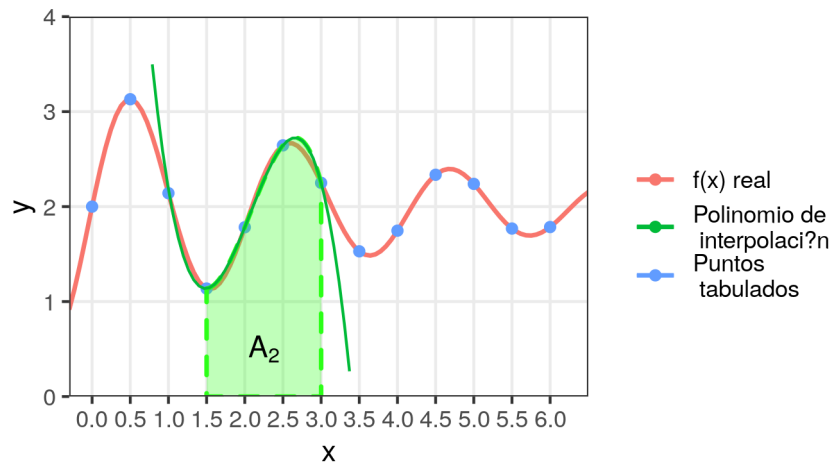
$$A_1 = 3.5531$$



- De manera semejante, se puede emplear la interpolación cúbica de Newton para obtener una aproximación de la integral entre x_3 y x_6 :

$$\int_{x_3}^{x_6} f(x)dx \cong A_2 = \frac{3}{8}h(y_3 + 3y_4 + 3y_5 + y_6)$$

$$A_1 = 3.1219$$



- Y sucesivamente para todos los intervalos:

$$\int_{x_i}^{x_{i+3}} f(x) dx \cong \frac{3}{8} h (y_i + 3y_{i+1} + 3y_{i+2} + y_{i+3}) \quad i = 0, 3, 6, \dots, n-3$$

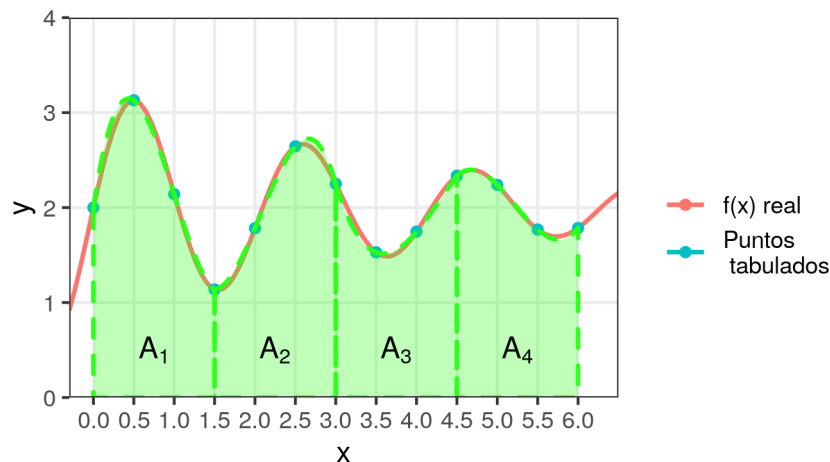
- De modo que la suma de estas áreas resulta ser la aproximación para la integral entre x_0 y x_n :

$$\begin{aligned} \int_{x_0}^{x_n} f(x) dx &\cong \sum_{\substack{i=0 \\ i \text{ múltiplo de } 3}}^{n-3} \frac{3}{8} h (y_i + 3y_{i+1} + 3y_{i+2} + y_{i+3}) \\ &= \frac{3}{8} h \left(y_0 + y_n + 2 \sum_{\substack{i=3 \\ i \text{ múltiplo de } 3}}^{n-3} y_i + 3 \sum_{\substack{i=1 \\ i \text{ no múltiplo de } 3}}^{n-1} y_i \right) \end{aligned}$$

- La fórmula hallada se conoce como **fórmula de Simpson de 3/8** y se la simboliza con:

$$A_{3/8} = \frac{3}{8} h \left(y_0 + y_n + 2 \sum_{\substack{i=3 \\ i \text{ múltiplo de } 3}}^{n-3} y_i + 3 \sum_{\substack{i=1 \\ i \text{ no múltiplo de } 3}}^{n-1} y_i \right)$$

- Para poder aplicarla, es necesario que la cantidad de intervalos en la tabla sea múltiplo de 3.
- Gráficamente:



- En el ejemplo: $A_{3/8} = 12.4088$.
- El valor exacto es: $\int_0^6 f(x) dx = 12.2935$, con lo cual el error relativo de la aproximación con la fórmula trapecial fue: 9.4 %.

5.2. Derivación numérica

- Para aproximar la derivada de una función en un punto, nuevamente haremos uso del polinomio interpolador de Newton:

$$\begin{aligned} f(x) &\cong y_0 + k\Delta y_0 + \frac{k(k-1)}{2!}\Delta^2 y_0 + \frac{k(k-1)(k-2)}{3!}\Delta^3 y_0 + \dots \\ &= y_0 + k\Delta y_0 + \left(\frac{k^2}{2} - \frac{k}{2}\right)\Delta^2 y_0 + \left(\frac{k^3}{6} - \frac{k^2}{2} + \frac{k}{3}\right)\Delta^3 y_0 + \dots \end{aligned}$$

- Se debe derivar con respecto a x el miembro derecho de la expresión anterior, aplicando la Regla de la Cadena ya que $k = (x - x_0)/h$.
- Por simplicidad, lo mostraremos sólo con el polinomio interpolador cuadrático.
- Aproximación de la derivada con el polinomio interpolador cuadrático de Newton:

$$\begin{aligned} f(x) &\cong y_0 + k\Delta y_0 + \frac{k^2}{2}\Delta^2 y_0 - \frac{k}{2}\Delta^2 y_0 \\ k &= \frac{x - x_0}{h} \implies \frac{\partial k}{\partial x} = \frac{1}{h} \end{aligned}$$

$$\begin{aligned} f'(x) &\cong \Delta y_0 \frac{1}{h} + \Delta^2 y_0 k \frac{1}{h} - \frac{\Delta^2 y_0}{2} \frac{1}{h} \\ &= \frac{1}{h} \left[\Delta y_0 + \left(k - \frac{1}{2}\right) \Delta^2 y_0 \right] \end{aligned}$$

Retomando el Ejemplo 1 del capítulo anterior (integración): vamos a aproximar el valor de $f'(3.4)$.

