Un estudio sobre ligaduras por condicionamiento

APLICACIÓN A LA SERIE DE TEMPERATURAS GLOBALES

JOSÉ ANTONIO LÓPEZ DÍAZ AEMET, MADRID

Ligaduras globales en Física y climatología

En los modelos climáticos (y de predicción del tiempo) la atmósfera se representa por ecuaciones en puntos de una rejilla que describen la evolución temporal del sistema por medio de ecuaciones físicas. Este punto de vista responde a la visión dinámica, en la que el sistema evoluciona en el tiempo de forma progresiva, de tal forma que las ecuaciones básicamente relacionan las derivadas temporales de las variables con otras cantidades. En la Mecánica este punto de vista es el que representa la segunda ecuación de Newton que, como es sabido, relaciona la derivada segunda respecto al tiempo de la posición del sistema con la fuerza aplicada.

Pero cualquier estudiante de Mecánica sabe que existe otro punto de vista que, aunque matemáticamente equivalente al anterior, resulta más intrigante cuando se reflexiona sobre él. Me refiero a los principios variacionales que permiten derivar la trayectoria del sistema a partir de cantidades de carácter global, una globalidad que hace desaparecer aparentemente la evolución progresiva en el tiempo puesto que esta magnitud es integrada a lo largo de cada posible trayectoria. Este enfoque variacional resulta mucho más intrigante para nosostros que el anterior dinámico, puesto que nuestra inteligencia está desarrollada para analizar los fenómenos en su discurrir temporal paso a paso, y con los principios variacionales nos surge inmediatamente la cuestión de cómo el sistema "conoce" las posibles trayectorias completas en el futuro (antes de tiempo) para luego "escoger" la que el principio variacional dicta. Por supuesto es muy cierto que matemáticamente ambos enfoques conducen al mismo resultado (por ejemplo el principio de mínima acción conduce a la ecuación de Newton y viceversa), pero la forma holística en que se formulan los principios variacionales los hace misteriosos. Misterio que no hace sino aumentar cuando a menudo los principios variacionales en Física permiten una formulación más sintética de las leyes físicas: pensemos en el principio de Fermat para los rayos de luz, que dice simplemente que el rayo sigue una trayectoria de tiempo mínimo (con más precisión estacionaria en cuanto a longitud óptica) y que incluye las leyes de la reflexión y refracción de la luz como casos particulares; o el ya mencionado de mínima acción, maravilloso en su simplicidad (para los estudiantes de física de mi época es legendario el comienzo del tomo de Mecánica de Landau en que establece las bases de la Mecánica a partir de este principio); en fisica teórica es muy notable que las ecuaciones de la relatividad general de Einstein se pueden deducir del mucho más sintético en su formulación principio variacional de Hilbert (que por unos días no llegó por esta vía antes que Einstein a sus ecuaciones), o la formulación variacional de Feynman de la electrodinámica cuántica.

Podemos plantearnos si es posible un enfoque parecido al holístico de los principios variacionales mencionados en climatología. No me refiero claro está a la búsqueda de un principio variacional teórico que abarque de forma sintética, como en los ejemplos antes mencionados, la evolución del clima. Esto parece imposible por la complejidad del sistema climático. Pero quizá podamos desarrollar un marco conceptual para algún problema particular que destile algo de la cautivadora fragancia holística de los grandes principios variacionales de la Física. Creo que algo de esto es posible introduciendo un marco probabilístico, que ya se sabe que el clima es un sistema tan complejo que la vía determinista tiene un recorrido muy corto o incluso nulo.

