Pontificia Universidad Javeriana de Cali

Ingeniería de sistemas y computación

Interpolación

Laboratorio Computación Cientifica

Autores:
Andrés Felipe Delgado y Ana María García

Abril 24 del 2020

1 Introducción

En el presente laboratorio se pondrán en práctica los conceptos vistos en clase correspondientes a la interpolación.

La interpolacion consiste en ajustar una función a unos datos tal que la función sea equivalente a dichos datos

El problema más sencillo de interpolación en una dimensión es de la siguiente forma:

Dado un conjunto de datos: $(t_i, y_i), i = 1, .., n$ con $t_1 < t_2 < ... < t_n$
Se busca una funcion tal que

$f(t_i) = y_i, i = 1, ..., n$

En general se puede expresar como

$f(t_i) = \sum_{j=1}^{n} x_j \phi_j(t_i) = y_i$

Donde ϕ_j son las funciones base y f es la funcion interpolante.

Se implementaron tres tipos de interpolacion en el lenguaje de programacion Matlab.
El primero es interpolacion Polinomial con Monomios,el segundo es interpolacion de Newton y finalmente interpolacion lineal a trozos.

En el presente documento primero se especificará la función interpolante con sus respectivas funciones base para cada tipo de interpolación. Luego se presentará os pequeñas bases de datos con las que se puso a prueba las interpolaciones. Se presentarán los resultados y se concluirá las ventajas y desventajas que tiene cada tipo de iterpolación respecto al resto.

2 Interpolación Polinomial con Monomios

Es la interpolación más sencilla que emplea polinomios.

Existe un unico polinomio de grado n-1 que pasa por todos los puntos. A este polinomio se le conoce como el polinomio interpolante y tiene la siguiente forma:

$p_{n-1} = x_1 + x_2 t + x_3 t^2 + ... + x_n t^{n-1}$

Note que en este caso la funcion base es monomial y tiene la forma:

$\phi_j(t) = t^{j-1}, j = 1...n$

Para construir el polinomio interpolante es necesario conocer los coeficientes x_i que estan determinados por el siguiente sistema lineal:

$$\begin{Bmatrix} 1 & t_1 & ... & t_1^{n-1} \\ 1 & t_2 & ... & t_2^{n-1} \\ ... & ... & ... & ... \\ 1 & t_n & ... & t_n^{n-1} \end{Bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{Bmatrix} = \begin{Bmatrix} y_1 \\ y_2 \\ ... \\ y_n \end{Bmatrix}$$

Este sistema se puede resolver como el de minimos cuadrados, pero limitado sólo al caso donde siempre existe el mismo número de incógnitas y de ecuaciones. Este fue expuesto a profundidad en el anterior laboratorio por medio de Cholesky y eliminación Gaussiana. En este laboratorio no se volverá a detallar la forma en que se soluciona este sistema.

Como tambien se expuso en el anterior laboratorio, mientras se resuelve el sistema lineal la matriz A que se genera es amenudo mal condicionada, sobre todo para un orden alto, ya que las funciones son menos distinguibles conforme el grado del polinomio aumenta, lo cual va haciendo las columnas de A linealmente independientes lo que eventualmente genera un sobre-ajuste.

Cuando mostremos los resultados, se indicara un ejemplo de sobre-ajuste.

La complejidad del algoritmo es la de la solucion del sistema lineal, es decir O(n^2), donde n es el número de datos que ingresan como parámetro.

3 Interpolación de Newton

Para un conjunto de datos $(t_i, y_i), i = 1...n$, el polinomio interpolante de Newton tiene la forma:

$$p_{n-1} = x_1 + x_2(t - t_1) + x_3(t - t_1)(t - t_2) + ... + x_n(t - t_1)(t - t_2)...(t - t_{n-1})$$

Por lo tanto, las funciones base de Newton tienen la forma:

$$\phi_j(t) = \prod_{k=1}^{j-1}(t - t_k), j = 1, ..., n$$

Lo cual forma una matriz A triangular inferior, por lo que los coeficientes x_i, i=1...n se pueden hallar mediante sustitucion sucesiva hacia adelante O(n^2).