x_k	y_k	Δy_k	$\Delta^2 y_k$	$\Delta^3 y_k$	$\Delta^4 y_k$	$\Delta^5 y_k$
2	0,3010	0,1761	-0,0511	0,0230	-0,0127	0,0081
3	0,4771	0,1250	-0,0281	0,0103	-0,0046	-
4	0,6021	0,0969	-0,0178	0,0057	-	-
5	0,6990	0,0791	-0,0121	-	-	-
6	0,7781	0,0670	-	-	-	-
7	0,8451	-	-	-	-	-

- $x = 3, 4$
- $x_0 = 3$
- $h = 1$
- $k = \frac{x - x_0}{h} = \frac{3.4 - 3}{1} = 0.4$
- $\Delta y_0 = 0.1250$; $\Delta^2 y_0 = -0.0281$

$$f'(x) \cong \frac{1}{h} \left[\Delta y_0 + \left(k - \frac{1}{2}\right) \Delta^2 y_0 \right] = 0.1250 + (-0.1)(-0.0281) = 0.12781$$

Nota: esta fórmula se conoce como *aproximación por diferencias hacia adelante*, pero se pueden lograr aproximaciones más precisas de otras formas, por ejemplo, haciendo que el punto de interés x esté en el centro del rango del polinomio interpolador (*aproximación por diferencias centrales*).

