Función de Correlación de dos puntos para una distribución de puntos en 3 dimensiones.

Universidad de Guanajuato, División de Ciencias e Ingenierías
Programación Básica
Brandon Emmanuel Hernandez Corona
José Misael Rojas Sánchez

I. Introducción

En las ultimas decadas han existido numerosos avances tecnológicos, lo que se ha traducido en astronomía en un gran cantidad de datos observacionales y en un incremento en limites de simulaciones alcanzables.

Es por esto que fue necesario recurrir a diversas herramientas para estudiar estos nuevos datos. Una de estas funciones correlación ($\xi(r)$), que es una herramienta estadística que ayuda a estudiar sistemas de muchos cuerpos, nos dan información sobre la estructura evolución de sistemas, por ejemplo en la bioestadistica cuando relacionamos poblaciones de especies especificas, también puede ser utilizada para el estudio de galaxias en en un sistema de n-cuerpos gravitatorios, etc.

II. Teoría

La correlación es una medida de la relación(covariación) entre dos variables cuantitativas continuas (x,y). Esta medida o índice de correlación -r- puede variar entre -1 y +1, ambos extremos indicando correlaciones perfectas, negativa y positiva respectivamente. Un valor de r = 0 indica que no existe relación lineal entre las dos variables. Una correlación positiva indica que ambas variables varían en el mismo sentido. Una correlación negativa significa que ambas variables varían en sentidos opuestos. Lo interesante del índice de correlación es que r es en sí mismo una medida del tamaño del efecto, que suele interpretarse de la siguiente manera:

- correlación **despreciable**: $r < |0{,}1|$
- correlación **baja**: $|0{,}1| < r <= |0{,}3|$
- correlación **mediana**: $|0{,}3| < r <= |0{,}5|$
- correlación **fuerte o alta**: $r > |0{,}5|$

II-A. Estimadores.

Existen definiciones alternas para la función de correlación, estas funciones de n-puntos, la que nos interesa estudiar es la función de correlación de dos puntos.

Mide el exceso de probabilidad de encontrar una galaxia a una distancia r de otra galaxia arbitraria en un elemento de volumen σV. Para entender esto pensemos en una distribución de masa como una distribución de objetos puntuales; la probabilidad de encontrar un objeto en un elemtno infinitesimal δV es:

$$\delta P = n\delta V \tag{1}$$

con n la densidad media de objetos, debido a que suponemos isotropía, n es independiente de la posición. Luego, la función correlación de dos puntos se define como la probabilidad conjunta de encontrar un objeto en los elementos de volumen δV_1 y δV_2 separados por $r_1 2$, esta probabilidad es dada por:

$$\delta P = n^2 \delta V_1 \delta V_2 [1 + \xi(r_1 2)] \tag{2}$$

Si las posiciones del objeto están correlacionadas se tiene que $\xi > 0$, pero si estas no están anticorrelacionas, se tiene que -1 <= ξ <0.

Por otro lado, debemos considerar que para hacer un mejor uso de la función correlación es necesario introducir una segunda expresión que la defina, la cual es mas simple de usar en el momento en que se esta trabajando con datos experimentales o de alguna simulación.

Esta se considera como un cubo con un lado L, donde se encuentran contenida un par de galaxias a una distancia $r + dr$ en una muestra de galaxias ($DD(r)$) y el número de pares esperado en la ausencia de de clustering ($RR(r)$), que se encuentran a una distancia $r + dr$, tomando en cuenta los limites de la muestra y considerando una distribución homogenea. Explicitamente la expresión tiene la siguiente forma:

$$1 + \xi(r) = \frac{DD(r)}{RR(r} \tag{3}$$

- Para estimar $DD(r)$ se deben calcular las distnacias que hay entre los objetos de la muestra, tenindo estos datos podemos pasar a clasificar el número de galaxias que hay por rango de distancia, finalmente este número correspondera al valor DD(r).
- Para estimar $RR(r)$ consideraremos una distribución de poisson para las objetos en un volumen definido, la expresión para esto es:

$$RR(r) = \frac{2\pi(r_e^3 - r^3)N_g^2}{3V} \tag{4}$$

donde N_g es el número de objetos que hay en el volúmen, el cual dependera de las dimensiones de la muestra de la cual se obtuvieron los datos; $r_e = r + dr$ donde dr es el radio del casquete esferico dado por el radio interior r.

III. PROGRAMA.

Estimar la función de correlación del grupo de datos usado utilizaremos la ecuación (3). Esto sginifica que el codigo necesario para calcular datos de la muestra, debe de realizar un conteo de la cantidad de datos cuyas dirancias esten dentro de un cierto ranfo de distnacias, calcular $RR(r)$ para ese rango de distancias y luego realizar el ajuste correpondiente.

Entonces, lo anterior se puede considerar para determinar las distancias entre los objetos de la muestra. Para esto consideramos las columnas correspondientes a las coordenadas x, y, y z de cada objeto. Estos datos son guardados en nuestro arreglo para que el manejo sea más eficiente. Entonces teniendo los valores de x, y, y z, la distancia se calcula por medio de la siguiente formula:

$$r = \sqrt{dx^2 + dy^2 + dz^2} \tag{5}$$

donde dx, dy y dz son dados por las diferencias entre las correspondientes coordenadas de los objetos que se calculan; o sea:

$$dx = x_1 - x_2, dy = y_1 - y_2, dz = z_1 - z_2 \tag{6}$$

Teniendo las distancias calculadas y guardadas en este arreglo se pasa a crear un nuevo arreglo donde cada matriz correspondera a un rango de distancia distinto. Con este arreglo se comienza a contar la cantidad de pares que hay por cada ranfo de distancia $(DD(r))$, luego el número que arroje se guarda en la matriz correspondiente.

Al final calculamos los $RR(r)$ por medio de la ecuacion (4) para cada rango de distancia con el que definimos el arreglo anterior. Con estos valores se pueden obtener las cantidades obtenidas de $DD(r)$ y este nuevo valor vuleve a ser guardado en el arreglo y en la matriz de las distancias. El valor que obtengamos para el final sera determinado como el estimador de función correlación de los puntos.

REFERENCIAS

Peebles, Phillip J. E. The large-scale structure of the universe .Princeton University Press (1980) Davis M., Peebles P., 1982, ApJ, 267, 465D. Davis M. et al, 1988, ApJ, 333L, 9D.