La técnica del “*fingerprint* óptimo” para el problema de detección en cambio climático y su enfoque tensorial en Hasselmann (1997)

JOSÉ ANTONIO LÓPEZ DÍAZ METEORÓLOGO DEL ESTADO

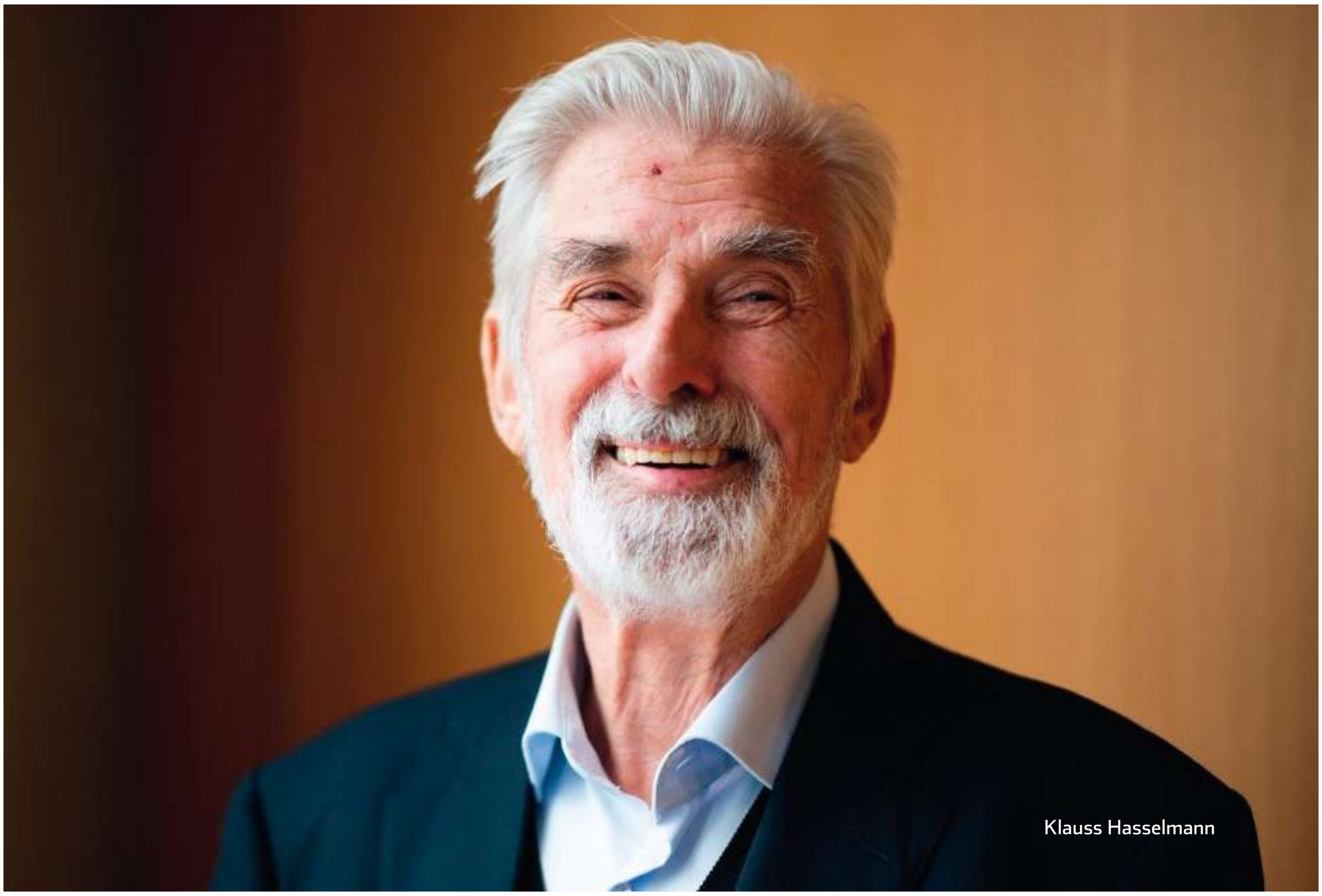
Las reflexiones que siguen surgieron al hilo de la lectura del artículo de Klaus Hasselmann (1997), recientemente galardonado con el premio Nobel en Física por sus contribuciones a la modelización climática y su tratamiento estocástico. Una de las contribuciones señaladas de Hasselmann ha sido la técnica de la huella digital (*fingerprint*) óptima para el problema acuciante de la detección de señales climáticas embebidas en la variabilidad climática natural, técnica esencial para fundamentar objetivamente la influencia antropogénica en el clima. En el citado artículo introduce el autor un enfoque tensorial que llama la atención en este tipo de técnicas climatológicas en que el énfasis no está en las propiedades puramente geométricas de los objetos.

