

# Similaridad entre Series Temporales y Modelo Logístico Poblacional

---

---

**Autor**

Tomás Lottersberger

---

10 de octubre de 2024

## Resumen

En este informe serán desarrollados y comparados métodos de diferenciación y resolución numérica de ecuaciones diferenciales. Se basará en establecer un análisis detallado sobre por qué el método de diferenciación por diferencias finitas centradas es más eficiente que si se realizan por diferencias finitas hacia adelante o hacia atrás. Además, se buscará detallar y experimentar sobre la eficiencia del método de Runge-Kutta por sobre el método de Euler para la resolución de ecuaciones diferenciales.

## 1. Introducción

Dentro de los métodos numéricos, dos áreas muy amplias se basan en la diferenciación numérica y en la resolución numérica de ecuaciones diferenciales, analizaremos distintos métodos que se encargan de estas dos funcionalidades.

### 1.1. Cálculo de similaridad de series temporales

Consideremos un conjunto de datos  $T_i$  con mediciones de temperatura  $x(t)$  de la región  $i$

$$T_i = (x(t_1), x(t_2), \dots, x(t_n))$$

Estas mediciones se realizan por año, mes y día en distintas ciudades del planeta y se descargan del portal [kaggle.com/datasets/sudalairajkumar/daily-temperature-of-major-cities](https://kaggle.com/datasets/sudalairajkumar/daily-temperature-of-major-cities).

Se analizará un par de series temporales  $(T_i, T_j)$  donde  $i, j$  serán dos ciudades parte de los datos registrados. El objetivo será establecer una comparación no entre los valores de las temperaturas  $x(t)$ , sino en la variación  $\frac{\delta x}{\delta t}$  de las series, es decir, en las derivadas de las funciones que estas provocan.

### 1.2. Resolución de ecuación diferencial

En matemáticas biológicas, se puede representar la dinámica de las poblaciones utilizando

diversas ecuaciones diferenciales. Uno de los modelos más conocidos es el modelo logístico, donde el tamaño de la población en un tiempo  $t$ , denotado como  $N(t)$ , sigue una dinámica descrita por la siguiente ecuación:

$$\frac{dN}{dt} = rN - \frac{rN^2}{K} = rN \frac{K - N}{K} \quad (1)$$

donde  $r$  representa la tasa intrínseca de crecimiento poblacional y  $K$  es la capacidad de carga, es decir, el número máximo de individuos que pueden mantenerse con los recursos disponibles en el sistema.

Una variante posible del modelo logístico incluye el efecto Allee, que toma en cuenta la posibilidad de extinción cuando el número de individuos está por debajo de un cierto umbral, debido a que no se reproducen por la dificultad de encontrar otros de la misma especie. Este modelo establece, por lo tanto, que una población necesita un número mínimo de individuos para sobrevivir y reproducirse con éxito.

La ecuación que describe esta dinámica poblacional se presenta a continuación:

$$\frac{dN}{dt} = rN(1 - \frac{N}{K})(\frac{N}{A} - 1) \quad (2)$$

donde  $r$  es la tasa de crecimiento poblacional intrínseca,  $K$  es el número máximo de individuos que los recursos del sistema pueden sustentar, y  $A$  representa el tamaño mínimo de población necesario para su supervivencia, es decir, el umbral mínimo de supervivencia. Es lógico suponer que  $A$  es menor que  $K$ . Se espera que las dinámi-

cas varíen dependiendo de si la población inicial  $N(t = 0)$  es mayor, igual o menor que  $K$ , y si es mayor, igual o menor que  $A$ .

Dado este modelo, evaluaremos diferentes métodos numéricos utilizados para la resolución de ecuaciones diferenciales y veremos cómo uno de ellos es más conveniente que otro. Además, observaremos distintos resultados según las condiciones iniciales que sean dadas, se experimentará con distintos  $N_0$  y velocidades de tiempo para evaluar cómo varía el comportamiento del modelo con ellos.