Unidad 6

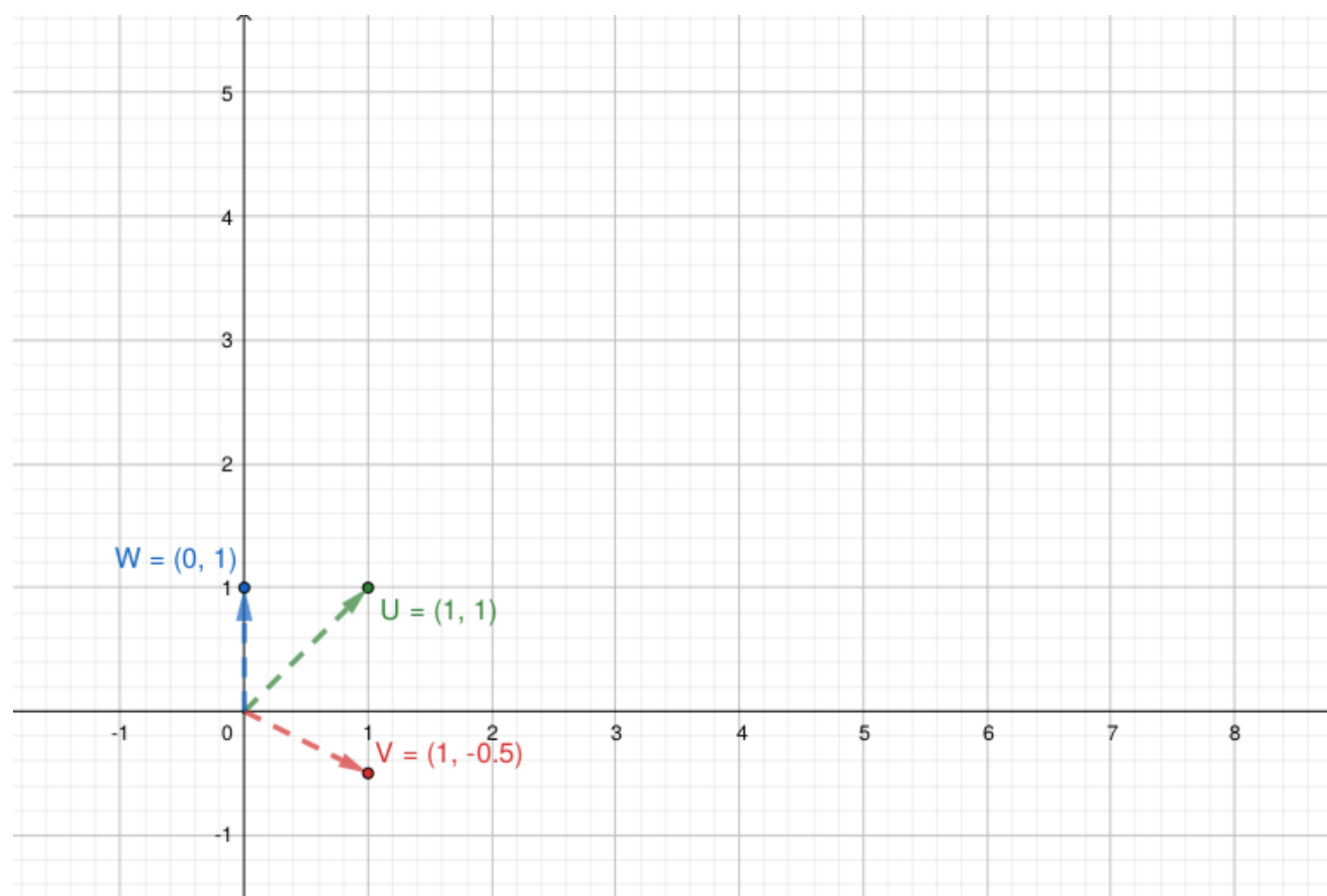
Autovalores y Autovectores

6.1. Generalidades

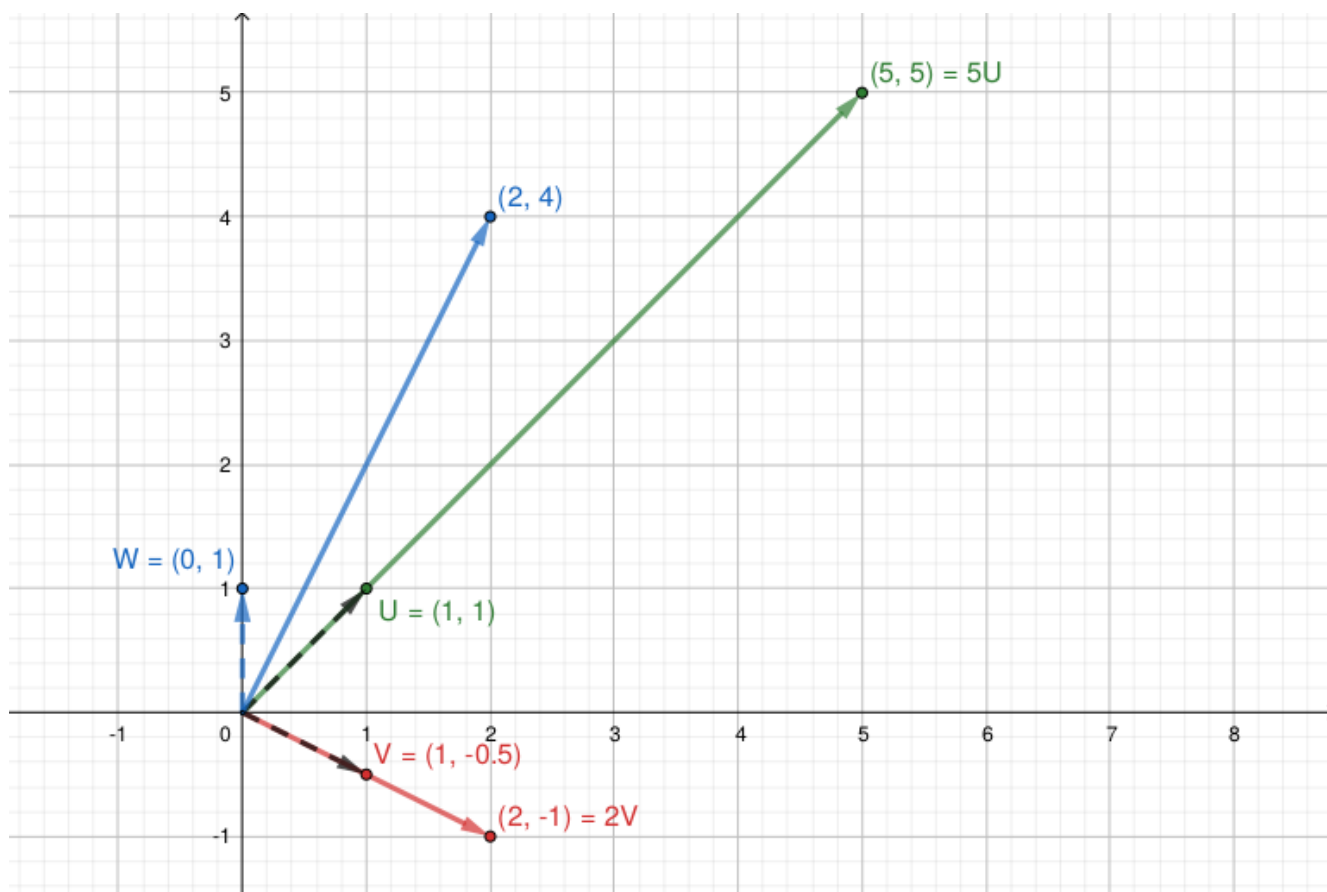
- Los **autovalores** y **autovectores** son esas cosas raras que aparecen por todos lados pero nunca terminamos por entender.
- El objetivo de esta unidad es ver métodos para su cálculo, pero antes vamos a repasar qué son (**informalmente, sin rigurosidad**, el que avisa no traiciona...)
- En muchas disciplinas los objetos que se estudian se representan con *vectores* (ej. \mathbf{x} , \mathbf{y}) y las cosas que se hacen con ellos son *transformaciones lineales*, que se representan como *matrices* (ej. \mathbf{A}).
- Así, en muchas situaciones las relaciones que importan entre esos objetos/vectores se expresan como:

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

- Esto abarca desde sistemas de ecuaciones lineales (presentes casi en todos lados en ciencia) hasta problemas muy sofisticados en ingeniería.
- Ahora bien, en general no es muy fácil mirar a la matriz \mathbf{A} y directamente darse cuenta qué es lo que va a pasar cuando se la multipliquemos a \mathbf{x} .
- Sin embargo, podríamos encontrar casos donde haya una relación muy simple entre el vector \mathbf{x} y el vector resultado $\mathbf{y}=\mathbf{A}\mathbf{x}$.
- Por ejemplo, si miramos la matriz $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ y se la multiplicamos al vector $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, ¡nos da como resultado el mismo vector \mathbf{x} !
- Es decir, que para ese vector, es muy fácil ver qué aspecto tiene $\mathbf{A}\mathbf{x}$.
- Se puede generalizar esta observación con el concepto de **autovectores**.
- Un **autovector** de una matriz \mathbf{A} es cualquier vector \mathbf{x} para el que sólo cambia su escala cuando se lo multiplica con \mathbf{A} , es decir: $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, para algún número λ real o complejo, que recibe el nombre de **autovalor**.
- Entonces si una matriz \mathbf{A} describe algún tipo de sistema, los autovectores son aquellos vectores que, cuando pasan por el sistema, se modifican en una forma muy sencilla.
- Por ejemplo, si \mathbf{A} describe operaciones geométricas, en principio \mathbf{A} podría estirar y rotar a los vectores, sin embargo, a sus autovectores lo único que puede hacerles es estirarlos, no rotarlos.
- Sea: $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$ y $\mathbf{w} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.
- En este gráfico podemos ver los vectores antes de transformarlos (multiplicarlos) mediante \mathbf{A} :



- Y en este gráfico podemos ver como quedan luego de la transformación:



- \mathbf{u} y \mathbf{v} no cambiaron su dirección, sólo su norma: son **autovectores** de \mathbf{A} , asociados a los autovalores 5 y 2.
- En cambio, la matriz \mathbf{A} modificó la dirección de \mathbf{w} , entonces no es un autovector.

6.1.1. Definición

- Dada una matriz \mathbf{A} cuadrada de orden n , llamamos **autovector** o **vector propio** de \mathbf{A} a todo vector \mathbf{x} de orden n cuya dirección no se modifica al transformarlo mediante \mathbf{A} .
 - Transformarlo mediante \mathbf{A} significa realizar el producto \mathbf{Ax} dando como resultado un nuevo vector de orden n .
 - Que la dirección de \mathbf{x} no se modifique significa que el nuevo vector debe ser múltiplo escalar de \mathbf{x} , es decir, igual a $\lambda\mathbf{x}$, con $\lambda \in \mathbb{C}$, que recibe el nombre de **autovalor** o **valor propio** de \mathbf{A} .
- Lo anterior se resume en la siguiente expresión: \mathbf{x} es un autovector y λ es un autovalor de \mathbf{A} si:

$$\mathbf{Ax} = \lambda\mathbf{x}, \quad \mathbf{x} \neq \mathbf{0}, \quad \lambda \in \mathbb{C}$$

Observación

- Se debe observar que si \mathbf{x} es un autovector con el autovalor λ entonces cualquier múltiplo escalar de \mathbf{x} diferente del vector nulo es también un autovector con el autovalor λ .

6.1.2. Propiedades

- Dada una matriz \mathbf{A} cuadrada de orden n :
 - \mathbf{A} tiene n autovalores, $\lambda_1, \lambda_2, \dots, \lambda_n$, los cuales no necesariamente son todos distintos.
 - $\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$.
 - $\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$.
 - Los autovalores de \mathbf{A}^k son $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$.
 - Si \mathbf{A} es real y simétrica todos sus autovalores son reales y los autovectores correspondientes a distintos autovalores son ortogonales.

- Si \mathbf{A} es triangular los valores propios son los elementos diagonales.
- Los autovalores de una matriz y su transpuesta son los mismos.
- Si \mathbf{A} tiene inversa, los autovalores de \mathbf{A}^{-1} son $1/\lambda_1, 1/\lambda_2, \dots, 1/\lambda_n$.
- Los autovalores de $\alpha\mathbf{A}$ son $\alpha\lambda_1, \alpha\lambda_2, \dots, \alpha\lambda_n$, $\alpha \in \mathbb{R}$.
- Las matrices \mathbf{A} y $\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ (forma cuadrática) tienen los mismos valores propios.

