

Trabajo final: Diseño y ejecución de un experimento

Juan Cantero Jimenez y Emilio Vazquez Bandera

3/6/2022

Introducción

Los sistemas de funciones iteradas, SFI, son construcciones matemáticas que permiten representar conjuntos de puntos que presentan autosimilaridad, autoafinidad o autocorrelación. En este caso estamos interesados en SFI que apliquen trasformaciones afijas, así estas SFI tomaran la forma:

$$f_1(x, y) = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c_5 \\ c_6 \end{bmatrix} \quad (1)$$

$$f_2(x, y) = \begin{bmatrix} c_7 & c_8 \\ c_9 & c_{10} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} c_{11} \\ c_{12} \end{bmatrix} \quad (2)$$

de forma más específica se ha representado un SFI con dos atractores. Para la representación de el atractor codificado en este SFI se hace uso del algoritmo de iteración aleatoria. Primero se escoge un punto aleatorio del espacio métrico, en este caso de dos dimensiones, seleccionado el punto se escoge una de las funciones que pertenecen al SFI de forma aleatoria, se aplica esta función al punto inicial. Este mismo procedimiento se repetira sobre el resultado de la primera iteración. Las coordenadas de los puntos en las distintas iteraciones se representan en el espacio métrico obteniendo la representación del atractor codificado en el SFI.

En este trabajo se hace uso del algoritmo propuesto en (1) para la obtención de motivos visualmente relevantes. Así se generan los coeficientes del SFI de forma aleatoria, se obtiene la representación por el metodo de iteración aleatoria mientras se va evaluando la calidad visual de estos mediante dos coeficientes, la dimensión fractal y el exponente máximo de Lyapunov. Siguiendo con las conclusiones obtenidas en (1) aquellos atractores visualmente relevantes serán los que posean una dimensión fractal mayor que 1 y un exponente máximo de Lyapunov menor que -0.2. Si cumple estos requisitos la representación se cosiderará visualmente relevante, éxito, en caso contrario no se considera relevante, fracaso. Las Figuras 1, 2 y 3 muestran tres SFI visualmente relevantes.

En este trabajo se quiere averiguar si existen diferencias estadísticamente significativas entre la media del valor de la dimensión fractal de los SFI visualmente relevantes en función del número de atractores codificados en el SFI. También se quiere averiguar si la estrategia usada para generar los coeficientes de forma aleatoria, mediante una distribución uniforme discretizada o continua, posee algún efecto sobre las medias de la dimensión fractal.

Lectura de los datos

Se cargán y se exploran la estructura del dataset para comprobar que las variables poseen el formato requerido.

```
data = read.csv("dinamica.csv")
str(data)

## 'data.frame': 30000 obs. of 5 variables:
##   $ X          : int 1 2 3 4 5 6 7 8 9 10 ...
##   $ L          : num -0.514 -0.441 -0.453 -0.277 -0.387 ...
##   $ FD         : num 1.1 1.28 1.09 1.12 1.02 ...
```