Empezaré con una breve introducción al análisis tensorial, lo justo para que el lector poco familiarizado con esta teoría pueda seguir estas reflexiones. A este respecto quiero recordar el excelente texto de análisis tensorial en coordenadas cartesianas escrito por el célebre meteorólogo y profesor universitario Francisco Morán Samaniego. Continuaré mostrando la conexión del análisis tensorial con las funciones de densidad multivariantes que permite la combinación de la invariancia típicamente tensorial con la incertidumbre estocástica en la formulación del problema de la detección. Por último, termino desgranando la deducción de la huella digital óptima con el enfoque tensorial que hace Hasselmann, y, como consecuencia del análisis, concluyo apuntando una posible generalización de la técnica.

Breve introducción a los tensores

Aunque los tensores alcanzan toda su potencia y elegancia en coordenadas curvilíneas, para los propósitos presentes basta considerar transformaciones lineales. El análisis tensorial se desarrolló a partir de la necesidad de expresar de forma general las propiedades invariantes de entes geométricos, como los vectores ordinarios de la física, frente a cambios de coordenadas. Una forma rápida de ver la necesidad de los tensores consiste en partir del producto escalar de dos vectores.

Si expresamos los vectores en un sistema de referencia ortonormal (ejes perpendiculares y vectores base de longitud unidad) la expresión del producto escalar toma la forma conocida $\mathbf{x.y} = x^i y^i$ donde uso la convención habitual en análisis tensorial de sumar sobre índices repetidos. Si cambiamos a otro sistema de referencia por un giro en el espacio del sistema de referencia ortonormal sabemos que la expresión del producto escalar no cambia en los nuevos ejes, notando con primas las nuevas componentes el producto escalar es $\mathbf{x.y} = x^{'i} y^{'i}$. Si en cambio pasamos a un sistema de referencia oblicuo por medio de una matriz $\mathbf{A}$ (no ortogonal), de forma que $\mathbf{x'} = \mathbf{A\,x}$ y la expresión análoga para y, vemos que la expresión del producto escalar ya no es la misma. Aparece entonces la necesidad de distinguir entre dos tipos de vectores por así decir. Los vectores llamados contravariantes se transforman según la ley anterior, los llamados covariantes según la ley $\mathbf{x'} = (\mathbf{A}^T)^{-1}\,\mathbf{x}$. Las componentes contravariantes se designan con superíndices (como hice antes), las de los covariantes con subíndices. Se puede entonces verificar que se tiene $\mathbf{x.y} = x^{'i} y'_i = x'_i y^{'i}$; de esta forma se garantiza la invariancia del producto escalar frente a cualquier cambio de coordenadas lineal (se dice que el producto escalar es un escalar tensorial). En los sistemas de referencia ortonormales las componentes contra- y covariantes de un vector coinciden, y esta definición es coherente, pues aplicando la ley de transformación anterior para un giro del sistema ortonormal, la matriz $\mathbf{A}$ es ortogonal, y por tanto ambas componentes se transforman igual. Se generaliza lo anterior a expresiones de escalares por medio de matrices. Si en un sistema ortonormal formamos el escalar $C_{ij}\, x^i y^j$, para garantizar que el escalar no cambia ante un cambio lineal de coordenadas la matriz debe transformarse según $\mathbf{C'} = (\mathbf{A}^T)^{-1}\,\mathbf{C}\,\mathbf{A}^{-1}$. En este caso, como sugiere la notación para los subíndices de la matriz, se dice que la matriz es un tensor doblemente covariante. Si partimos de la expresión del mismo escalar pero en la forma $C^{ij} x_i y_j$ obtenemos la ley de transformación de la matriz en forma doblemente contravariante $\mathbf{C'} = \mathbf{A\,C\,A}^T$, nótese que el inverso de una matriz en forma doblemente covariante se transforma con una matriz doblemente contravariante, y viceversa. Se pueden escribir también expresiones para el mismo escalar como producto de los vectores en forma covariante y la matriz en forma doblemente contravariante, o en forma mixta siguiendo razonamientos análogos. De nuevo en un sistema de referencia ortonormal las componentes de una matriz contra- o covariantes, o mixtas son idénticas. La propiedad básica de los tensores es que una igualdad de tensores del mismo tipo es válida en cualquier sistema de referencia, de tal forma que si dos tensores del mismo tipo son idénticos en un sistema de referencia determinado, lo son en todos. Una operación básica es la suma respecto de un índice repetido de tipo contra y de tipo covariante, que da un tensor de un tipo sin ambos índices, lo que generaliza lo dicho para el producto escalar.