6.2. Obtención los autovalores y autovectores

- A partir de la expresión anterior:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \Rightarrow \mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0} \Rightarrow (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$$

- Esto es un sistema de ecuaciones lineales con matriz de coeficientes $\mathbf{A} - \lambda\mathbf{I}$ y vector de términos independientes $\mathbf{0}$, es decir, es un **sistema homogéneo** y como tal tiene solución no nula si: $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. La razón es:
 - Como ya sabemos, un sistema $\mathbf{C}\mathbf{x} = \mathbf{b}$ tiene solución única si $\det \mathbf{C} \neq 0$.
 - En nuestro caso, siendo $\mathbf{b} = \mathbf{0}$, si el determinante de $(\mathbf{A} - \lambda\mathbf{I})$ es distinto de cero, la solución única tiene que ser la trivial, $\mathbf{x} = \mathbf{0}$.
 - Por lo tanto, para que el sistema tenga soluciones no nulas (los autovectores) requerimos que el determinante sea igual a 0, lo cual significa que el sistema va a tener infinitas soluciones (infinitos autovectores).
- El desarrollo de la expresión anterior conduce a un polinomio de grado n en la incógnita λ que igualado a cero es llamado **ecuación característica** y su resolución permite hallar los autovalores.

Ejemplo

$$\mathbf{A} = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{bmatrix} \Rightarrow$$

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 5-\lambda & -2 & 0 \\ -2 & 3-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{vmatrix} = \dots = -\lambda^3 + 9\lambda^2 - 18\lambda + 6 = 0$$

- Las soluciones de la ecuación característica son $\lambda_1 = 6.2899$, $\lambda_2 = 2.2943$ y $\lambda_3 = 0.4158$, los cuales son los autovalores de \mathbf{A} .
- Hallar la ecuación característica ya es demasiado trabajoso para $n = 3$, y mucho más será para mayor n ... por eso veremos métodos que directamente nos den los coeficientes de esta ecuación.
- Pero nos faltan los autovectores!
- Para eso hacemos uso de la definición: \mathbf{x} es un autovector de \mathbf{A} asociado al autovalor λ si $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$.
- Tomamos uno de los autovalores, por ejemplo, $\lambda_1 = 6.2899$ y resolvemos el sistema de ecuaciones que la expresión anterior plantea:

$$(\mathbf{A} - 6.2899\mathbf{I})\mathbf{x} = \mathbf{0} \Rightarrow \begin{bmatrix} -1.2899 & -2 & 0 \\ -2 & -3.2899 & -1 \\ 0 & -1 & -5.2899 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} -1.2899x_1 - 2x_2 = 0 \\ -2x_1 - 3.2899x_2 - x_3 = 0 \\ -x_2 - 5.2899x_3 = 0 \end{cases} \Rightarrow \begin{cases} x_1 = 8.2018x_3 \\ x_2 = -5.2899x_3 \\ x_3 \in \mathbb{R} \end{cases}$$

- Como se puede ver la solución de este sistema homogéneo no es única, representando los infinitos autovectores asociados a $\lambda_1 = 6.2899$. Por ejemplo, si elegimos $x_3 = 1$, obtenemos el autovector:

$$\mathbf{x}_1 = \begin{bmatrix} 8.2018 \\ -5.2899 \\ 1 \end{bmatrix}$$

- En general, se resuelve informando el autovector de norma 1 que sí es único.
- De la misma forma se procede con los restantes autovalores λ_2 y λ_3 .

Resumen 1: Obtener autovalores y autovectores

- **Paso 1:** desarrollar la expresión de $\det(\mathbf{A} - \lambda \mathbf{I})$ para obtener la ecuación característica (muy engorroso para $n > 3$):

$$f(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^n + b_1 \lambda^{n-1} + \dots + b_{n-1} \lambda + b_n = 0$$

- **Paso 2:** resolver la ecuación característica para hallar los autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$. Dependiendo de n , podemos hacerlo a mano, con la calculadora o con los métodos de la Unidad 2.
- **Paso 3:** tomar cada autovalor λ_i y resolver el sistema de ecuaciones lineales $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{x} = \mathbf{0}$. No nos sirven los métodos de la Unidad 3 porque este sistema es compatible indeterminado, realizarlo “a mano” y dar una expresión para los infinitos autovectores o informar el autovector de norma 1.

6.3. Método de Krylov

- Como ya mencionamos, el desarrollo de $\det(\mathbf{A} - \lambda \mathbf{I})$ para obtener la ecuación característica tal como lo vimos en el ejemplo inicial se vuelve engorroso rápidamente.
- El método de Krylov permite obtenerla de manera sencilla, basándose en el siguiente teorema:
- **Teorema de Cayley-Hamilton:** toda matriz cuadrada \mathbf{A} verifica su propia ecuación característica. Es decir, siendo la ecuación característica:

$$f(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^n + b_1 \lambda^{n-1} + \dots + b_{n-1} \lambda + b_n = 0,$$

se verifica que:

$$f(\mathbf{A}) = \mathbf{A}^n + b_1 \mathbf{A}^{n-1} + \dots + b_{n-1} \mathbf{A} + b_n \mathbf{I} = \mathbf{0}_{n \times n}$$

Ejemplo

$$\mathbf{A} = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\mathbf{A}^2 = \begin{bmatrix} 29 & -16 & 2 \\ -16 & 14 & -4 \\ 2 & -1 & 1 \end{bmatrix}$$

$$\mathbf{A}^3 = \begin{bmatrix} 177 & -108 & 18 \\ -108 & 78 & -18 \\ 18 & -18 & 6 \end{bmatrix}$$

$$f(\mathbf{A}) = f(\mathbf{A}) = \mathbf{A}^3 + b_1\mathbf{A}^2 + b_2\mathbf{A} + b_3\mathbf{I} = \mathbf{0} \Rightarrow$$

$$\begin{bmatrix} 177 & -108 & 18 \\ -108 & 78 & -18 \\ 18 & -18 & 6 \end{bmatrix} + b_1 \begin{bmatrix} 29 & -16 & 2 \\ -16 & 14 & -4 \\ 2 & -1 & 1 \end{bmatrix} + b_2 \begin{bmatrix} 5 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

$$+ b_3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} 177 + 29b_1 + 5b_2 + b_3 & -108 - 16b_1 - 2b_2 & 18 + 2b_1 \\ -108 - 16b_1 - 2b_2 & 78 + 14b_1 + 3b_2 + b_3 & -18 - 4b_1 - b_2 \\ 18 + 2b_1 & -18 - 4b_1 - 1b_2 & 6 + 2b_1b_2 + b_3 \end{bmatrix} =$$