Figure 1: Ejemplo de representación gráfica de un IFS con dos atractores

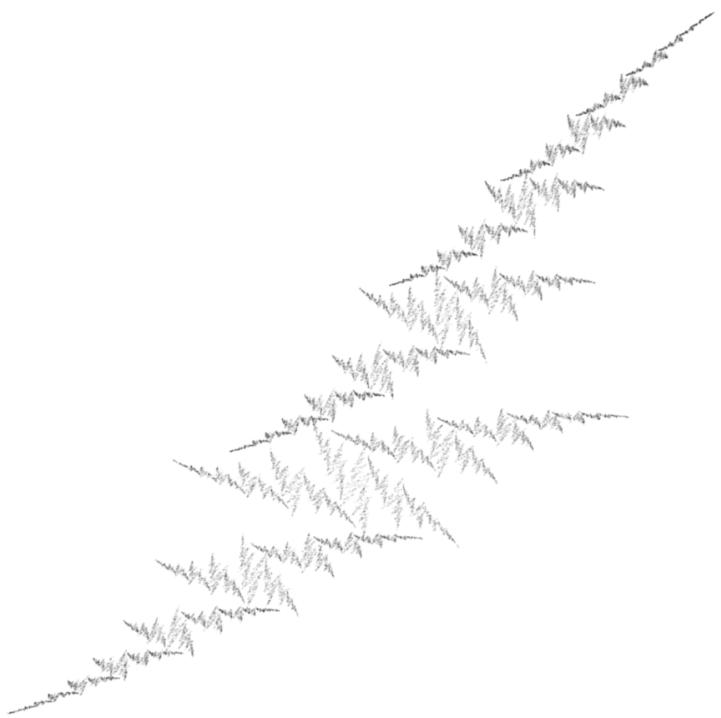


Figure 2: Ejemplo de representación gráfica de un IFS con dos atractores



Figure 3: Ejemplo de representación gráfica de un IFS con dos atractores

```

## $ Atractores: int 2 2 2 2 2 2 2 2 2 ...
## $ Generador : chr "continuo" "continuo" "continuo" "continuo" ...
data$Atractores <- factor(data$Atractores)
data$Generador <- factor(data$Generador)
str(data)

## 'data.frame': 30000 obs. of 5 variables:
## $ X : int 1 2 3 4 5 6 7 8 9 10 ...
## $ L : num -0.514 -0.441 -0.453 -0.277 -0.387 ...
## $ FD : num 1.1 1.28 1.09 1.12 1.02 ...
## $ Atractores: Factor w/ 3 levels "2","3","4": 1 1 1 1 1 1 1 1 1 ...
## $ Generador : Factor w/ 2 levels "continuo","discreto": 1 1 1 1 1 1 1 1 1 ...

```

Análisis de los datos

Puesto que estamos interesados en como cambia la media de la dimensión fractal, variable FD, en función de el tipo de generador de los coeficientes así como del número de atractores que poseen estos SFI debemos de crear un modelo que considere estas como efectos fijos y la interacción de estos:

```

aov1 <- aov(FD ~ Atractores + Generador + Atractores*Generador, data=data)
summary(aov1)

```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)						
## Atractores	2	78.9	39.44	600.702	<2e-16 ***						
## Generador	1	0.0	0.00	0.076	0.783						
## Atractores:Generador	2	0.0	0.02	0.310	0.734						
## Residuals	29994	1969.1	0.07								
## ---											
## Signif. codes:	0	'***'	0.001	'**'	0.01	'*'	0.05	'. '	0.1	' '	1

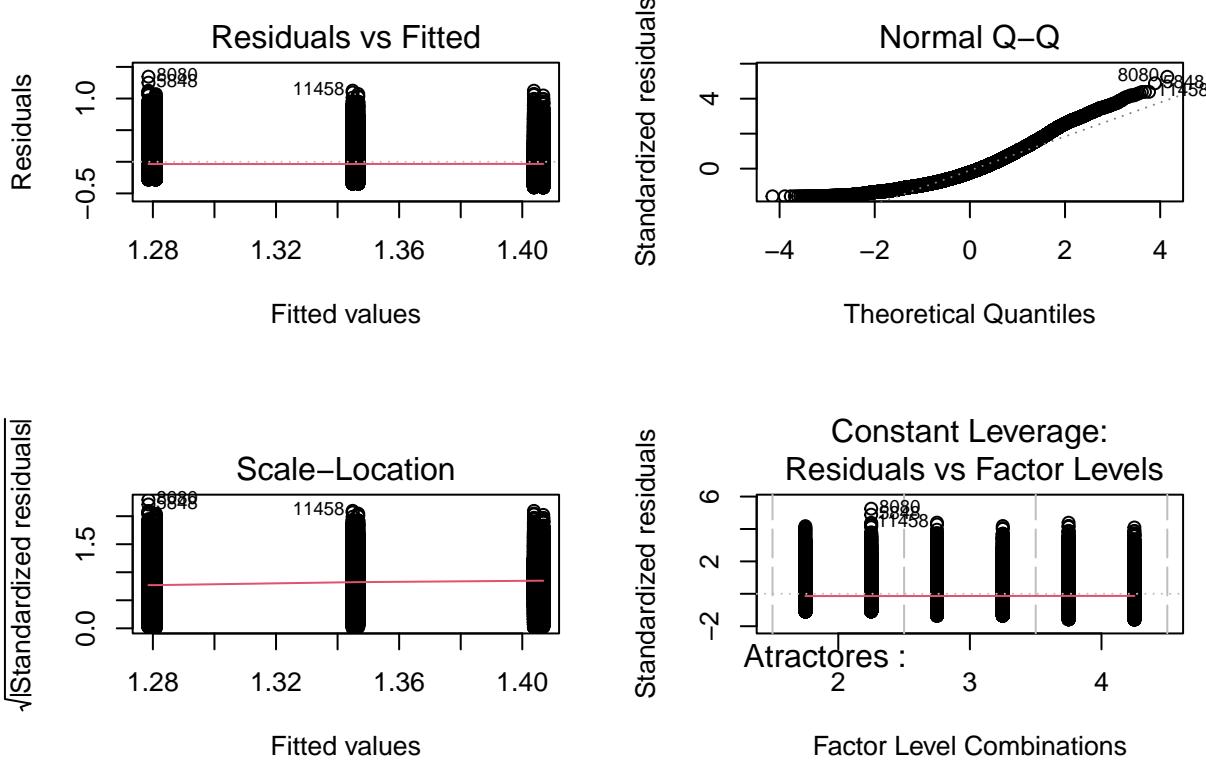
Como se puede observar, el único efecto fijo que posee influencia estadísticamente significativa sobre la media de la dimensión fractal es el número de atractores que posee el SFI.

