

Reto 2 – Interpolación en variables climáticas

José Mario Arias(ariasjosem@javeriana.edu.co), Lida Guerrero(lida.guerrero@javeriana.edu.co), Juan Duarte(duarte-juan@javeriana.edu.co), Santiago Diaz(diaz.sandres@javeriana.edu.co), Camilo Cruz (cruz_camilo@javeriana.edu.co)

Abstract- In this paper, a study of the behavior of climatic variables from different meteorological stations in Brazil will be presented. An analysis will be made thanks to different interpolation methods in addition to the error presented.

Resumen-En el presente documento se encontrará un estudio del comportamiento de las variables climáticas de diferentes estaciones meteorológicas en Brasil. Se hará un análisis gracias a distintos métodos de interpolación además del error que se presente.

I. Introducción

El proceso de predicción de datos por medio de una función se conoce como interpolación, para estos cálculos matemáticos contamos con diferentes métodos que permiten realizar un ajuste de curva, en este caso se implementara un ajuste para datos referentes al clima, utilizando las herramientas que ofrece el programa RStudio.

II. Definición del problema

En muchas circunstancias, no se cuenta con un sistema que este midiendo constantemente las diferentes variables de clima, por lo que se recurre a predicciones por medio de datos anteriores, para estas predicciones se debe encontrar el método que más se ajuste a un valor correcto.

Se cuenta con datos referentes al clima de Brasil en el mes de abril de 2013, se busca ajustar una curva para los valores faltantes por medio de interpolaciones, en este caso utilizaremos los valores de temperatura y humedad relativa del aire.

III. Herramientas y librerías

Se usará la herramienta Rstudio ya que viene siendo la herramienta más trabajada del curso y más sencilla,

además permite una fácil manipulación de tablas y archivos .xlsx debido a que los datos de estudio se encontraban en un archivo de Excel. Dentro de las librerías usadas las cuales las provee Rstudio existen funciones las cuales implementan los diferentes algoritmos de interpolación. Las funciones usadas fueron spline() la cual usa el spline cubico, splinefun() la cual usa la interpolación de Hermite y aprox() la cual usa interpolación lineal.

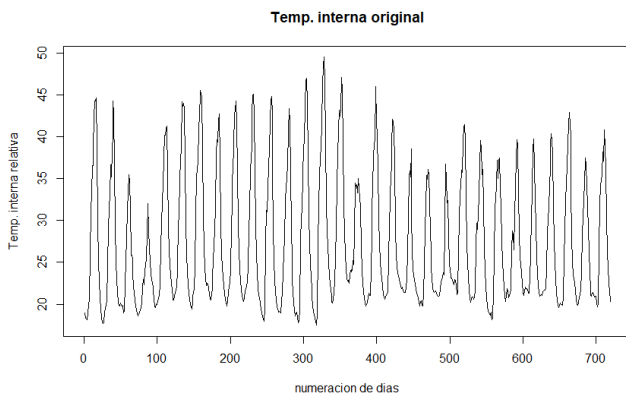
Las librerías usadas fueron: PolynomF la cual según la página de R implementa operaciones con polinomios univariantes en R, incluyendo aritmética de polinomios, búsqueda de ceros, trazado y algunas operaciones con listas de polinomios. [1] Pracma la cual según la página de R proporciona un gran número de funciones de análisis numérico y álgebra lineal, optimización numérica, ecuaciones diferenciales, series temporales, además de algunas funciones matemáticas especiales muy conocidas. Utiliza los nombres de las funciones de MATLAB para simplificar la portabilidad. [2] y readxl la cual sirve para leer archivos xlsx.

Para los errores en los resultados se van a analizar los errores máximos, mínimos, medios y el error medio cuadrático (EMC).

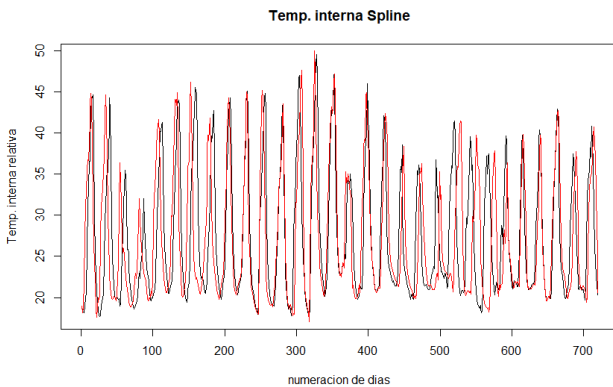
IV. Punto 1

Como eran varias estaciones se para este punto se tomó la de Aiuaba la cual contaba con 720 datos por variable. De las distintas variables se eligió la de temperatura interna.

Primero se mostrará la gráfica con los datos sin aplicar ningún método.