## 2. Métodos pre-existentes

### 2.1. Diferenciación numérica

#### 2.1.1. Serie de Taylor

En esencia, la serie de Taylor proporciona un medio para predecir el valor de una función en un punto en términos del valor de la función y sus derivadas en otro punto. En particular, el teorema establece que cualquier función suave puede aproximarse por un polinomio.

$$f_{(x_{i+1})} = f_{(x_i)} + f'_{(x_i)}h + \frac{f''_{(x_i)}h^2}{2!} + \dots + \frac{f^{(n)}_{(x_i)}h^n}{n!} + R_n \quad (3)$$

#### 2.1.2. Diferencias finitas

Es un método que se obtiene de expandir la serie de Taylor de múltiples maneras. Más precisamente, podemos obtener tres formas de diferencias finitas.

Truncando la serie de Taylor (3) después del término con la primera derivada y despejando la derivada, se obtiene:

$$f'_{(x_i)} = \frac{f_{(x_{i+1})} - f_{(x_i)}}{x_{i+1} - x_i} + R_n \quad (4)$$

A esto se le conoce como la **diferencia finita hacia adelante** y a “ $h$ ” se le llama el tamaño del paso o incremento; esto es, la longitud del intervalo sobre el cual se realiza la aproximación. Se le llama diferencia “hacia adelante”, porque usa los datos en  $(i)$  e  $(i + 1)$  para estimar la derivada.

También, la serie de Taylor (3) se puede expandir hacia atrás para calcular un valor anterior sobre la base del valor actual, truncando la ecuación después de la primera derivada y reordenando los términos se obtiene:

$$f'_{(x_i)} = \frac{f_{(x_i)} - f_{(x_{i-1})}}{h} + R_n \quad (5)$$

donde el error es  $O(h)$  y 5 es la **diferencia finita hacia atrás**.

Una tercera forma de reescribir la ecuación 3 es la siguiente:

$$f'_{(x_i)} = \frac{f_{(x_{i+1})} - f_{(x_{i-1})}}{2h} - R_n \quad (6)$$

donde el error es  $O(h^2)$  y 6 se conoce como la **diferencia finita centrada**.

### 2.2. Resolución numérica de ecuaciones diferenciales

#### 2.2.1. Método de Euler

El método de Euler es un procedimiento de integración numérica para resolver ecuaciones diferenciales ordinarias (EDOs) a partir de un valor inicial dado, es el más simple de los métodos numéricos para resolver un problema de valor inicial. Es un método de primer orden, lo que significa que el error local es proporcional al cuadrado del tamaño del paso, y el error global es proporcional al tamaño del paso.

El método de Euler consiste en encontrar iterativamente la solución de una ecuación diferencial de primer orden y valores iniciales conocidos para un rango de valores. Partiendo de un valor

inicial  $x_0$  y avanzando con un paso  $h$ , se pueden obtener los valores de la solución de la siguiente manera:

$$Y(k+1) = Y_k + h * f(x_k, Y_k) \quad (7)$$

donde  $Y$  es solución de la ecuación diferencial y  $f$  es la ecuación diferencial en función de las variables independientes.

### 2.2.2. Método de Runge-Kutta

Uno de los métodos más utilizados para resolver numéricamente problemas de ecuaciones diferenciales ordinarias con condiciones iniciales es el método de Runge-Kutta de cuarto orden, el cual proporciona un pequeño margen de error con respecto a la solución real del problema y es fácilmente programable en un software para realizar las iteraciones necesarias. Y es sumamente útil para casos en los que la solución no puede hallarse por los métodos convencionales (como separación de variables).

El método Runge-Kutta para este problema está dado por la siguiente ecuación:

$$y(i+1) = y_i + \frac{1}{6}[k_1 + 2k_2 + 2k_3 + k_4] \quad (8)$$

Para  $i=0, \dots, n-1$ . La solución se da a lo largo del intervalo  $(x_0, x_0+hn)$ . Donde:

$$\begin{aligned} k_1 &= hf(x_i, y_i) \\ k_2 &= hf(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}) \\ k_3 &= hf(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}) \\ k_4 &= hf(x_i + h, y_i + k_3) \end{aligned}$$

## 3. Métodos propuestos

### 3.1. Medida de similaridad

#### 3.1.1. Errores en diferencias finitas

Para optimizar la manera en la que se aproximará la variación de la función provocada por las temperaturas registradas de una ciudad en un intervalo determinado, primero se analizará qué método exacto de las diferencias finitas es el más adecuado para hacerlo.