Tensores en funciones de densidad multivariante

Comencemos con el caso de una dimensión, con una variable aleatoria (v.a.) X, con función de densidad $f_X(t)$. Si pasamos a otra v.a. Y por una transformación de escala de parámetro s, Y = s X, podemos interpretar la nueva v.a. Y de dos formas matemáticamente equivalentes: una forma activa, en que realmente pensamos en otra v.a. Y relacionada con X por una multiplicación, que es quizá la más habitual, y otra digamos pasiva, en que pensamos que lo que hemos hecho es expresar la misma cantidad aleatoria X en otra unidad s veces más pequeña, como cuando con una longitud pasamos de metros a centímetros, en que s =100. Este segundo caso es el que corresponde a un cambio de coordenadas en consonancia con el análisis tensorial. La función de densidad de probabilidad (pdf) de Y se obtiene expresando que en la transformación se conserva el escalar tensorial probabilidad infinitesimal en un intervalo, o sea, $f_Y(y)\,dy = f_X(x)\,dx$ para $y = s\,x$; de donde $f_Y(t) = 1/s\, f_X(t/s)$. Vemos que, aparte del factor 1/s dentro del argumento de f_X, que tiene en cuenta la transformación de coordenadas para un mismo punto del eje, aparece el factor 1/s que tiene en cuenta el cambio de longitud al pasar de dx a dy. Generalizando al caso N-dimensional, dado un vector aleatorio **X** con N componentes, su función de densidad f_X es un función positiva de N argumentos tal que integrada a todo el espacio N-dimensional vale la unidad. Si se aplica una transformación lineal al vector **X** dada por una matriz A, **Y = A X**, que vamos a interpretar en lo sucesivo de forma pasiva como un cambio de coordenadas en el espacio N-dimensional para la misma magnitud aleatoria, la función de densidad de **Y** toma la forma $f_Y(\mathbf{t}) = \det(\mathbf{A})^{-1} f_X(\mathbf{A}^{-1}\mathbf{t})$, donde el determinante jacobiano tiene en cuenta la relación de elementos de volumen. Podría parecer que en el caso multivariante, una pdf como $f(x+y^2)$ en dos dimentsiones no se presta a ponerla en la forma anterior tras una transformación lineal por no aparecer el vector **t**=(x, y) en forma explícita. Pero podemos extraer cada una de sus componentes multiplicando escalarmente por los vectores unitarios estándar, del tipo (1,0) y (0,1) en este ejemplo en dos dimensiones, y ya aparece el vector **t** en forma explícita: se tiene $f(x+y^2) = f((\mathbf{t.e}_1) + (\mathbf{t.e}_2)^2)$. De todas formas en un caso de una función similar en más variables sin duda la expresión final es engorrosa y no se suele expresar una pdf multivariante con una “matriz de escala” **A** incluida explícitamente.