Podemos pensar en una variable de carácter global, como la temperatura media global, a la que imponemos en un marco probabilístico restricciones o ligaduras de carácter también global. En el caso de la temperatura global podemos postular que se obtiene como media de N temperaturas parciales, que supondremos en principio independientes entre sí. La forma más directa de imponer una ligadura global a las N temperaturas parciales es por medio del condicionamiento; consideremos pues que las N temperaturas parciales tienen todas la misma distribución, en principio sin especificar, pero con media cero (después de pasar a anomalías) y que son independientes entre sí. A estas variables aleatorias imponemos la condición de que su suma (y por tanto su media) sea cero. Con estos datos ya podemos hallar con cálculos sencillos que la media de cada variable parcial después de condicionar sigue siendo cero (para demostrarlo expresar que E[suma de X_i | suma de X_i = 0] = 0 y utilizar la simetría de las variables

Un estudio sobre ligaduras por condicionamiento

X_i para concluir que E[Xi | suma de X_i = 0] =0). Además podemos también ver que el coeficiente de correlación entre un par cualquiera de variables después de condicionar vale -1/(N-1) (para probarlo expresar que E[(suma de $X_i)^2$ | suma de X_i = 0] = 0, desarrollar la suma al cuadrado y usar la simetría entre las variables X_i). Esta coeficiente de correlación negativo, que expresa asociación negativa entre las variables, responde al hecho intuitivo de que la ligadura sobre la suma hace que a una variación positiva grande de una variable tienda a corresponder una de sentido opuesto de otra para conseguir que la suma de todas sea cero. Disminuye con N dado que al aumentar N el efecto de compensación entre variables requerido para conseguir suma nula es menor. Así pues la ligadura destruye la independencia estricta entre variables que hemos supuesto de partida, induciendo entre ellas un correlación negativa que decrece con N.

El caso gaussiano

Podemos avanzar más en el estudio de los efectos de ligaduras si suponemos que las variables parciales tienen distribución gaussiana. Y esto es así debido a una propiedad sorprendente a priori que tiene la distribución gaussiana: la función de densidad compuesta de varias variables gaussianas indenpendientes con media cero y la misma desviación estándar depende exclusivamente de la suma de los cuadrados de las coordenadas. Esta propiedad, que se deduce inmediatamente de la dependencia básica de cada función de densidad individual de la forma Exp [- x^2], tiene consecuencias profundas en la teoría de probabilidades y estadística. En esencia la propiedad indicada significa que la función de densidad conjunta tiene simetría esférica (al ser la suma de los cuadrados de las coordenadas igual al radio al cuadrado del punto en el espacio euclídeo N-dimensional); esto produce un completo acoplamiento de las propiedades estadísticas y geométricas. Por ejemplo la independencia estadística en este modelo coincide con la ortogonalidad geométrica en el espacio euclídeo N-dimensional. A menudo en exposiciones elementales de la estadística no se resalta este hecho que es rico en consecuencias. Pensemos en dos muy importantes: la inferencia estadística con el modelo de regresión lineal con errores gaussianos, con el uso de la F de Fisher-Snedecor para valorar la significación estadística de la regresión, descansa en último análisis en esta propiedad; igualmente la inferencia estadística para el periodograma con ruido blanco gaussiano, que es la base para la inferencia estadística espectral, también descansa en este hecho (ver por ejemplo Brockwell & Davis, Time Series: Theory and Methods, Springer Series in Statistics para una exposición clara). Cabe también mencionar que esta propiedad le sirvió a Gauss para encontrar la integral normalizadora de la función de densidad gaussiana, o sea, la integral entre 0 e infinito de Exp[-x^2]. El célebre truco de magia que usó consistió en expresar la integral al cuadrado como integral extendida al plano en virtud del teorema de Fubini para integrales múltiples, y luego pasar a coordenadas polares lo que reduce la integral a dos integraciones elementales.

Retrato de Carl Friedrich Gauss, por Christian Albrecht Jensen

Pues resulta que esta maravillosa propiedad también simplifica mucho el problema de condicionamiento que estamos considerando. La esencia del argumento es como sigue: el condicionar a suma cero supone buscar la densidad sobre el hiperplano N-1- dimensional que pasa por el origen definido por la condición suma de X_i = 0. Esto para una distribución de las X's arbitraria no es sencillo debido a que la integral normalizadora no es en general simple. Pero en el caso de distribuciones parciales gaussianas, en virtud de la simetría esférica, la distribución sobre el hiperplano de suma nula es idéntica a la de cualquier hiperplano que pase por el origen (la simetría esférica hace que todos los hiperplanos por el origen tengan las mismas propiedades). Así que la densidad sobre el hiperplano suma cero es la misma que sobre el hiperplano X_N = 0, por ejemplo. Pero la densidad sobre este último es simplemente la inducida por las X_i para i=1, 2, ..., N-1. O sea que el problema queda reducido en esencia a una sencilla reducción de la dimensionalidad.