Para las diferencias finitas hacia adelante y hacia atrás el error puede expresarse específicamente con la siguiente ecuación:

$$error = |\frac{h^2}{3}f'''(x_0)|$$

Mientras que aproximar por el método de diferencias finitas centradas podemos obtener un error expresado por la siguiente ecuación:

$$error = |\frac{h^2}{6}f'''(x_0)|$$

Dadas estas ecuaciones y teniendo en cuenta que  $|\frac{h^2}{6}f'''(x_0)| \leq |\frac{h^2}{3}f'''(x_0)|$  entonces se puede determinar con seguridad que el error con diferencias finitas centradas es menor que el error cometido al aproximar con diferencias finitas hacia adelante o hacia atrás.

Para demostrar la diferencia que ocurre cuando se aproxima se observará en la figura 1 con la serie temporal  $T_i$  la cual grafica la variación ocurrida en las temperaturas en Buenos Aires durante octubre del 2019.

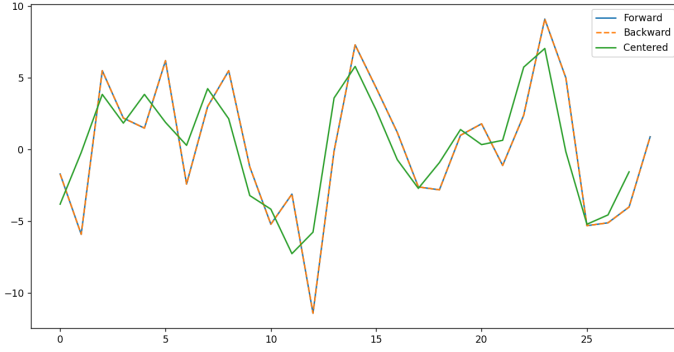


Figura 1: Comparación de la variación de la serie  $T_i$  aproximando mediante diferencias finitas hacia adelante, hacia atrás y centradas respectivamente.

### 3.1.2. Proposición de aproximación

Habiendo analizado la consistencia del método de diferencias finitas en sus tres formas, determinamos que el error cometido en la forma centrada es menor a los dos restantes, por lo que éste se utilizará para aproximar las derivadas de las distintas series temporales nombradas en el punto 1.1.

## 3.2. Elección de método numérico para resolución de ecuación diferencial

Para seleccionar un método de resolución de ODE's adecuado comenzaremos por buscar la solución analítica de la ecuación diferencial 1 y comparar dicha solución con la solución brindada por los métodos de Euler y Runge-Kutta.

### 3.2.1. Resolución analítica

La ecuación 1 se puede reescribir como:

$$\frac{1}{N(1 - \frac{N}{K})} dN = r dt$$

Simplificando el lado izquierdo descomponiendo en fracciones parciales obtenemos:

$$(\frac{1}{N} + \frac{1}{K - N}) dN = r dt$$

Luego, integrando ambos lados de la ecuación:

$$\ln(\frac{N}{K - N}) = rt + C$$

Despejando  $N(t)$ :

$$N(t) = \frac{K e^C e^{rt}}{1 + e^C e^{rt}}$$

Usando la condición inicial  $N(0) = N_0$ , se obtiene:

$$e^C = \frac{N_0}{K - N_0}$$

Finalmente, la solución analítica del modelo logístico es:

$$N(t) = \frac{K N_0 e^{rt}}{K + N_0 (e^{rt} - 1)} \quad (9)$$

### 3.2.2. Comparación con métodos

Utilizando como 'ground truth' los resultados obtenidos de la resolución analítica (9) en diferentes valores de tiempo, evaluaremos los resultados obtenidos utilizando los métodos de Euler y Runge-Kutta para la ecuación diferencial original.

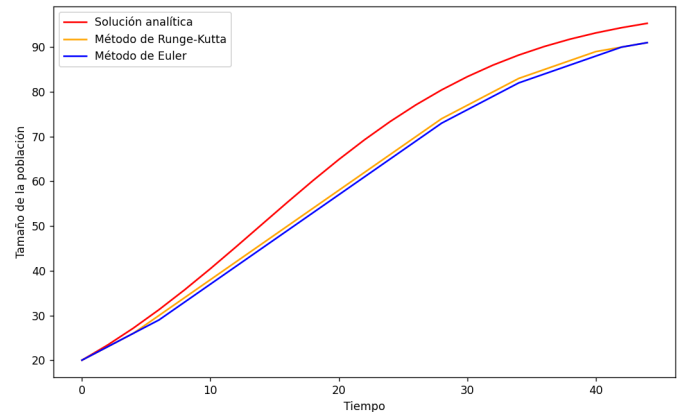


Figura 2: Comparación de error de los métodos de Euler y Runge-Kutta respecto a la resolución analítica de la ecuación diferencial.