Para asegurar que este resultado es válido será necesario comprobar las hipótesis de aplicabilidad del modelo.

```

par(mfrow=c(2,2))
plot(aov1)

```



```
ks.test(rstudent(aov1), pt, df=nrow(data)-2-2)

## Warning in ks.test(rstudent(aov1), pt, df = nrow(data) - 2 - 2): ties should not
## be present for the Kolmogorov-Smirnov test
```

```
##
## One-sample Kolmogorov-Smirnov test
##
## data: rstudent(aov1)
## D = 0.08899, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

Puesto que los datos divergen de la normalidad, vease qq-Plot así como resultado del test de Kolmogorov-Smirnov, se intentará realizar una trasformación de tipo Box-Cox sobre los datos

```
lambda <- EnvStats:::boxcox(data$FD, lambda = c(-10,10), optimize = TRUE)$lambda
lambda

## [1] -1.557733

data$FD_transformed <- (data$FD^lambda -1)/lambda
aov2 <- aov(FD_transformed ~ Atractores + Generador + Atractores*Generador, data=data)
summary(aov2)

##                               Df Sum Sq Mean Sq F value Pr(>F)
## Atractores                  2    17.5   8.749 703.186 <2e-16 ***
## Generador                   1     0.0   0.001   0.083   0.773
## Atractores:Generador        2     0.0   0.005   0.367   0.693
## Residuals                  29994   373.2   0.012
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

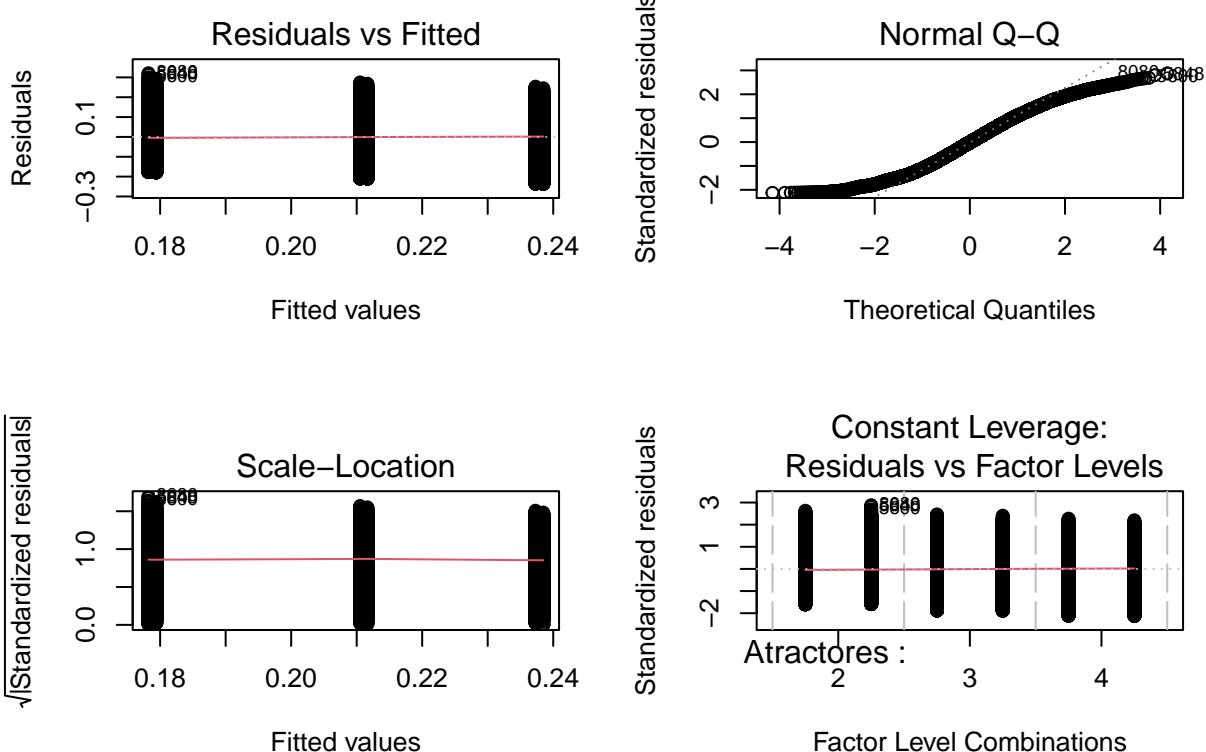
ks.test(rstudent(aov2), pt, df=nrow(data)-2-2)