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Cualquiera de las columnas constituyen un sistema de tres ecuaciones lineales en las incógnitas b_1, b_2 y b_3 , los coeficientes de la ecuación característica.
- Podemos usar el siguiente artificio para generar un único sistema de ecuaciones:

$$f(\mathbf{A}_{n \times n}) = \mathbf{0}_{n \times n} \Rightarrow f(\mathbf{A})\mathbf{y} = \mathbf{0}\mathbf{y} = \mathbf{0}_{n \times 1} \quad \forall \mathbf{y}_{n \times 1} \in \mathbb{R}^n$$

- Por ejemplo, tomando $\mathbf{y} = [1 \ 0 \ 0]^t$, nos queda:

$$f(\mathbf{A})\mathbf{y} = \mathbf{0}\mathbf{y}$$

$$\Rightarrow (\mathbf{A}^3 + b_1\mathbf{A}^2 + b_2\mathbf{A} + b_3\mathbf{I})\mathbf{y} = \mathbf{0}$$

$$\Rightarrow \mathbf{A}^3\mathbf{y} + b_1\mathbf{A}^2\mathbf{y} + b_2\mathbf{A}\mathbf{y} + b_3\mathbf{y} = \mathbf{0}$$

$$\Rightarrow \begin{bmatrix} 177 & -108 & 18 \\ -108 & 78 & -18 \\ 18 & -18 & 6 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b_1 \begin{bmatrix} 29 & -16 & 2 \\ -16 & 14 & -4 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} +$$

$$+ b_2 \begin{bmatrix} 5 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 177 \\ -108 \\ 18 \end{bmatrix} + b_1 \begin{bmatrix} 29 \\ -16 \\ 2 \end{bmatrix} + b_2 \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix} + b_3 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} 29 & 5 & 1 \\ -16 & -2 & 0 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} -177 \\ 108 \\ -18 \end{bmatrix}$$

- Lo anterior no es más que un sistema de tres ecuaciones lineales, $\mathbf{C}\mathbf{b} = \mathbf{d}$, donde:
 - el vector incógnitas es $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, los coeficientes de la ecuación característica.
 - el vector de términos independientes es $\mathbf{d} = -\mathbf{A}^3\mathbf{y} = \begin{bmatrix} -177 \\ 108 \\ -18 \end{bmatrix}$.
 - la matriz de coeficientes es $\mathbf{C} = [\mathbf{A}^2\mathbf{y} \quad \mathbf{A}\mathbf{y} \quad \mathbf{y}] = \begin{bmatrix} 29 & 5 & 1 \\ -16 & -2 & 0 \\ 2 & 0 & 0 \end{bmatrix}$
- Dependiendo de n , podemos resolver este sistema “a mano”, con la calculadora o con algunos de los métodos de la Unidad 3.
- En el ejemplo, el resultado es: $b_1 = -9, b_2 = 18$ y $b_3 = -6$.
- La ecuación característica entonces es:

$$\lambda^3 - 9\lambda^2 + 18\lambda - 6 = 0$$

- Esta ecuación coincide con la que obtuvimos en la sección anterior.
- A partir de aquí, se debe continuar desde el Paso 2 del Resumen 1 para hallar los autovalores y sus respectivos autovectores.

Resumen 2: Método de Krylov

- **Qué necesita:** la matriz \mathbf{A} y un vector \mathbf{y} .
- **Qué nos da:** un sistema de ecuaciones para obtener los coeficientes de la ecuación característica.
- **Paso 1:** elegir un vector \mathbf{y} de dimensión $n \times 1$.
- **Paso 2:** crear la matriz de coeficientes $\mathbf{C} = [\mathbf{A}^{n-1}\mathbf{y} \quad \dots \quad \mathbf{A}^2\mathbf{y} \quad \mathbf{A}\mathbf{y} \quad \mathbf{y}]$.
- **Paso 3:** crear el vector de términos independientes $\mathbf{d} = -\mathbf{A}^n\mathbf{y}$, de dimensión $n \times 1$.

- **Paso 4:** resolver el sistema $\mathbf{Cb} = \mathbf{d}$, donde el vector de incógnitas \mathbf{b} son los coeficientes de la ecuación característica.
- **Paso 5:** formar la ecuación característica y continuar desde el Paso 2 del Resumen 1 para hallar los autovalores y sus respectivos autovectores.

6.4. Método de Faddeev-LeVerrier

- Este método propone hallar los coeficientes b_k de la ecuación característica:

$$f(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}) = \lambda^n + b_1 \lambda^{n-1} + \dots + b_{n-1} \lambda + b_n = 0$$

mediante el siguiente cálculo iterativo:

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{A} & b_1 &= -\text{tr}(\mathbf{M}_1) \\ \mathbf{M}_k &= \mathbf{A}(\mathbf{M}_{k-1} + b_{k-1} \mathbf{I}) & b_k &= -\frac{\text{tr}(\mathbf{M}_k)}{k} \quad k = 2, 3, \dots, n \end{aligned}$$

- Este método se deriva a partir de propiedades de matrices conjugadas.
- En nuestro ejemplo, tenemos:

$$\begin{aligned} \mathbf{M}_1 &= \mathbf{A} = \begin{bmatrix} 5 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{bmatrix} & b_1 &= -\text{tr}(\mathbf{M}_1) = -9 \\ \mathbf{M}_2 &= \mathbf{A}(\mathbf{M}_1 + b_1 \mathbf{I}) = \begin{bmatrix} -16 & 2 & 2 \\ 2 & -13 & 5 \\ 2 & 5 & -7 \end{bmatrix} & b_2 &= -\frac{\text{tr}(\mathbf{M}_2)}{2} = 18 \\ \mathbf{M}_3 &= \mathbf{A}(\mathbf{M}_2 + b_2 \mathbf{I}) = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} & b_3 &= -\frac{\text{tr}(\mathbf{M}_3)}{3} = -6 \end{aligned}$$

- La ecuación característica entonces es:

$$\lambda^3 - 9\lambda^2 + 18\lambda - 6 = 0$$

- A partir de aquí, otra vez se debe continuar desde el Paso 2 del Resumen 1 para hallar los autovalores y sus respectivos autovectores.
- Este método también sirve para calcular \mathbf{A}^{-1} .
- Por Cayley-Hamilton, ya sabemos que:

$$f(\mathbf{A}) = \mathbf{A}^n + b_1 \mathbf{A}^{n-1} + \dots + b_{n-2} \mathbf{A}^2 + b_{n-1} \mathbf{A} + b_n \mathbf{I} = \mathbf{0}$$