En la figura 2 se puede observar que, si bien hay cierto error en los métodos numéricos utilizados, ambos tienen una función acorde al crecimiento verdadero. Además, el método de Runge-Kutta es notablemente más parecido al ground truth ya que éste posee un error menor al método de Euler.

## 4. Experimentación

En esta sección, se presentarán los resultados obtenidos de analizar los diferentes métodos numéricos propuestos en los estudios de caso que se plantearon al comienzo de este informe.

### 4.1. Cálculo de similaridad entre series temporales

Se pondrá en práctica la aproximación de derivadas determinada en el punto 3.1.2. para dos series temporales  $(T_i, T_j)$  donde  $T_i, T_j$  corresponderán a las temperaturas registradas en las ciudades  $i, j$ , para ello se propondrán distintos  $i, j$  de experimentación.

#### 4.1.1. Ciudades del mismo hemisferio

En este caso de estudio,  $i, j$  corresponderán a dos ciudades ubicadas en el mismo hemisferio, más específicamente, se ha propuesto a  $i$  como la ciudad de *Seúl (Corea del Sur)* en el intervalo de años entre 2017 y 2019 y a  $j$  como la ciudad de *Moscú (Rusia)* en el mismo intervalo. Luego de utilizar el método propuesto para aproximar la variación (véase la derivada) de la serie, se calculó la siguiente similaridad entre ambas:

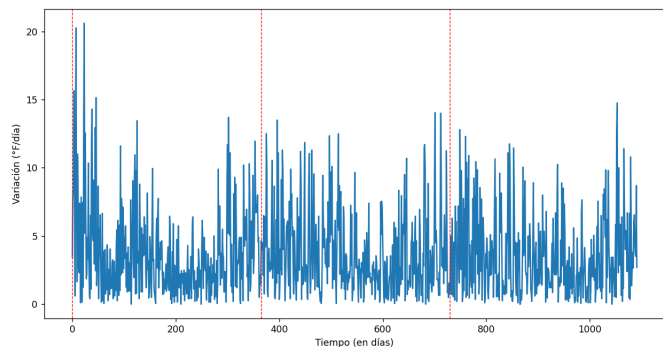


Figura 3: Cálculo de similaridad entre las variaciones de las series temporales  $(T_i, T_j)$  con  $(i, j) = (\text{Seúl}, \text{Moscú})$ .

Teniendo en cuenta que la línea vertical de color rojo en la figura 3 marca un cambio del año corriente, se puede observar un patrón en las variaciones de ambas series temporales, donde éstas varían de manera diferente en los extremos y en la época de mayo/abril de los años, mientras que en el resto se mantiene una variación más lineal entre ambas. Los motivos de dichas diferencias o similaridades exceden el contenido del presente informe.

#### 4.1.2. Ciudades de hemisferios opuestos

Con el objetivo de hacer un cálculo diferente, se ha propuesto que  $(T_i, T_j)$  sean, en principio, series de características menos similares que las propuestas en el punto 4.1.1., por ésto,  $i$  seguirá representando a la ciudad de *Seúl (Corea del Sur)*, pero  $j$  ahora representará a la ciudad de *Buenos Aires (Argentina)*, ubicada en el hemisferio opuesto a  $i$ . Los intervalos de tiempo evaluados permanecen iguales.

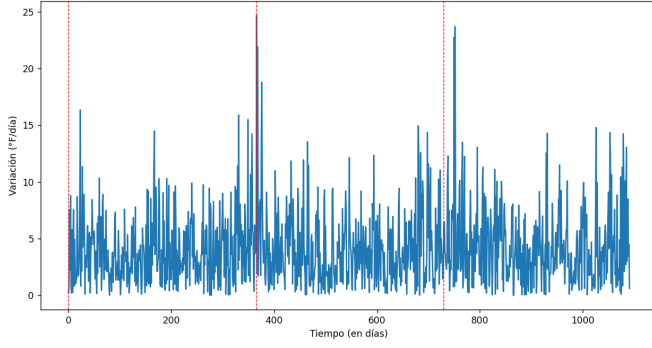


Figura 4: Cálculo de similaridad entre las variaciones de las series temporales  $(T_i, T_j)$  con  $(i, j) = (\text{Seúl}, \text{Buenos Aires})$ .