Para concretar más el argumento anterior sólo falta algo de álgebra: la idea es que si pasamos en el espacio N-dimensional de la base ortonormal estándar a una base ortonormal con un vector en la dirección (1,1,....,1) (por ejemplo la base de la transformada de Fourier discreta, ver el libro de Brokwell & Davis antes mencionado, en que el vector de unos se asocia a frecuencia de Fourier cero) entonces la distribución sin condicionar es simplemente la suma de N variables gaussianas por los vectores de la nueva base. Y la distribución condicionada a suma cero es la suma de las N-1 variables gaussianas por los vectores de la base quitando el de unos. Esto da además un algoritmo sencillo de simulación de la citada distribución condicionada, que sólo requiere generar N-1 variables aleatorias gaussianas y multiplicarlas por una matriz.

Vemos además que podemos generalizar lo anterior, en el sentido de que si condicionamos, además de a suma nula, a otras combinaciones lineales nulas con la condición de que las combinaciones lineales sean ortogonales entre sí y respecto al vector de unos que representa la suma, el problema se reduce en esencia a disminuir la dimensionalidad en el número k de ligaduras ortogonales impuestas.

También, con algo de trabajo adicional, podemos relajar las ligaduras, en el sentido de imponer no una igualdad estricta a cero, sino que el valor de la combinación lineal o suma sea un v.a. normal con varianza dada. En concreto, para variables X_i gaussianas estándar, si imponemos varianza de la suma normalizada (para que el vector de unos tenga norma unidad) s1 y varianza del resto de las combinaciones lineales ortogonales normalizadas s2, .., sk, si hay en total k ligaduras, se obtiene que la varianza de cualquier variable Xi después de condicionar vale:

$$1/N \, (N - k + \Sigma \, sj) \qquad (1)$$

Y que el valor medio de la covarianza entre dos variables (promedio para todas las parejas X_i y Xj con i ≠ j) vale:

$$1/N \, (s1 -1 -1 /(N\text{-}1) \, (s_2+s_3+..+ s_{k-1} -(k\text{-}1))) \qquad (2)$$

Vemos que todas las ligaduras contribuyen a la reducción en la varianza de forma equitativa, pero que el efecto medio en la covarianza es mayor para la ligadura en la media s1 que para las demás combinaciones ortogonales (que tienen un efecto del orden de 1/N el anterior).

El cociente entre los dos valores anteriores (1) y (2) arroja un valor típico medio del coeficiente de correlación (no es exactamente el coeficiente de correlación medio porque la media del cociente no es el cociente de las medias)

$$\frac{(N\text{-}1) \, s1 - \Sigma^{\prime} s -(N\text{-}k)}{(N\text{-}1) \, (N\text{-}k+\Sigma s)} \qquad (3)$$

donde el sumatorio con la prima indica suma extendida a todas las s menos a s1. Para el caso de que todas las varianzas de las ligaduras sean iguales a s la expresión anterior queda:

$$\frac{(N\text{-}k) \, (s\text{-}1)}{(N\text{-}1) \, (N+k \, (s\text{-}1))} \qquad (4)$$

Desarrollando para valores pequeños de s hasta el primer orden el valor típico medio del coeficiente de correlación queda -1/(N-1) + s/(N-k) N /(N-1) + O(s^2). Por tanto el efecto de orden cero en s fundamental se debe a la restricción en la suma (término -1((N-1)) y el efecto de las otras ligaduras es para N moderado fundamentalmente s/(N-k) con signo contrario al anterior. No obstante para valores de s menores que la unidad (ligaduras restrictivas de la varianza, recordemos que la varianza sin ligaduras de cada combinación lineal normalizada es la unidad) vemos de la expresión completa que el valor típico del coeficiente medio es negativo.