- Premultiplicando por \mathbf{A}^{-1} nos queda:

$$\begin{aligned} \mathbf{A}^{-1}(\mathbf{A}^n + b_1 \mathbf{A}^{n-1} + \dots + b_{n-2} \mathbf{A}^2 + b_{n-1} \mathbf{A} + b_n \mathbf{I}) &= \mathbf{A}^{-1} \mathbf{0} \\ \mathbf{A}^{n-1} + b_1 \mathbf{A}^{n-2} + \dots + b_{n-2} \mathbf{A} + b_{n-1} \mathbf{I} + b_n \mathbf{A}^{-1} &= \mathbf{0} \\ \mathbf{A}^{-1} &= -\frac{1}{b_n} (\mathbf{A}^{n-1} + b_1 \mathbf{A}^{n-2} + \dots + b_{n-2} \mathbf{A} + b_{n-1} \mathbf{I}) \\ \mathbf{A}^{-1} &= -\frac{1}{b_n} (\mathbf{M}_{n-1} + b_{n-1} \mathbf{I}) \end{aligned}$$

... donde el último reemplazo se deduce a partir de la fórmula iterativa vista antes:

$$\begin{aligned}
 \mathbf{M}_1 &= \mathbf{A} \\
 \mathbf{M}_2 &= \mathbf{A}(\mathbf{M}_1 + b_1 \mathbf{I}) = \mathbf{A}(\mathbf{A} + b_1 \mathbf{I}) = \mathbf{A}^2 + b_1 \mathbf{A} \\
 \mathbf{M}_3 &= \mathbf{A}(\mathbf{M}_2 + b_2 \mathbf{I}) = \mathbf{A}(\mathbf{A}^2 + b_1 \mathbf{A} + b_1 \mathbf{I}) = \mathbf{A}^3 + b_1 \mathbf{A}^2 + b_2 \mathbf{A} \\
 \mathbf{M}_4 &= \mathbf{A}(\mathbf{M}_3 + b_3 \mathbf{I}) = \mathbf{A}^4 + b_1 \mathbf{A}^3 + b_2 \mathbf{A}^2 + b_3 \mathbf{A} \\
 &\vdots \\
 \mathbf{M}_{n-1} &= \mathbf{A}(\mathbf{M}_{n-2} + b_{n-2} \mathbf{I}) = \mathbf{A}^{n-1} + b_1 \mathbf{A}^{n-2} + b_2 \mathbf{A}^{n-3} + \dots + b_{n-2} \mathbf{A}
 \end{aligned}$$

Resumen 3: Método de Faddeev-LeVerrier

- **Qué necesita:** la matriz \mathbf{A}
- **Qué nos da:** los coeficientes de la ecuación característica.
- **Paso 1:** calcular los coeficientes b_k de la ecuación característica con la fórmula recursiva:

$$\begin{aligned}
 \mathbf{M}_1 &= \mathbf{A} & b_1 &= -\text{tr}(\mathbf{M}_1) \\
 \mathbf{M}_k &= \mathbf{A}(\mathbf{M}_{k-1} + b_{k-1} \mathbf{I}) & b_k &= -\frac{\text{tr}(\mathbf{M}_k)}{k} \quad k = 2, 3, \dots, n
 \end{aligned}$$

- **Paso 2:** formar la ecuación característica y continuar desde el Paso 2 del Resumen 1 para hallar los autovalores y sus respectivos autovectores.

6.5. Método de Aproximaciones Sucesivas o de las Potencias

- **Definición:** si λ es un autovalor de \mathbf{A} tal que en valor absoluto es mayor que cualquier otro autovalor, se dice que es un **autovalor dominante** y sus autovectores se llaman **autovectores dominantes**.
- El **método de las potencias** dice que si \mathbf{A} tiene un autovalor dominante y \mathbf{v} es su autovector normalizado, la sucesión \mathbf{x}_k a partir de cualquier \mathbf{x}_0 no nulo converge a \mathbf{v} :

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1}$$

- El autovalor correspondiente está dado por el **cociente de Rayleigh**: si \mathbf{x} es un autovector de \mathbf{A} , entonces su correspondiente autovalor es:

$$\lambda = \frac{(\mathbf{A}\mathbf{x})^t \mathbf{x}}{\mathbf{x}^t \mathbf{x}}$$

- Se llama método de las potencias porque:

$$\begin{aligned}
 \mathbf{x}_1 &= \mathbf{A}\mathbf{x}_0 \\
 \mathbf{x}_2 &= \mathbf{A}\mathbf{x}_1 = \mathbf{A}^2 \mathbf{x}_0 \\
 \mathbf{x}_3 &= \mathbf{A}\mathbf{x}_2 = \mathbf{A}^3 \mathbf{x}_0 \\
 &\vdots \\
 \mathbf{x}_k &= \mathbf{A}\mathbf{x}_{k-1} = \mathbf{A}^k \mathbf{x}_0
 \end{aligned}$$

- Este método tiende a producir aproximaciones en donde los elementos de \mathbf{x} tienen gran magnitud, lo cual produce problemas (errores de desbordamiento, *overflow error*).
- Por eso, en la práctica se añade un escalamiento en cada paso iterativo, dividiendo por el elemento de mayor magnitud del paso anterior.
- **Método de las potencias:** si \mathbf{A} tiene un autovalor dominante, la siguiente sucesión c_k converge al mismo mientras que la sucesión \mathbf{x}_k converge a uno de sus autovectores dominantes:

$$\mathbf{x}_k = \frac{1}{c_k} \mathbf{A} \mathbf{x}_{k-1}$$

donde c_k es la coordenada de mayor tamaño de $\mathbf{A} \mathbf{x}_{k-1}$ y \mathbf{x}_0 es cualquier vector no nulo.

- Retomando nuestro ejemplo:

k	\mathbf{x}_k	$\mathbf{A} \mathbf{x}_k$	c_k	$\mathbf{x}_{k+1} = \mathbf{A} \mathbf{x}_k / c_k$	Error (L_2)
0	$[1 \ 1 \ 1]^t$	$[3 \ 0 \ 0]^t$	3	$[1 \ 0 \ 0]^t$	1.4142
1	$[1 \ 0 \ 0]^t$	$[5 \ -2 \ 0]^t$	5	$[1 \ -0.4 \ 0]^t$	0.4
2	$[1 \ -0.4 \ 0]^t$	$[5.8 \ -3.2 \ 0.4]^t$	5.8	$[1 \ -0.5517 \ 0.0690]^t$	0.1667
3	$[1 \ -0.5517 \ 0.0690]^t$	$[6.1034 \ -3.7241 \ 0.6207]^t$	6.1034	$[1 \ -0.6102 \ 0.1017]^t$	0.0690
4	$[1 \ -0.6102 \ 0.1017]^t$	$[6.2203 \ -3.9322 \ 0.7119]^t$	6.2203	$[1 \ -0.6322 \ 0.1144]^t$	0.0254
...
16	$[1 \ -0.644972 \ 0.1219239]^t$	$[6.2899 \ -4.0568 \ 0.7669]^t$	6.2899	$[1 \ -0.644972 \ 0.1219241]^t$	3.956E-7

6.5.1. Variantes del método de las potencias

Método de las potencias inversas

- Es una variante del método de las aproximaciones sucesivas o de las potencias.
- Permite hallar el menor autovalor de \mathbf{A} .
- Se aplica el método a \mathbf{A}^{-1} para hallar su mayor autovalor.
- Pero como los autovalores de \mathbf{A}^{-1} son los recíprocos de los de \mathbf{A} , el autovalor así hallado es el recíproco del menor autovalor de \mathbf{A} .