Hay un caso en que la pdf transformada toma una forma particularmente sencilla, y además expresando la transformación exclusivamente en función de la propia matriz de covarianza del vector aleatorio **X** (en el caso unidimensional siempre se puede sustituir el parámetro s de la transformación de escala por la desviación estándar de X pues es fácil ver que ambos son proporcionales). Sucede esto cuando la pdf multivariante es función tan solo del módulo de **X**, $|\mathbf{X}| = (\mathbf{X.X})^{1/2}$. En este caso la pdf tiene simetría esférica, es decir, todos los puntos en la N-esfera $\mathbf{X.X} = R^2$ tienen la misma función de densidad. Como consecuencia de esta simetría esférica, $E[\mathbf{X}] = \mathbf{0}$, pues no puede haber ninguna dirección del espacio preferida, y la matriz de covarianza $\mathbf{S}_X$ es un múltiplo de la matriz unidad, es decir, con elementos no nulos positivos iguales solo a lo largo de la diagonal. En efecto, la cuádrica de dicha matriz, de ecuación $\mathbf{t}^T \mathbf{S}_X\, \mathbf{t} = \mathbf{q}$, siendo q una constante positiva arbitraria tiene que ser una esfera debido a la simetría esférica, y esto implica lo afirmado. Así pues siempre podemos, multiplicando **X** por una constante, suponer que tiene matriz de covarianza unidad. Si pasamos al vector transformado **Y =**

La técnica del "*fingerprint* óptimo" para el problema de detección en cambio climático y su enfoque tensorial en Hasselmann (1997)

A X, tenemos que la matriz de covarianza se transforma en $\mathbf{S}_Y = E[\mathbf{Y}^T \mathbf{Y}] = \mathbf{A}\, \mathbf{S}_X\, \mathbf{A}^T = \mathbf{A}\, \mathbf{A}^T$, y por tanto es un tensor doblemente contravariante según la regla que vimos en el parágrafo anterior, lo que se deduce también de la expresión de sus componentes, $S_X^{ij} = E[X^i X^j]$. Aplicando la ecuación de transformación de la pdf tras algunas operaciones se obtiene $f_Y(\mathbf{t}) = \det(\mathbf{S}_Y)^{-1/2}\ f_X(\mathbf{t}^T\ \mathbf{S}^{-1}_Y\ \mathbf{t})$, en la que se ha expresado el determinante jacobiano también en función de $\mathbf{S}_Y$ utilizando la regla del determinante producto. Vemos que, en efecto, en dicha pdf no aparece la matriz de transformación **A** explícitamente, tan solo la matriz de covarianza. La distribución de este tipo más importante con diferencia es la normal multivariante, en que la f_X es una exponencial. Esta se obtiene partiendo de N v.a. normales estándar e independientes estocásticamente, con lo que su pdf conjunta multivariante es el producto de las pdf de cada v.a., y en el exponente de la exponencial aparece la dependencia con **X.X** que demuestra la simetría esférica. Si se expresa ahora la pdf del vector **Y** obtenido por la transformación lineal se obtiene la forma habitual de la ley normal multivariante con matriz de covarianza arbitraria.

Se puede obtener la anterior ley de transformación para la pdf de **X** con simetría esférica razonando tensorialmente. En el sistema inicial ortonormal en que suponemos expresado el vector **X**, el escalar del que depende f_X es el producto escalar **t.t**, que al ser la matriz de covarianza $\mathbf{S}_X$ la identidad, se puede expresar como el escalar tensorial $S_{X,ij}\ t^i\ t^j$, por tanto invariante en un cambio de coordenadas (no olvidemos que en un sistema ortonormal las componentes contra- y covariantes coinciden). En este escalar aparece la matriz de covarianzas en forma doblemente covariante, y por tanto es igual al inverso de la matriz de covarianzas que es doblemente contravariante. Expresando el escalar usando la notación habitual matricial se obtiene $\mathbf{t}^T\ \mathbf{S}_Y^{-1}\ \mathbf{t}$ para el vector transformado **Y**, válida para cualquier sistema de coordenadas por su carácter tensorial.