Con las nuevas ciudades propuestas, observamos en la figura 4 que las principales variaciones entre ellas se encuentran en exactamente los principios de los años, y, en menor medida, en las mitades de ellos.

#### 4.1.3. Misma ciudad, diferentes años

Para finalizar con un tercer cálculo de similaridad diferente, se ha propuesto analizar la variación en la temperatura de una ciudad pero en años diferentes, para esto, ambas series pertenecerán a la ciudad de *Seúl (Corea del Sur)* y las variables  $i, j$  ahora representarán los años analizados, los cuales han sido elegidos arbitrariamente.

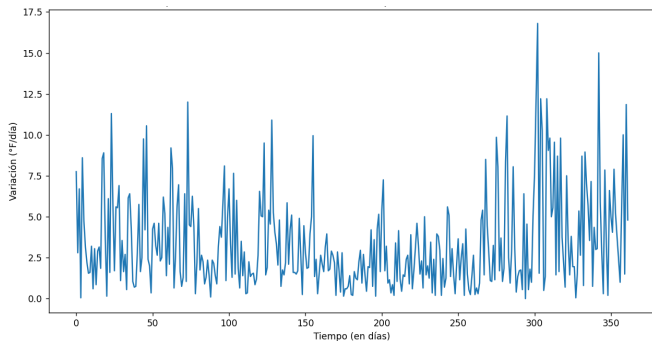


Figura 5: Cálculo de similaridad entre las variaciones de las series temporales  $(T_i, T_j)$  correspondientes a las temperaturas de Seúl con  $(i, j) = (1995, 2017)$ .

En los nuevos términos, se puede observar el cambio que ha habido en las variaciones de las temperaturas de Seúl luego de 22 años, donde vemos que en las fechas correspondientes al invierno en Corea del Sur (finales de noviembre, diciembre, enero y febrero) se ha producido un gran cambio en la manera en la que varían las temperaturas.

## 4.2. Resolución de ecuación diferencial

Dado que el modelo de ecuación 2 no tiene una solución analítica directa ya que involucra dos términos no lineales (N y A) interactuando de manera más compleja, solo se puede resolver numéricamente con métodos como el de Euler y Runge-Kutta. Por este motivo utilizaremos ambos métodos para tener una idea de cómo se comporta este modelo, y cómo reacciona ante diferentes condiciones iniciales.

### 4.2.1. Variación de métodos respecto al paso del tiempo $dt$

Se sabe que el método de Euler es más preciso cuanto menor es el paso en el tiempo  $dt$  de una muestra a otra, con la variante 2 del modelo logístico poblacional se experimentará con distintos  $dt$  para ver cómo afecta en la tasa de crecimiento y qué tanto se parece a los resultados del método de Runge-Kutta, el cual tiene un error menor.

Para este análisis tomaremos las siguientes condiciones iniciales fijas:

- K = 100: Capacidad máxima de supervivencia.
- A = 10: Umbral mínimo de supervivencia.
- r = 0.1: Tasa de crecimiento poblacional.
- $N_0 = 20$ : Semilla inicial de personas vivas.

Y la siguiente condición inicial variable:

$dt = (2, 1, 0.1)$ : Paso del tiempo hasta la muestra  $N(t+1)$

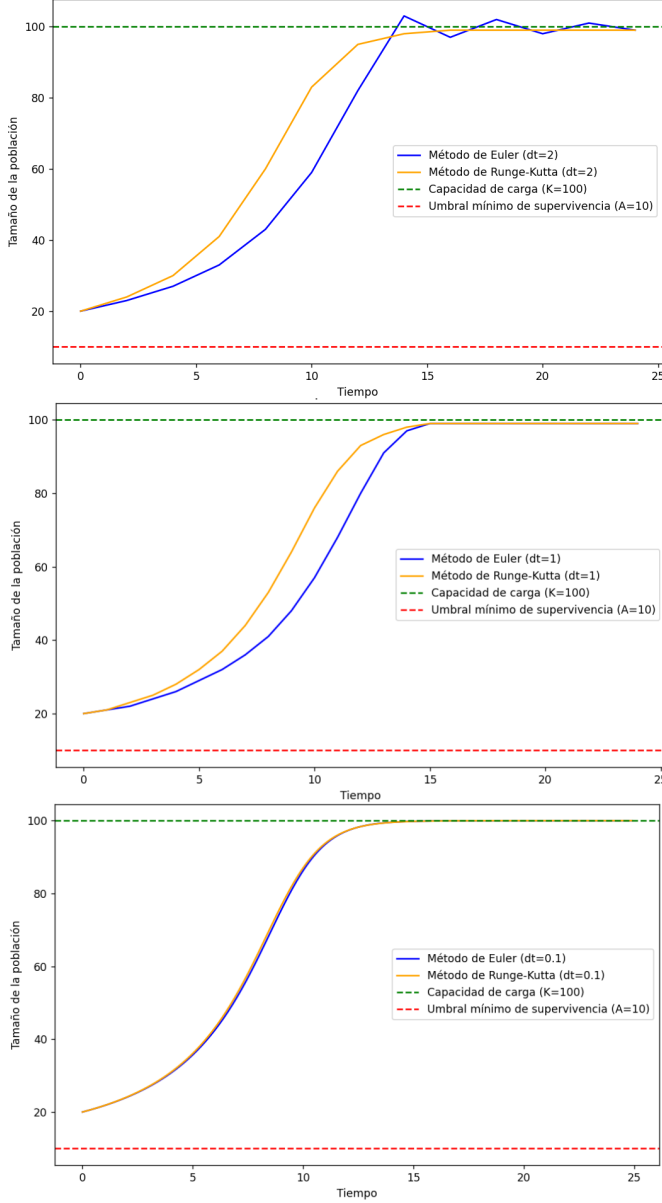


Figura 6: Comparación de los resultados arrojados por los métodos de Euler y Runge-Kutta para  $dt = (2, 1, 0.1)$ .

En la figura 6 se puede observar que, con la semilla inicial  $N_0 = 20$ , la población crecerá exponencialmente hasta la capacidad máxima de

personas con vida que puede haber, esto se deduce de que, cuando  $N_0 \geq A$  las personas con vida encuentran a otras de su misma especie para reproducirse y, cuantas más haya, más rápido crecerá la población.

Además se demuestra lo dicho con anterioridad, que el método de Euler será más preciso cuanto más chico sea el paso del tiempo  $dt$  con respecto a la próxima muestra, se puede observar que con un paso de tiempo  $dt = 2$  los resultados distan mucho más a los arrojados por el método de Runge-Kutta, incluso sobrepasando la capacidad máxima que el modelo debería permitir, mientras que cuando disminuimos el valor de  $dt$  los resultados son más acertados hasta arrojar unos idénticos a los de Runge-Kutta.

#### 4.2.2. Efecto Allee

El efecto Allee, así llamado por W.C. Allee, se produce cuando, a partir de cierto umbral, el tamaño poblacional de una especie es tan reducido que la tasa de supervivencia y/o la tasa reproductiva desciende debido a que los individuos no se reproducen al no encontrarse con más individuos de la misma población. Aunque el efecto Allee se produce de manera individual, afecta a toda la población debido a que no hay reclutamiento de nuevos individuos.

#### 4.2.3. Variación de crecimiento respecto a la semilla inicial $N_0$

Debido a la teoría del efecto Allee, se quiere ver cómo éste repercute en nuestra variante del modelo logístico poblacional, se usará el método de Runge-Kutta para el caso de estudio.

Para este análisis tomaremos las siguientes condiciones iniciales fijas:

$K = 100$ : Capacidad máxima de supervivencia.

$A = 10$ : Umbral mínimo de supervivencia.

$r = 0.1$ : Tasa de crecimiento poblacional.



$dt = 2$ : Paso del tiempo hasta la muestra  
 $N(t + 1)$

Y la siguiente condición inicial variable:

$N_0 = (5, 20, 110)$ : Semilla inicial de personas  
vivas.