Método de las potencias con deflación (o de Hotelling)

- Es una variante del método de las aproximaciones sucesivas o de las potencias.
- Una vez hallado el mayor autovalor λ_1 es posible encontrar el segundo mayor autovalor aplicando el mismo método sobre la matriz $\mathbf{A}_2 = \mathbf{A} - \lambda_1 \mathbf{u} \mathbf{u}^t$, donde $\mathbf{u} = \mathbf{x} / \|\mathbf{x}\|$, con \mathbf{x} el autovector hallado para λ_1 .
- Si $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ son los autovalores de \mathbf{A} , entonces $\{0, \lambda_2, \dots, \lambda_n\}$ son los de \mathbf{A}_2 .
- Repitiendo este proceso se encuentran los restantes autovalores.

Resumen 4: Método de las Aproximaciones Sucesivas o de las Potencias

- **Qué necesita:** la matriz \mathbf{A} y un vector inicial \mathbf{x}_0 .
- **Qué nos da:** el autovalor dominante de \mathbf{A} y su autovector.
- **Paso 1:** elegir un vector inicial \mathbf{x}_0 de dimensión $n \times 1$.
- **Paso 2:** repetir el siguiente proceso iterativo estableciendo un criterio para la convergencia:

$$\mathbf{x}_k = \frac{1}{c_k} \mathbf{A} \mathbf{x}_{k-1}$$

- **Paso 3:** al finalizar, c_k aproxima al autovalor dominante y \mathbf{x}_k a uno de sus autovectores.
- **Modificación 1:** hacer lo mismo con \mathbf{A}^{-1} nos da el recíproco del menor autovalor de \mathbf{A} y uno de sus autovectores.
- **Modificación 2:** aplicar sucesivamente este método modificando \mathbf{A} como establece Hotelling para hallar todos los autovalores.

Anexo: Teoremas útiles

Teorema del Valor Intermedio o de Bolzano

Sea f una función real continua en un intervalo cerrado $[a, b]$ con $f(a)$ y $f(b)$ de signos contrarios, es decir, $f(a) \cdot f(b) < 0$. Entonces existe al menos un punto c del intervalo abierto (a, b) con $f(c) = 0$.

Teorema del Valor Medio

Dada cualquier función f continua en el intervalo $[a, b]$ y derivable en el intervalo abierto (a, b) , entonces existe al menos algún punto c en el intervalo (a, b) tal que la tangente a la curva en c es paralela a la recta secante que une los puntos $(b, f(b))$ y $(a, f(a))$. Es decir:

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

Teorema de Taylor

Sea $k \geq 1$ un entero y la función $f : \mathbb{R} \rightarrow \mathbb{R}$ diferenciable k veces en el punto $x_0 \in \mathbb{R}$. Entonces existe una función $h_n : \mathbb{R} \rightarrow \mathbb{R}$ tal que:

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n}_{\text{Polinomio de Taylor de orden } n} + \underbrace{h_n(x)(x - x_0)^{n+1}}_{\text{resto}}$$

y $\lim_{x \rightarrow x_0} h_n(x) = 0$. Esta es la llamada **forma de Peano del resto**.

El polinomio que aparece en el teorema de Taylor se denomina **polinomio de Taylor de orden n** .

Existen diversas fórmulas explícitas para el resto. Una de ellas es la **forma de Lagrange**:

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

para algún número real ξ entre x_0 y x , siendo f diferenciable $n + 1$ veces.

Teorema del Punto Fijo

Dadas las siguientes condiciones:

- f es una función continua en el intervalo $[a, b]$
- $f(x) \in [a, b] \quad \forall x \in [a, b]$
- f' existe en (a, b) con $|f'(x)| \leq m < 1 \quad \forall x \in (a, b)$

Si x_0 es cualquier número en $[a, b]$, entonces la sucesión definida por

$$x_n = f(x_{n-1}), \quad n \geq 1,$$

converge al único punto fijo que f posee en $[a, b]$.

Demostración

1. Existencia de un punto fijo

- Si $f(a) = a$ o $f(b) = b$, la existencia del punto fijo es obvia.
- Entonces suponemos que $f(a) \neq a$ y $f(b) \neq b \implies f(a) > a$ y $f(b) < b$ por (b).
- Definimos $h(x) = f(x) - x$, continua en $[a; b]$ tal que:

$$h(a) = f(a) - a > 0 \text{ y } h(b) = f(b) - b < 0$$

- Por el Teorema del Valor Intermedio:

$$\exists p \in (a; b) \text{ tal que } h(p) = 0 \implies f(p) - p = 0$$

$$\implies p = f(p) \implies p \text{ es un punto fijo de } f$$

- Por lo tanto, f tiene al menos un punto fijo.

2. Unicidad del punto fijo

Sean p y q dos puntos fijos de f , $p \neq q$, $p, q \in [a; b]$:

$$\begin{aligned} |p - q| &= && \text{por ser puntos fijos} \\ = |f(p) - f(q)| && \text{Teorema Valor Medio } \xi \text{ entre } p \text{ y } q \\ = |f'(\xi)(p - q)| && \text{Valor absoluto de un producto} \\ = |f'(\xi)||p - q| && \text{por (c)} \\ \leq m|p - q| && \text{por (c)} \\ \leq |p - q| \\ \implies |p - q| < |p - q| \end{aligned}$$

Lo cual es una contradicción que proviene de la única suposición: $p \neq q$. Por lo tanto $p = q$ y el punto fijo es único.

3. Convergencia del proceso iterativo

Por (b), dado que $f : [a; b] \rightarrow [a; b]$, la sucesión $\{x_n\}_{n=0}^{\infty}$ está definida $\forall n \geq 0$ y $x_n \in [a; b] \quad \forall n$.

$$\begin{aligned} |x_n - p| && p \text{ es punto fijo} \\ = |x_n - f(p)| && \text{Recordando que: } x_n = f(x_{n-1}) \\ = |f(x_{n-1}) - f(p)| && \text{Teorema Valor Medio } \xi \text{ entre } p \text{ y } x_{n-1} \\ = |f'(\xi)||x_{n-1} - p| && \text{por (c)} \\ \leq m|x_{n-1} - p| \implies |x_n - p| \leq m|x_{n-1} - p| \end{aligned}$$

Del mismo modo podríamos ver que $|x_{n-1} - p| \leq m|x_{n-2} - p|$.