## Warning in ks.test(rstudent(aov2), pt, df = nrow(data) - 2 - 2): ties should not
## be present for the Kolmogorov-Smirnov test

##
## One-sample Kolmogorov-Smirnov test
##
## data: rstudent(aov2)
## D = 0.036247, p-value < 2.2e-16
## alternative hypothesis: two-sided

par(mfrow=c(2,2))
plot(aov2)

```



puede observar como la transformación de Box-Cox de los datos mejora ligeramente la normalidad de los datos si atendemos al qq-plot, sin embargo el test de Kolmogorov-Smirnov sigue descartando la hipótesis nula de normalidad, esto puede darse al gran tamaño muestral de los datos que hace que se detecte cualquier ligera desviación de la normalidad.

Si se atiende ahora a la hipótesis de homocedasticidad para el modelo con la variable transformada, esta puede ser evaluada observando el plot Residual Vs Fitted o el de Scale-Location, así como mediante un test de Levene

```

grupos <- cut(aov2$fitted.values, quantile(aov2$fitted.values, (0:3)/3),
include.lowest = TRUE)
lawstat::levene.test(rstandard(aov2), grupos)

##
## Modified robust Brown-Forsythe Levene-type test based on the absolute
## deviations from the median
##

```

```
## data: rstandard(aov2)
## Test Statistic = 5.0962, p-value = 0.006125
```

La hipótesis de homocedasticidad se descarta en base al test de Levene. De nuevo, este resultado puede ser consecuencia del gran tamaño muestral lo que permite detectar desviaciones mínimas de la homocedasticidad, pues si observamos los plots antes mencionados se puede observar como la dispersión de los datos parece ser homocedástica.

Por último se realizará un test a posteriori de tipo Tukey para comparar las medias de los grupos entre ellos.

```
aov2.tukey <- TukeyHSD(aov2, which = "Atractores")
aov2.tukey
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = FD_transformed ~ Atractores + Generador + Atractores * Generador, data = data)
##
## $Atractores
##      diff      lwr      upr p adj
## 3-2 0.03227041 0.02857324 0.03596758 0
## 4-2 0.05907430 0.05537713 0.06277147 0
## 4-3 0.02680389 0.02310672 0.03050106 0
```

En base a estos resultados, se puede concluir que los SFI que posean cuatro atractores posee un mayor valor medio de dimensión fractal así como que los que poseen dos atractores tendrán el menor valor medio de dimensión fractal entre las configuraciones de SFI ensayadas.

Conclusiones

A la luz de los datos podemos concluir:

-No se cumplen las condiciones de aplicabilidad de un análisis de la varianza, aunque la hipótesis de normalidad mejora aplicando una transformación de tipo Box-Cox con $\lambda = -1.56$. El hecho de que descarte la hipótesis de homocedasticidad mediante un test de Levene puede deberse al gran tamaño muestral, puesto que no se aprecia heterocedasticidad en los gráficos de valores predichos frente a residuos.

-El modelo indica que el número de atractores posee una influencia significativa en la media de la variable respuesta, mientras que la estrategia por la que se generaron los coeficientes no.

-Los SFI con cuatro atractores poseen la media más alta mientras que los SFI con 2 atractores posee el menor valor medio en la dimensión fractal de entre las configuraciones escogidas.

Bibliografía

- 1-. Sprott, J. C. (1994). Automatic generation of iterated function systems. Computers & graphics, 18(3), 417-425.