Una aplicación a la serie de temperaturas globales

Veamos cómo el esquema conceptual anterior se puede aplicar a la serie de temperaturas medias globales. Podemos suponer que la temperatura media global de cada año se obtiene de aplicar una ligadura a la media como la discutida antes a un conjunto de N variables que representan componentes de la temperatura global. Por tanto necesitamos generalizar el caso de ligadura a suma (o media) cero considerado antes a ligadura en la media igual a m(t) para el año t. Por fortuna, de nuevo, el caso gaussiano es particularmente sencillo: si suponemos las N v.a. componentes con distribución gaussiana idéntica e independientes entre sí, la distribución después de condicionar a media m(t) es la misma que la que se obtiene tras condicionar a media nula y sumar a todas las componentes m(t). Esto se debe al hecho de que el efecto de condicionar en la componente según el vector de unos (que representa la media) es independiente del resto de componentes según los otros vectores de la base (pues la independencia en este modelo coincide como vimos con la ortogonalidad). Esto en sí me parece una propiedad notable de la distribución gaussiana; pensemos que, en general, para variables con idéntica distribución arbitraria, al condicionar a suma alta la densidad de puntos en el hiperplano correspondiente contendrá valores en las colas de la distribución correspondiente, por lo que no esperamos una relación sencilla con la densidad correspondiente a suma nula.

Así pues vemos que no aparece ninguna complicación esencial adicional al condicionar a media variable. Consideremos pues que tenemos una serie de medias globales de T años, para t = 1, 2, ..., T, siendo en cada año t la media m(t). En cada año las N variables componentes y_i i = 1, 2 ,..., N, se distribuyen como variables gaussianas de varianza σ^2 y media común μ arbitraria condicionadas a media m(t). Para este modelo la varianza de cualquiera de las N componentes y_i promediada en los T años tiene un valor esperado:

$$\langle var \; y_i \rangle = s^2(m(t)) + \sigma^2(1\text{-}1/N) \qquad (5)$$

donde el primer término indica la varianza muestral de m(t) y el segundo muestra que la varianza de cada componente se reduce por el factor (1 – 1/N) debido a la ligadura impuesta, como se vio antes. Para obtener esta expresión basta aplicar el teorema de la varianza total.

De forma similar se prueba que el valor medio en los T años de la covarianza entre dos cualesquiera de las componentes toma el valor esperado:

$$\langle cov \; (y_i \, , \, y_j) \rangle = s^2(m(t)) + \sigma^2/N \qquad (6)$$

En esta expresión vemos que hay dos contribuciones de signo opuesto a la covarianza de dos cualesquiera componentes: por una parte la tendencia global m(t) contribuye positivamente con su varianza muestral, por otra parte la ligadura impuesta cada año contribuye sustractivamente con $-\sigma^2 / N$.

De la anterior expresión vemos que para una varianza propia de las componentes σ^2 dada existe un valor de la varianza de la media global $s^2(m(t))$ que hace que el valor esperado medio de la covarianza entre dos componentes se anule. Este punto singular representa un equilibrio entre la tendencia a asociar positivamente las componentes que representa la varianza de su media y la tendencia a asociar negativamente a las mismas que representa la ligadura impuesta. Para el caso de tendencia global lineal de la forma

$$m(t)=\alpha t+\beta$$

se tiene que la varianza de la tendencia tomo la forma:

$$s^2(m(t))=\alpha^2/12\ (T^2-1) \sim (\alpha T)^2/12$$

El término σ^2 / N depende de la varianza de las componentes y de su número, por lo que su estimación precisa requiere el conocimiento de esas componentes. Sin embargo, si sólo conocemos la evolución de la media global, podemos