El metodo de Newton tiene un mejor balance entre el costo computacional del interpolante y el costo de evaluar los datos, en comparación con la interpolación polinomial con monomios. Esto se podrá evidenciar en la sección de resultados del presente documento.

4 Interpolación Polinomial a trozos

La interpolación a trozos surge como una alternativa que tiene como característica principal el poder ajustar un número grande de datos con polinomios de grado bajo, lo que mitiga el problema de sobre-ajuste.

Existen diversas interpolaciones a trozos, pero en esta practica nos centraremos en la más sencilla que es la lineal.

Dado un conjunto de datos (t_i, y_i) se emplea un polinomio distinto entre cada subintervalo $[t_i, t_{i+1}]$.

5 Resultados

Para calcular los resultados de los casos explicados más abajo, se utilizaron 22 datos en la interpolacion polinomial con monomios, pues a partir de los 26 datos, como se mencionó anteriormente, las columnas comienzan a convertirse en linealmente independientes. Por lo cual, la función de Cholesky no puede trabajar.

Por esta razon se hicieron dos tipos de pruebas: primero se probaron los 3 algoritmos con 22 datos para poder comparar entre todos y posteriormente se hizo la prueba con los 40 datos completos en la interpolacion de Newton y la interpolacion polinomial a trozos.

NOTA 1: En todos los graficos de resultados que se van a presentrar posteriormente, la linea de color rojo representa la interpolacion y la linea de color azul es la grafica generada con los datos reales.

NOTA 2: El error es calculado en el código.

5.1 Bases de datos

Para llevar a cabo un mejor análisis y comprobación de la correctitud de los algoritmos, se utilizaron dos bases de datos de ejemplo.

En cada conjunto se toma una cantidad de 40 datos donde cada dato indica un valor correspondiente a un año los cuales se representan en el eje x, mientras que en el eje y aparecen los datos de los conjuntos correspondientes.

5.1.1 Población Colombiana

El conjunto de datos fue extraído del siguiente sitio:
https://datosmacro.expansion.com/demografia/poblacion/colombia

El primer conjunto de datos con el que se hicieron las pruebas es un conjunto que indica el crecimiento de la población colombiana iniciando desde el año 1978 y finalizando en el año 2018. El comportamiento de los datos se muestra en la siguiente gráfica:

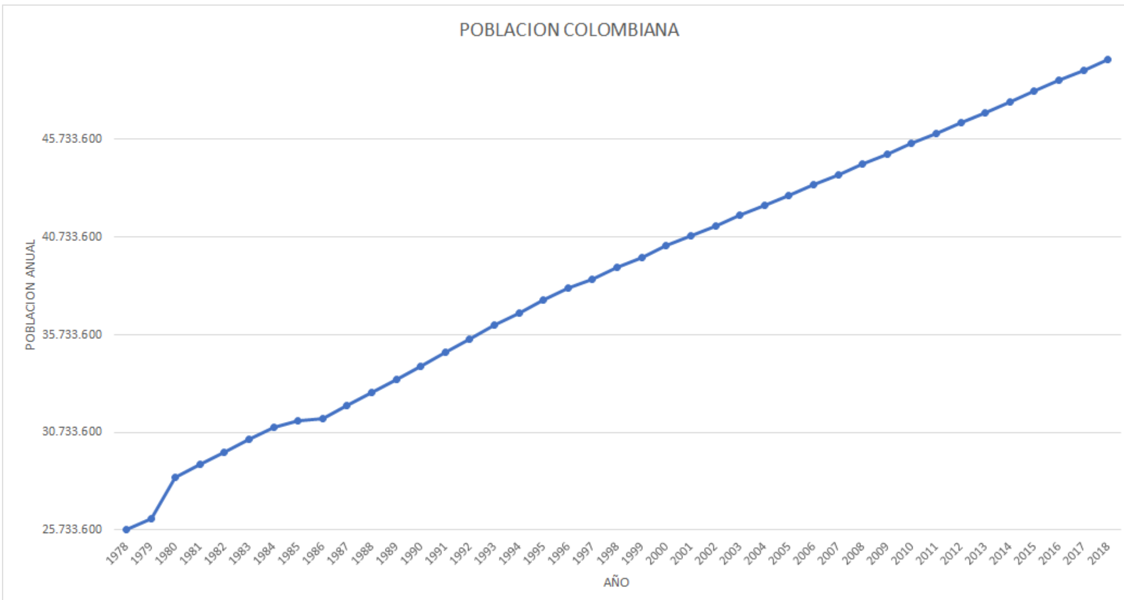

Figure 1: Poblacion Colombiana

Para transformar estos datos como entrada al programa en Matlab, se tuvieron que hacer ciertas modificaciones: los datos que representan los años con 1978-2018 se redujeron al rango 1-40. Además de esto, todos los datos del tamaño de la población dividieron entre un millón para bajar el rango en el que trabaja el programa. Estos cambios son equivalentes al conjunto real mostrado en la gráfica anterior.

INTERPOLACIÓN POLINOMIAL CON MONOMIOS:

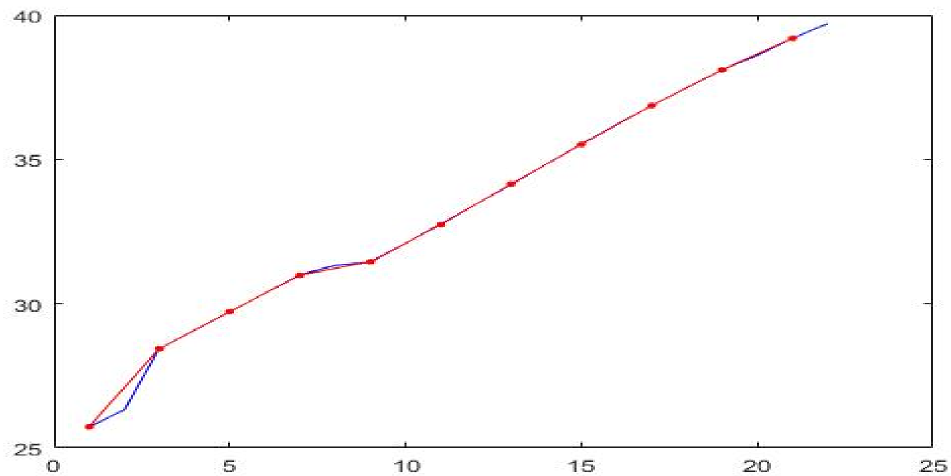

Figure 2: Interpolación Polinomial con monomios

INTERPOLACIÓN DE NEWTON:

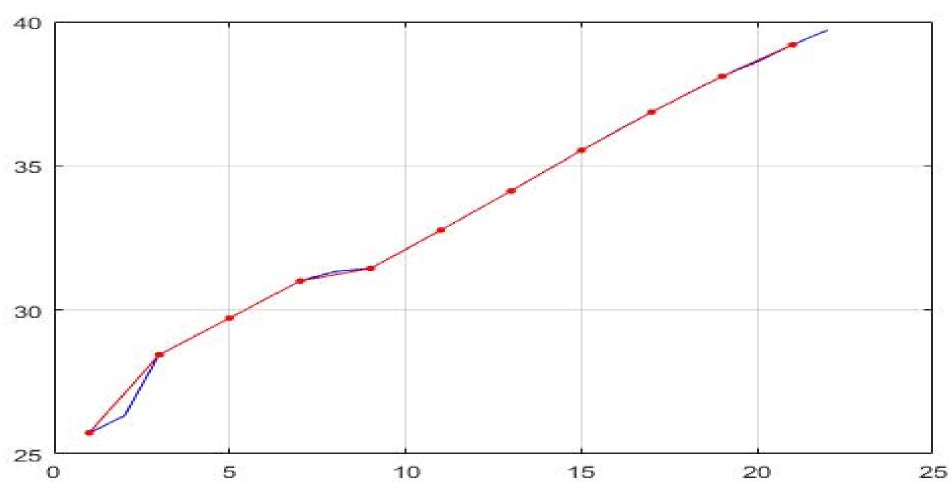

Figure 3: Interpolacion de Newton - 22 datos

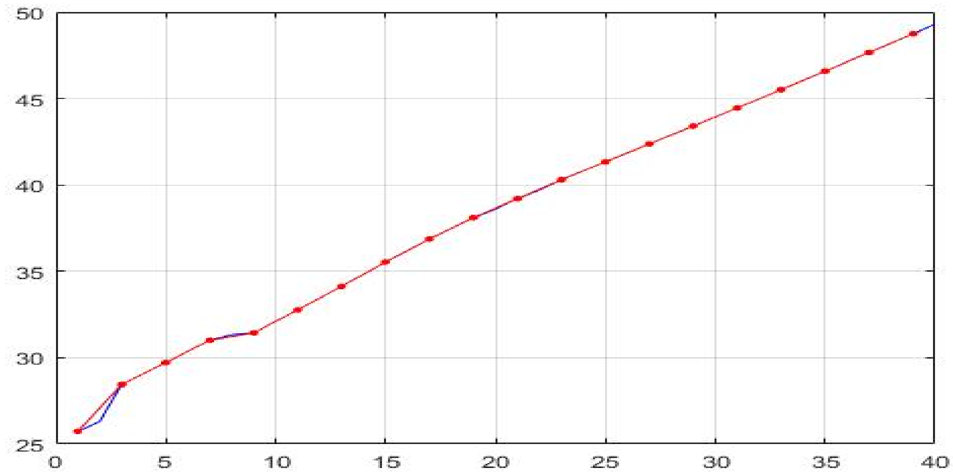

Figure 4: Interpolacion de Newton - 40 datos

INTERPOLACION POLINOMIAL A TROZOS:

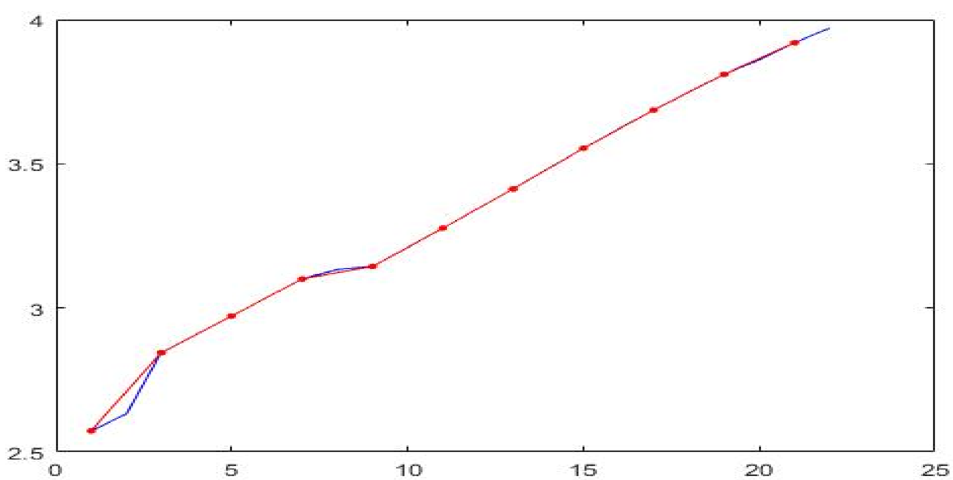

Figure 5: Interpolacion polinomial a trozos - 22 datos

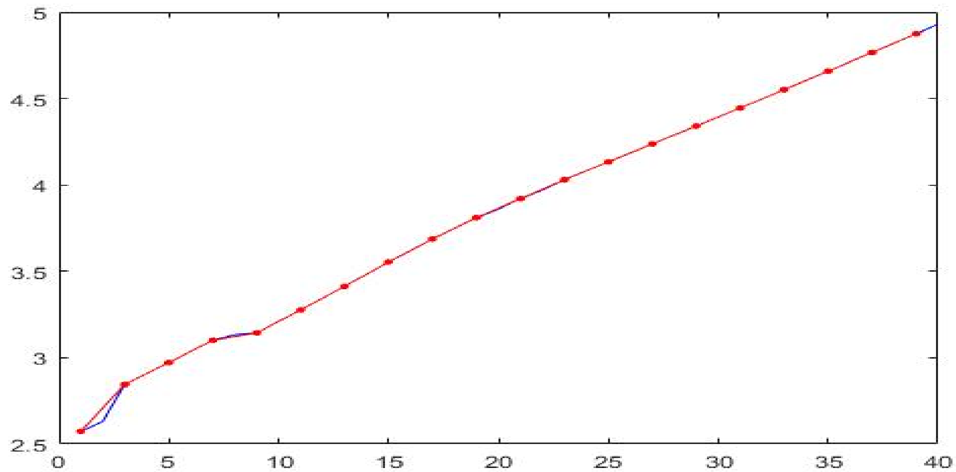

Figure 6: Interpolacion polinomial a trozos - 40 datos

5.1.2 Valor del dolar en Colombia

El conjunto de datos fue extraído del siguiente sitio:
https://dolar.wilkinsonpc.com.co/dolar-historico/

El segundo conjunto de datos con el que se hicieron las pruebas es un conjunto que indica el comportamiento del valor del dolar en Colombia iniciando desde el año 1979 y finalizando en el año 2019. El comportamiento de los datos se muestra en la siguiente gráfica:

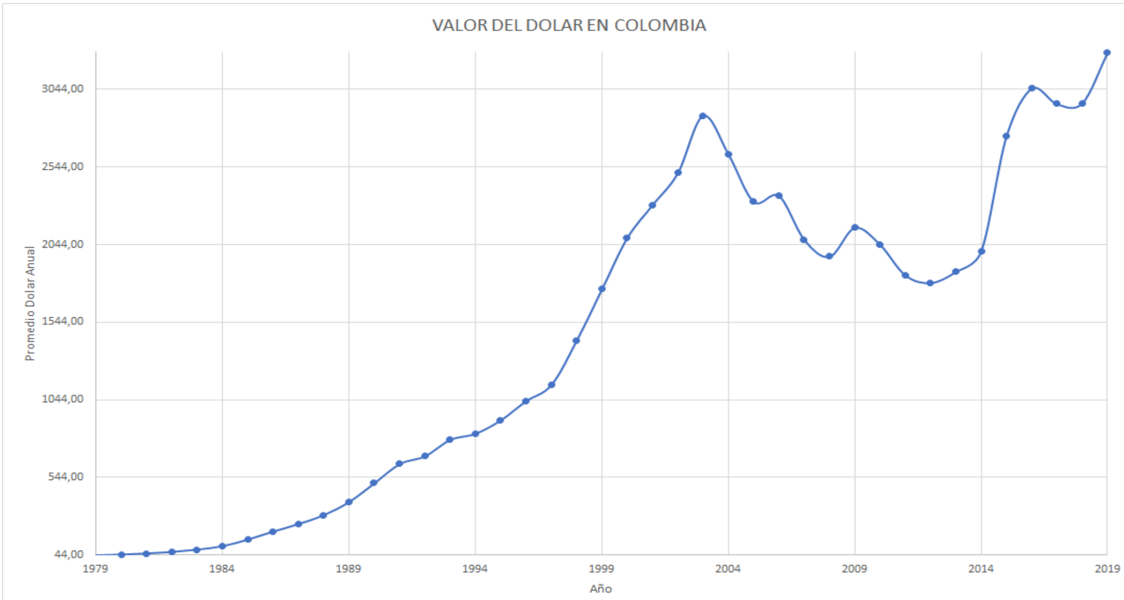

Figure 7: Valor Dolar Colombia

Al igual que en el punto anterior, para transformar estos datos como entrada al programa en Matlab, se tuvo que modificar los datos que representan los años con 1979-2019 reduciendolos al rango 1-40.

INTERPOLACIÓN POLINOMIAL CON MONOMIOS:

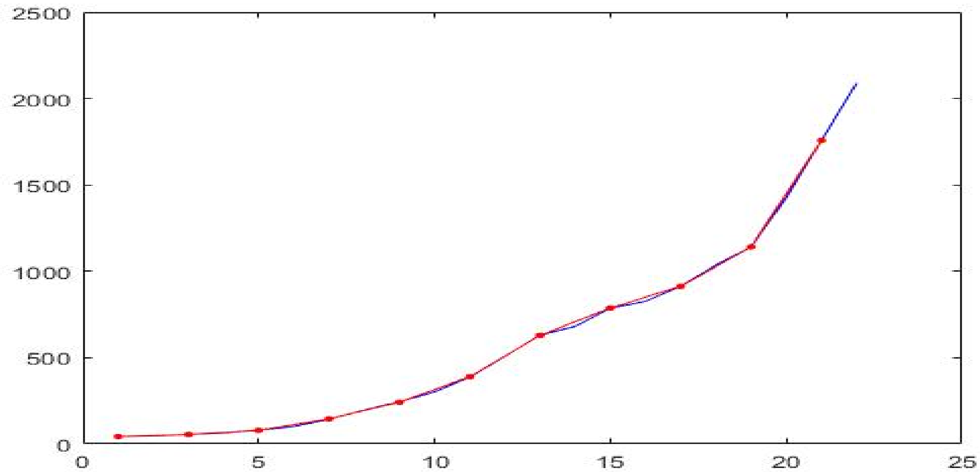

Figure 8: Interpolacion polinomial con monomios

INTERPOLACIÓN DE NEWTON:

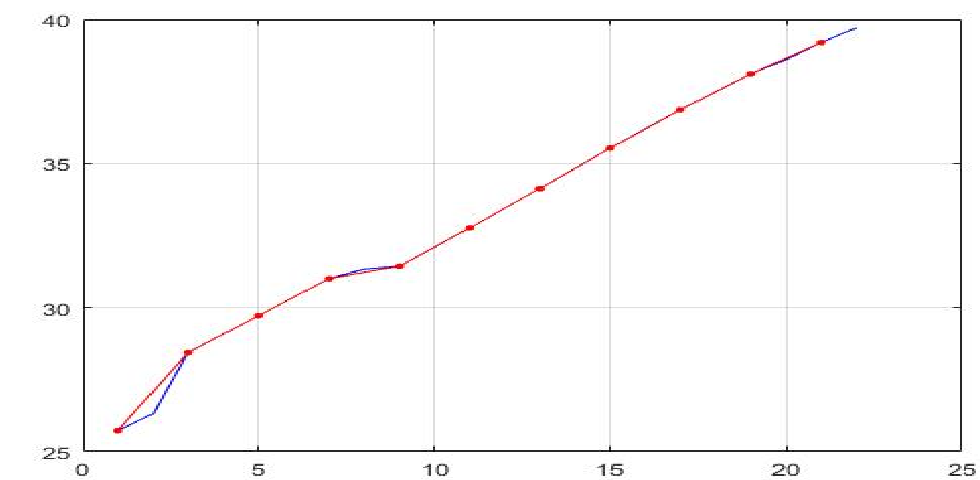

Figure 9: Interpolacion de Newton - 22 datos

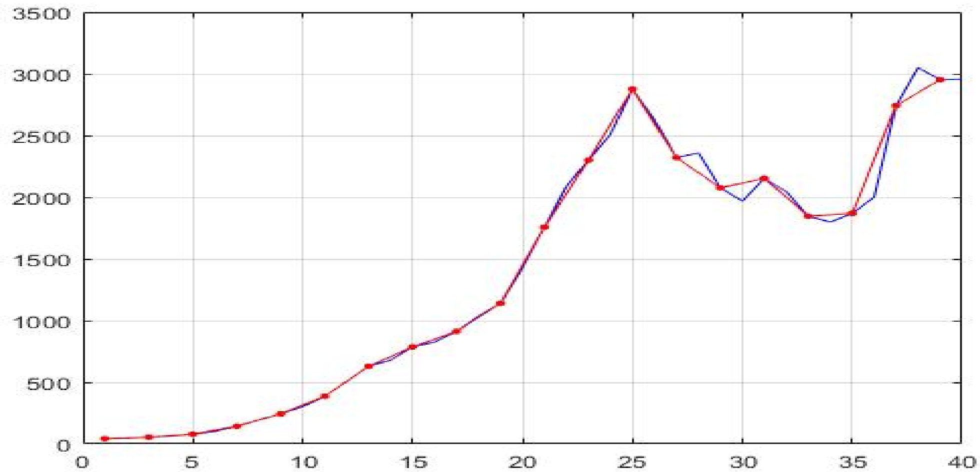

Figure 10: Interpolacion de Newton - 40 datos

INTERPOLACION POLINOMIAL A TROZOS:

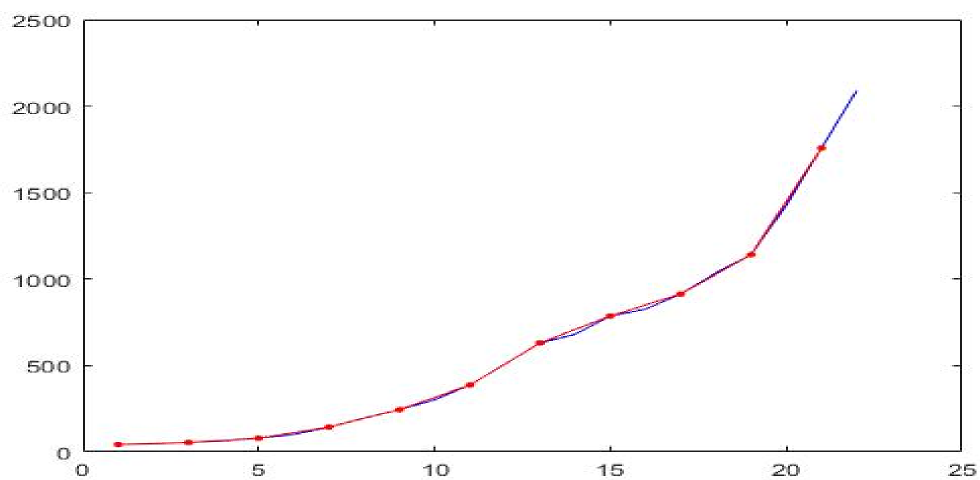

Figure 11: Interpolacion polinomial a trozos - 22 datos

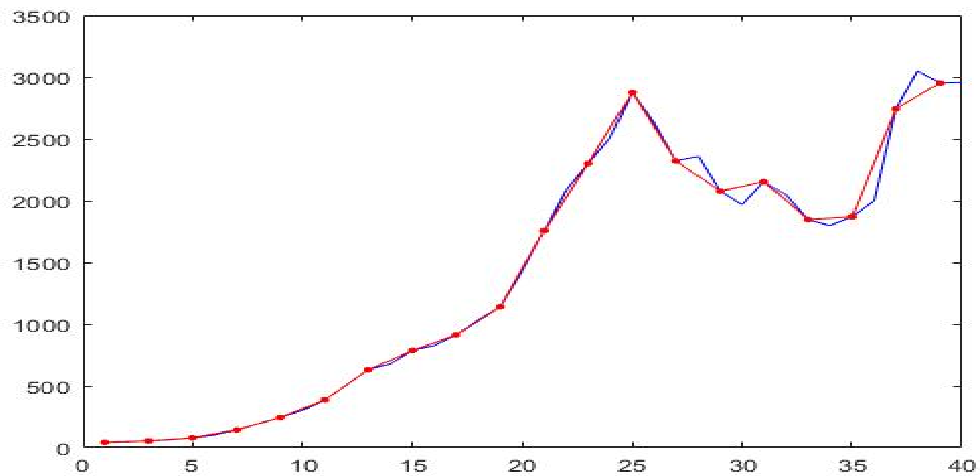

Figure 12: Interpolacion polinomial a trozos - 40 datos

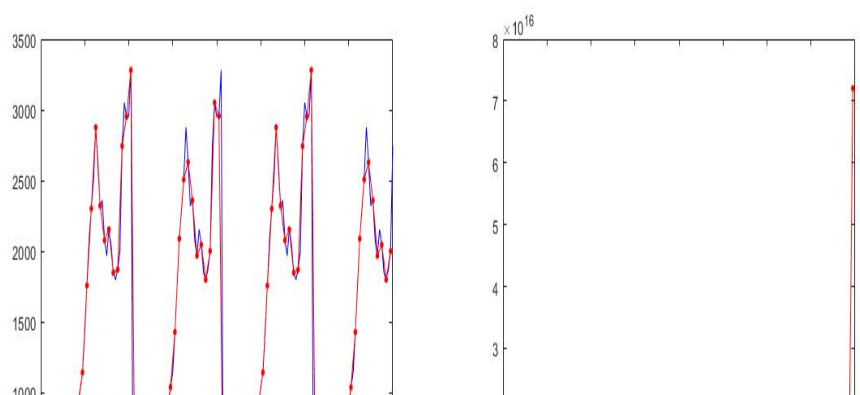

Figure 13: COMPARACIÓN: No sobre-ajuste en interpolación polinomial a trozos VS Sobre-ajuste en interpolación de Newton

6 Conclusiones

Despues de haber analizado a profundidad el funcionamiento de estos tres distintos métodos podemos obtener varias conclusiones:

- La interpolacion polinomial con monomios está muy limitada en cuanto a la cantidad de datos que puede interpolar en comparación con los otros dosmétodos. Sin embargo, para pequeñas bases de datos tiene una buena precisión y computacionalmente no es tán costosa.
- La interpolacion de Newton puede ajustar muchos más datos que la monomial, pero aún así tiene un límite como podemos ver en la gráfica ()SOBREAJUSTE. Se hizo una prueba con 160 datos y se presentó un sobre-ajuste. Computacionalente hablando, esta interpolacion suelen ser muy costosa conforme crece el número de datos, pues el número de multiplicaciones a realizar crece exponencialmente. Aunque en menor medida que en La Grange igual se requiere un gran número de operaciones para calcular el polinomio de interpolación.
- La interpolacion polinomial a trozos representa una mejora importante con respecto a las dos formas de interpolar anteriores, pues se basa en operaciones sencillas de bajo grado además de que por este mismo motivo no le afecta el sobre-ajuste, como lo podemos ver en la gráfica donde, con los mismos 160 datos, la interpolación a trozos se hizo correctamente, mientras que la de Newton sufrió un sobre-ajuste.