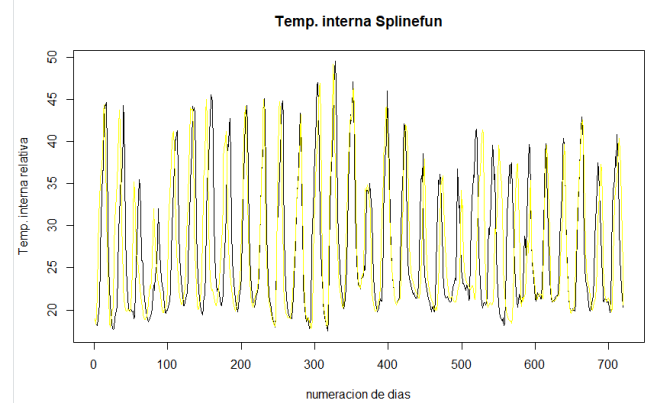
El primer método que se aplicara es el spline cubico el cual corresponde a la función spline() las líneas rojas corresponden a spline.



Y el error:

```
> cat(METODO DE SPLINE '\n')
METODO DE SPLINE
> cat("error EMC: ",emcs,'\n')
error EMC: 7.241257
> cat("error medio: ",(sum(espline))/tam,'\n')
error medio: 5.03517
> cat("error minimo: ",min(espline),'\n')
error minimo: 0
> cat("error maximo: ",max(espline),'\n')
error maximo: 22.97563
```

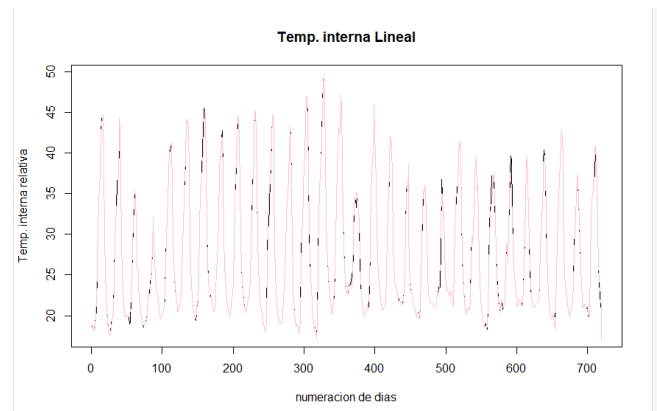
El segundo método que se aplicara es el de interpolación de Hermite el cual corresponde a la función splinefun() las líneas amarillas corresponden a splinefun.



Y el error:

```
METODO DE HERMITE
> cat("error EMC: ",emcc,'\n')
error EMC: 0.5492531
> cat("error medio: ",(sum(ec))/tam,'\n')
error medio: 0.1669731
> cat("error minimo: ",min(ec),'\n')
error minimo: 0
> cat("error maximo: ",max(ec),'\n')
error maximo: 3.876693
```

El tercer método que se aplicara es el de interpolación lineal el cual corresponde a la función aprox() las líneas rosadas corresponden a aprox()



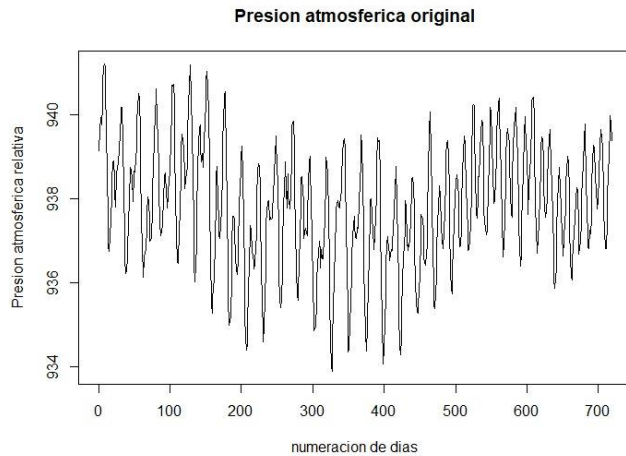
Y el error:

```
METODO LINEAL
> cat("error EMC: ",emcl,'\n')
error EMC: 7.159532
> cat("error medio: ",(sum(el))/tam,'\n')
error medio: 4.989762
> cat("error minimo: ",min(el),'\n')
error minimo: 0
> cat("error maximo: ",max(el),'\n')
error maximo: 22.91442
```

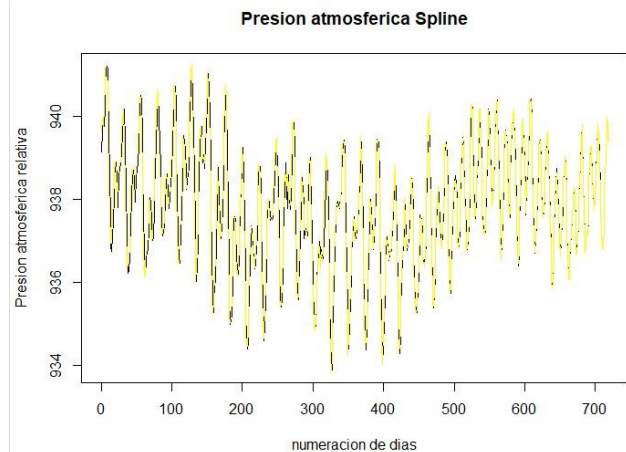
V. Punto2

Como eran varias estaciones se para este punto se tomó la de Araripe la cual contaba con 719 datos por variable. De las distintas variables se eligió la de presión atmosférica.