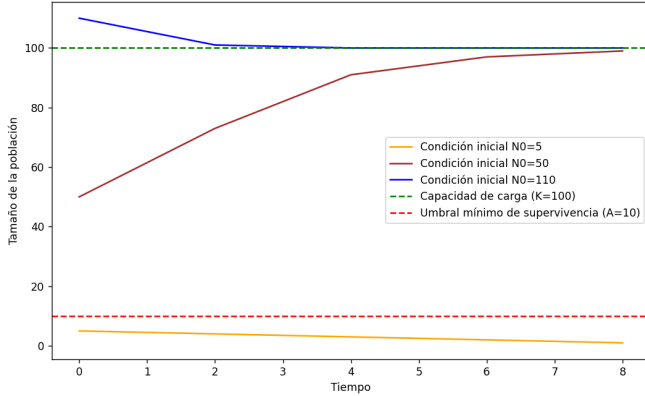


Figura 7: Comparación del crecimiento poblacional con  $N_0 = (5, 20, 110)$

En la figura 7 se puede observar que el efecto Allee se hace notar, al haber considerado el umbral mínimo de supervivencia en  $A=10$  cualquier  $N_0$  que sea menor que  $A$  hará que la población no pueda crecer y terminará por extinguirse.

Por otro lado, cuando se toma un  $N_0$  de manera que  $A \leq N_0 < K$  se tiene que la población sí podrá crecer, cuanto mayor sea  $N(t)$  entonces crecerá más rápido.

Finalmente se tiene que si  $N_0$  es mayor que  $K$ , al sobrepasarse la capacidad máxima de personas que pueden permanecer con vida en la población, habrá un decrecimiento en la población hasta que éste se ajuste exactamente en el límite  $N(t)=K$ .

## 5. Conclusiones

En este estudio, se examinaron en detalle diversas técnicas y métodos de diferenciación

numérica y técnicas para la resolución de ecuaciones diferenciales. Mediante un análisis meticuloso y una serie de experimentos se logró un entendimiento sólido de las fortalezas, limitaciones y aplicaciones de cada enfoque en diferentes contextos prácticos. Uno de los puntos más relevantes fue la comprensión de por qué el método de diferenciación mediante diferencias finitas centradas es más conveniente que el de diferencias finitas hacia adelante o hacia atrás, poniéndolo en práctica para aproximar derivadas de series temporales. Luego, dichas aproximaciones fueron usadas para calcular una similaridad entre ellas con el ejemplo de las temperaturas en las ciudades, donde se observaron patrones que podrían ser indicios de un motivo para que ocurran.

También se evaluó la resolución de ecuaciones diferenciales mediante los métodos numéricos de Euler y Runge-Kutta, donde se observó que éste último es el más eficiente a la hora de considerar un balance de resultados/costo computacional, aunque el método de Euler no arroja resultados tan alejados de los esperados, donde además estos son más acertados cuanto más pequeño es el paso del tiempo  $dt$ .

Finalmente se consideró el método de Runge-Kutta para analizar distintas condiciones iniciales en los  $N_0$  a evaluar, haciéndolo con tres  $N_0$  arbitrarios diferentes obtuvimos resultados dispersos, donde se estableció que un criterio para que una población pueda crecer se necesitará que la población inicial sea mayor o igual al umbral mínimo de supervivencia establecido, de otro modo, ésta será incapaz de sobrevivir debido a la falta de individuos que puedan reproducirse.

## 6. Bibliografía

1. Burden, R. L., Faires, J. D., & Burden, A. M. (2015). Numerical analysis.
2. Diferencias Finitas. Universidad de Cantabria. <https://www.giematic.unican.es/funciones-de-una-variable-segunda-parte-2/diferencias-finitas/>
3. Método de Euler. Wikipedia, La enciclopedia libre. [https://es.wikipedia.org/wiki/M%C3%A9todo\\_de\\_Euler](https://es.wikipedia.org/wiki/M%C3%A9todo_de_Euler)
4. Método de Runge-Kutta. Wikipedia, La enciclopedia libre. [https://es.wikipedia.org/wiki/M%C3%A9todo\\_de\\_Runge-Kutta](https://es.wikipedia.org/wiki/M%C3%A9todo_de_Runge-Kutta)
5. Efecto Allee. Wikipedia, La enciclopedia libre. [https://es.wikipedia.org/wiki/Efecto\\_Alee](https://es.wikipedia.org/wiki/Efecto_Alee)