La huella digital (*fingerprint*) óptima

El problema que se plantea Hasselmann es el de encontrar en un vector de datos climáticos (en general de muchas dimensiones) la sospecha fundada (de ahí el término huella digital) de que una hipotética señal climática ha dejado constancia de su presencia, por así decirlo. Si representamos en un sistema de referencia inicial, en consonancia con el espíritu tensorial, el vector aleatorio de N componentes de datos climáticos por **k**, el vector sospechoso de señal climática determinista por **g**, y el vector aleatorio de variabilidad climática natural por $\mathbf{k}^r$, postulamos la ecuación estocástica siguiente para **k**,

$$\mathbf{k} = d\,\mathbf{g} + \mathbf{k}^r \qquad (1)$$

Además Hasselmann postula para $\mathbf{k}^r$ una distribución normal multivariante. La constante d, si es distinta de cero, indicaría que efectivamente hay sospechas fundadas de la influencia de la señal climática postulada g (puede ser la señal que provoca el aumento de concentración de gases invernadero típicamente obtenida por simulación con modelos climáticos) en los datos observados climáticos **k**. Nos proponemos determinar la constante d de forma óptima buscando el vector **f** de forma que el producto escalar **f.k** distinga la presencia de la señal **g** del "ruido" de la variabilidad climática natural $\mathbf{k}^r$ de forma óptima según un criterio razonable. El criterio que propuso Hasselmann es maximizar la razón señal/ruido, expresada por $E[\mathbf{f}.\,\mathbf{k}]^2/\ E[\mathbf{f}.\,\mathbf{k}^r]^2$. Veamos cómo se puede resolver el anterior problema de forma elegante con un enfoque tensorial. Para ello abordamos el problema en un sistema de referencia privilegiado, que no es otro que un sistema de referencia ortonormal. Hemos visto que en ese sistema la distribución normal multivariante tiene una matriz de covarianza múltiplo de la unidad, y por tanto simetría esférica. Por tanto en este sistema no hay ninguna dirección privilegiada en la ec.(1) determinada por $\mathbf{k}^r$, la única dirección privilegiada es la determinada por **g**. Así pues el vector *fingerprint* óptimo ***f*** debe estar en la dirección de **g**. Esta deducción, que es evidente en dicho sistema, proporciona la solución general sin más que aplicar adecuadamente los tensores. Como la razón señal/ruido no cambia si se multiplica **f** por una constante, podemos tomar en este sistema ortonormal **f** = **g.** y por tanto d = **g. k.** Generalizando a una sistema de referencia arbitrario, en que la matriz de covarianza de la variabilidad natural es distinta en general de un múltiplo de la unidad, partimos de que hay que garantizar el carácter escalar tensorial de **f. k**, que es la cantidad que vamos a utilizar para la detección en el cociente señal ruido. En el sistema ortonormal esa cantidad es **g.k**, y replicando el razonamiento hecho arriba, lo expresamos en forma tensorial invariante usando la matriz de covarianza **S** en la forma $\mathbf{g}^T\ \mathbf{S}^{-1}\ \mathbf{k}$. Esta expresión, igualada a **f. k** demuestra que el vector huella dactilar óptimo es

$$\mathbf{f} = \mathbf{S}^{-1}\,\mathbf{g} \qquad (2)$$

donde se ha usado que **S** es simétrica. Esta es por tanto la solución completa al problema de detección por la huella dactilar óptima.

En la deducción anterior se ha postulado exclusivamente de la pdf de $\mathbf{k}^r$ que tiene simetría esférica en un sistema de referencia privilegiado, no tiene que tener necesariamente una distribución normal multivariante. Esta generalización podría quizá ser interesante si se plantea el problema de la detección de una señal climática en el contexto de extremos climáticos, en los que la hipótesis de normalidad no es razonable. Podría quizás proponerse una pdf para la variabilidad natural más adecuada con una cola más larga, pero eso sí, en la forma genérica exclusivamente dependiente de la matriz de covarianzas **S**:

$$f(\mathbf{t}) = A \det(\mathbf{S})^{-1/2}\ g(\mathbf{t}^T\ \mathbf{S}^{-1}\ \mathbf{t}) \qquad (3)$$

que garantiza como hemos visto la simetría esférica en un sistema privilegiado, donde A es una constante numérica de normalización y g un función arbitraria (que en el caso de la normal multivariante es una exponencial).

Referencias

✱ K. Hasselmann (1997): Multi-pattern fingerprint method for detection and attribution of climate change. Climate Dynamics (1997) 13 : 601Ð611