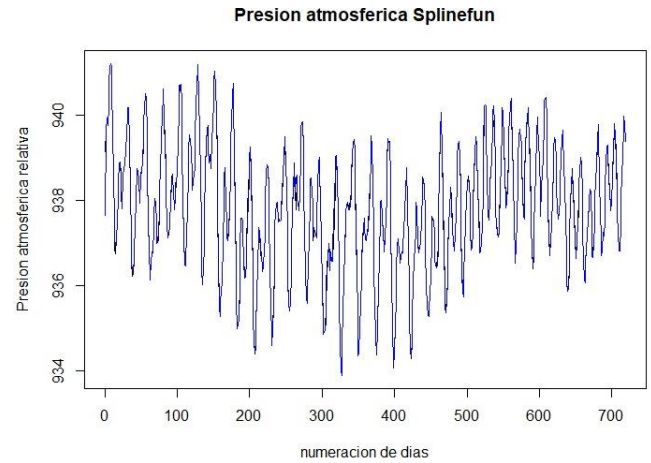
Primero se mostrará la gráfica con los datos sin aplicar ningún método.



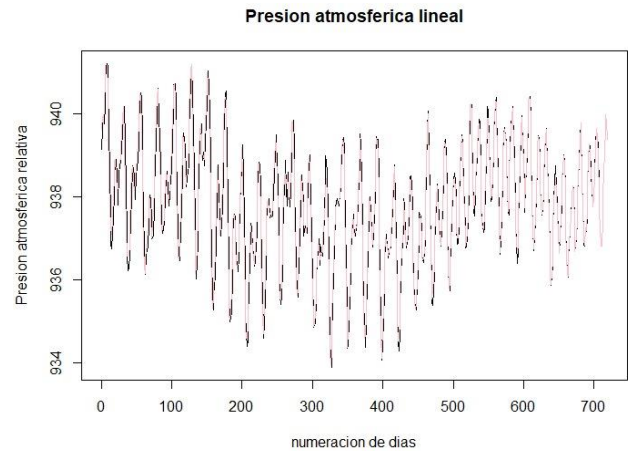
El primer método que se aplicara es el spline cubico el cual corresponde a la función spline() las líneas amarillas corresponden a spline.



El segundo método que se aplicara es el de interpolación de Hermite el cual corresponde a la función splinefun() las líneas azules corresponden a splinefun.



El tercer método que se aplicara es el de interpolación lineal el cual corresponde a la función aprox() las líneas rosadas corresponden a aprox()



En la siguiente imagen veremos los errores para cada método:

```
> cat("MÉTODO DE SPLINE \n")
MÉTODO DE SPLINE
> cat("error EMC: ", emcs, '\n')
error EMC: 0.075406
> cat("error medio: ", (sum(espline))/tam, '\n')
error medio: 0.027809
> cat("error mínimo: ", min(espline), '\n')
error mínimo: 0
> cat("error máximo: ", max(espline), '\n')
error máximo: 0.47373
> (1 - (sum(espline)/sum(datosoriginales)))*100
[1] 99.997
> #-----
> cat("MÉTODO DE HERMITE \n")
MÉTODO DE HERMITE
> cat("error EMC: ", emcc, '\n')
error EMC: 0.075406
> cat("error medio: ", (sum(ec))/tam, '\n')
error medio: 0.027809
> cat("error mínimo: ", min(ec), '\n')
error mínimo: 0
> cat("error máximo: ", max(ec), '\n')
error máximo: 0.4737308
> (1 - (sum(ec)/sum(datosoriginales)))*100
[1] 99.997
> #-----
> cat("MÉTODO LINEAL \n")
MÉTODO LINEAL
> cat("error EMC: ", emcl, '\n')
error EMC: 0.12199
> cat("error medio: ", (sum(el))/tam, '\n')
error medio: 0.040992
> cat("error mínimo: ", min(el), '\n')
error mínimo: 0
> cat("error máximo: ", max(el), '\n')
error máximo: 0.915
> (1 - (sum(el)/sum(datosoriginales)))*100
[1] 99.996
```

VI. Conclusiones

Los errores de los puntos se pueden ver afectados ya sea por el método o el tipo de datos evaluado, pues, aunque se utilizaron los mismos métodos se evidencia que en el segundo problema los valores son mucho más aproximados ya que los datos de temperatura así lo permiten al ser más suaves, a diferencia del primer punto donde se evidencian errores más grandes. Analizando las gráficas y los errores obtenidos, podemos ver que, si bien los puntos son muy acertados en todos los métodos, el mejor en ambos casos es el cual utiliza la función de splinefun que corresponde al método de Hermite, esto se puede comprobar comparando su error en los dos problemas.

VII. Bibliografía

- [1] PolynomF: Polinomials in R. Tomado de:
<https://cran.r-project.org/web/packages/PolynomF/index.html>
- [2] pracma: Practical Numerical Math Functions. Tomado de:
<https://cran.rstudio.com/web/packages/pracma/index.html>