Aplicando dicha desigualdad inductivamente resulta:

$$\begin{aligned} |x_n - p| &\leq m|x_{n-1} - p| \leq m^2|x_{n-2} - p| \leq \dots \leq m^n|x_0 - p| \\ \implies |x_n - p| &\leq m^n|x_0 - p| \end{aligned}$$

Como $0 < m < 1$, el segundo miembro decrece a medida que n aumenta, sin importar cuál es x_0 :

$$\lim_{n \rightarrow \infty} |x_n - p| \leq \lim_{n \rightarrow \infty} m^n|x_0 - p| = 0$$

Entonces $|x_n - p| \xrightarrow{n \rightarrow \infty} 0$, es decir el método converge hacia el punto fijo p .

Teorema de Newton-Raphson

Supongamos que la función F es continua, con derivada segunda continua en el intervalo $[a; b]$, y que existe un número $p \in [a; b]$ tal que $F(p) = 0$. Si $F'(p) \neq 0$, entonces $\exists \delta > 0$ tal que la sucesión $\{x_n\}_{n=0}^{\infty}$ definida por el proceso iterativo:

$$x_n = x_{n-1} - \frac{F(x_{n-1})}{F'(x_{n-1})} \quad n = 1, 2, \dots \quad (6.1)$$

converge a p cualquiera sea la aproximación inicial $x_0 \in [p - \delta; p + \delta]$.

Demostración

Aplicamos el **Teorema de Taylor** para la función F en el punto x_0 empleando un polinomio de grado $n = 1$ y el resto en la forma de Lagrange:

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + \frac{F''(\xi)}{2!}(x - x_0)^2$$

donde ξ es un número real entre x y x_0 .

Si tomamos $x = p$, sabiendo que $F(p) = 0$, nos queda:

$$0 = F(x_0) + F'(x_0)(p - x_0) + \frac{F''(\xi)}{2!}(p - x_0)^2$$

Si x_0 está suficientemente cerca de p , entonces el último término del segundo miembro en la igualdad anterior será pequeño comparado con los restantes y podemos despreciarlo:

$$0 \approx F(x_0) + F'(x_0)(p - x_0) \implies p - x_0 \approx -\frac{F(x_0)}{F'(x_0)} \implies p \approx x_0 - \frac{F(x_0)}{F'(x_0)}. \quad (6.2)$$

De esta manera podemos llamar

$$x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}$$

y hemos obtenido un nuevo valor x_1 más cercano a p que x_0 . Pensando del mismo modo, podemos escribir

$$p \approx x_1 - \frac{F(x_1)}{F'(x_1)}$$

de manera que

$$x_2 = x_1 - \frac{F(x_1)}{F'(x_1)}$$

es aún una mejor aproximación a p . Continuando de esta manera, queda establecida la regla general (6.1):

$$x_n = x_{n-1} - \frac{F(x_{n-1})}{F'(x_{n-1})} \quad n = 1, 2, \dots$$

Para garantizar la convergencia, debemos darnos cuenta que la iteración de Newton-Raphson es una iteración de punto fijo, por lo cual vale el **Teorema del Punto Fijo**:

$$x_n = x_{n-1} - \underbrace{\frac{F(x_{n-1})}{F'(x_{n-1})}}_{f(x_{n-1})}$$

Por lo tanto, el método será convergente siempre que $|f'(x)| \leq m < 1$:

$$f(x) = x - \frac{F(x)}{F'(x)}$$

$$f'(x) = 1 - \frac{[F'(x)]^2 - F(x)F''(x)}{[F'(x)]^2} = 1 - 1 + \frac{F(x)F''(x)}{[F'(x)]^2} = \frac{F(x)F''(x)}{[F'(x)]^2}.$$

Es decir, el metodo convergerá si:

$$|f'(x)| = \frac{|F(x)F''(x)|}{[F'(x)]^2} \leq m < 1$$

Por hipótesis, sabemos que $F(p) = 0$; luego $f'(p) = 0$. Como $f(x)$ es continua y $f'(p) = 0$, podemos encontrar $\delta > 0$ tal que $|f'(x)| < 1$ se cumple en el intervalo $[p - \delta, p + \delta]$. Por consiguiente, que $x_0 \in [p - \delta, p + \delta]$ es una condición suficiente para que x_0 sea el punto de partida de una sucesión $\{x_n\}_{n=0}^{\infty}$ que converge a la única raíz de $F(x) = 0$ en dicho intervalo, siempre que δ sea elegido tal que:

$$\frac{|F(x)F''(x)|}{[F'(x)]^2} < 1 \quad \forall x \in [p - \delta, p + \delta]$$

Deducción de la fórmula de recurrencia para el Método de Newton-Raphson de 2º Orden

Una modificación al método de N-R se deriva a partir de la utilización de un término más en el desarrollo por serie de Taylor de la función $F(x)$. Aplicamos el **Teorema de Taylor** para la función F en el punto x_0 empleando un polinomio de grado $n = 2$ y el resto en la forma de Lagrange:

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + \frac{F''(x_0)}{2!}(x - x_0)^2 + \frac{F''(\xi)}{3!}(x - x_0)^3$$

donde ξ es un número real entre x y x_0 .

Si tomamos $x = p$, sabiendo que $F(p) = 0$, nos queda:

$$0 = F(x_0) + F'(x_0)(p - x_0) + \frac{F''(x_0)}{2!}(p - x_0)^2 + \frac{F''(\xi)}{3!}(p - x_0)^3$$

Si x_0 está suficientemente cerca de p , entonces el último término del segundo miembro en la igualdad anterior será pequeño comparado con los restantes y podemos despreciarlo:

$$\begin{aligned} 0 &\approx F(x_0) + F'(x_0)(p - x_0) + \frac{F''(x_0)}{2!}(p - x_0)^2 \\ \Rightarrow 0 &\approx F(x_0) + (p - x_0) \left[F'(x_0) + \frac{F''(x_0)}{2!}(p - x_0) \right] \end{aligned}$$

Si dentro de los corchetes se reemplaza $(p - x_0)$ por $-\frac{F(x_0)}{F'(x_0)}$ según la relación vista en (6.2) y se despeja p , se obtiene una expresión para la primera aproximación x_1 , dando lugar a la fórmula iterativa:

$$x_n = x_{n-1} - \frac{F(x_{n-1})F'(x_{n-1})}{[F'(x_{n-1})]^2 - 0,5F(x_{n-1})F''(x_{n-1})} \quad n = 1, 2, \dots$$