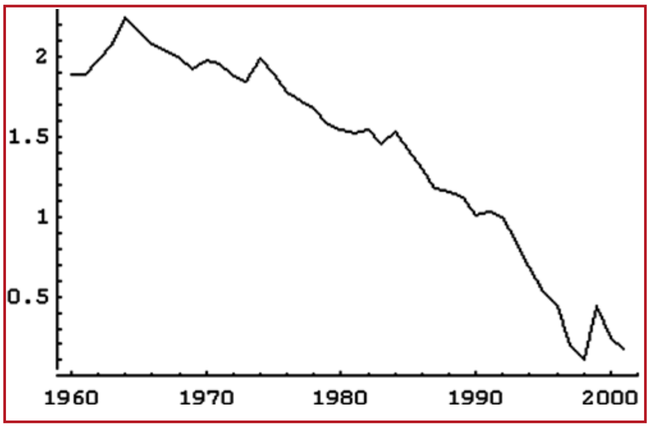

Fig.1: Factor q de (7) para la serie de temperaturas medias globales, para subseries comenzando en el año en abscisas y terminando en 2010

de forma aproximada pensar en estimarlo, a partir de su carácter de varianza de una media, por la varianza residual de las temperaturas en un modelo de regresión lineal. Tenemos que la condición de equilibrio quedaría en la forma:

$$q = |\alpha T| / (\sqrt{12}\,\sigma_m) \sim 1 \qquad (7)$$

En donde σ_m representa la desviación típica residual.

En la figura 1 he tomado la serie de temperaturas medias globales de la página web de la Climate Research Unit de la Universidad de East Anglia, que termina en 2010, y para cada año en abscisas a partir de 1960 he ploteado el parámetro q de la ecuación (7) para subseries de año de comienzo variable en abscisas y terminando en 2010. Vemos que el equilibrio entre la varianza de la tendencia y la contribución de

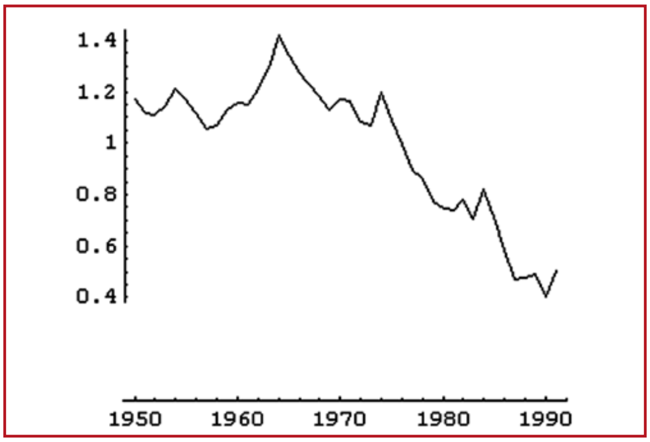

Fig.2: Como en la figura 1 para subseries terminando en 2000

los grados de libertad internos, que corresponde a q = 1, se establece a finales de los ochenta del pasado siglo. Para comienzos anteriores a este punto prepondera la contribución de la varianza de la tendencia, y que q alcanza un valor del orden 2 a principios de los 60 del pasado siglo.

En la figura 2 se hace el mismo análisis pero para las subseries terminando en el año 2000. Aquí vemos que para alcanzar el equilibrio de las dos componentes dado por q = 1 es necesario remontarse hasta mediados de los 70 del pasado siglo.

Por tanto para alcanzar el equilibrio terminando en 2010 bastan unos 20 años, pero terminando en 2000 se requieren unos cinco años más.

REFERENCIAS

- **LOS DATOS USADOS ESTÁN DISPONIBLES EN HTTP://WWW.CRU.UEA.AC.UK/CRU/INFO/WARMING/**
- **BROCKWELL & DAVIS, 1991: TIME SERIES: THEORY AND METHODS (2ª ED.),** Springer Series in Statistics
- **BROHAN, P.,** J.J. Kennedy, I. Harris, S.F.B. Tett and P.D. Jones, 2006: Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. J. Geophysical Research 111, D12106, doi:10.1029/2005JD006548
- **GRIMMETT & STIRZAKER, 1992: PROBABILITY AND RANDOM PROCESSES (2ª ED.),** Oxford Science Publications
- **KLIMOV, 1986: PROBABILITY THEORY AND MATHEMATICAL STATISTICS, MIR PUBLISHERS**