Entrega 1.1

Stefano Molina Martínez

Febrero 2016

1 ¿Qué es un producto tensorial?

Un producto tensorial es un método que funciona a través de juntar dos o más espacios vetoriales que forman un nuevo espacio vectorial que se define como el espacio de producto tensorial. El nuevo espacio resulta ser lineal en cada uno de los espacios que lo componen.

Una forma de pensar el producto tensorial de los vectores $v_1 \in V_1$ y $v_2 \in V_2$ es como funciones de una variable. El producto tensorial $v_1 \otimes v_2$ es una función en el plano cuyo valor en (x_1, x_2) es igual al producto $v_1(x_1)v_2(x_2)$.

Definición: Sean V y W espacios vectoriales. para $v_1, v_2 \in V$ y $w_1, w_2 \in W$, se dice que el símbolo $\otimes$ es el producto bilineal, que satisface para α_1, α_2 escalares:

$$(\alpha_1 v_1 + \alpha_2 v_2) \otimes w = \alpha_1 v_1 \otimes w + \alpha_2 v_2 \otimes w$$
$$v \otimes (\alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 v \otimes w_1 + \alpha_2 v \otimes w_2$$

A partir de lo anterior, $V \otimes W$ es el espacio tensorial de las combinaciones lineales $\sum \alpha_i v_i \otimes w_i$ para $v_i \in V$ y $w_i \in W$.
Nota: si $\{e_1, e_2, ..., e_n\}$ es una base en V y $\{e_1^{'}, e_2^{'}, ..., e_m^{'}\}$ una base en W, entonces su producto tensorial, $\{e_{ij} = e_i \otimes e_j^{'}\}$ es una base de $V \otimes W$, llamada *producto base*.

1.1 Componentes de un tensor

Sean V_1 y V_2 espacios vectoriales con bases $\{e_1, e_2, ..., e_n\}$ y $\{e_1^{'}, e_2{'}, ..., e_m^{'}\}$ respectivamente. Entonces $W = V_1 \otimes V_2$ es un espacio vectorial con dimensión mn con vectores base

$$\{e_{ij} = e_i \otimes e_j^{'}; i = 1, ...n; j = 1, ..., n\}$$

Cada punto $w \in W$ tiene una representación única como una combinación lineal de los vectores base:

$$w = w^{ij} e_i \otimes e_j^{'} = w^{ij} e_{ij}$$

Los w^{ij} son componentes del tensor w, cabe señalar que <u>no</u> es un tensor.

1.2 Transformación tensorial

Sean $\{e_1, e_2, ..., e_n\}$ una base en V_1 y $\{e_1^{'}, e_2^{'}, ..., e_m^{'}\}$ una base una base en V_2 y sean w_{ij} los componentes del tensor en $W = V_1 \otimes V_2$ con respecto al producto base $\{e_{ij} = e_i \otimes e_j^{'}$.

Ahora, consideremos el efecto de una nueva base $\{\bar{e_i} \in V_1\}$ y $\{\bar{e^{'}} \in V_2\}$, se escriben:

$$\bar{e_r} = K_r^i e_i \qquad \bar{e_r^{'}} = K_r^{'i} e_i^{'}$$

A partir de los términos de las bases originales. El producto base en W se transforma de $\{e_{ij}\}$ en $\{\bar{e_{ij}}\}$ como sigue

$$\bar{e_{rs}} = K_r^i K_s^{'j} e_i \otimes e_j^{'} = K_r^i K_s^{'j} e_{ij}$$

Ahora, en la notación $\bar{w^{ij}}$, el componente de w con respecto a la nueva base, tenemos:

$$w = \bar{w^{rs}} \bar{e_{rs}} = \bar{w^{rs}} K_r^i K_s^{'j} e_{ij} = w^{ij} e_{ij}$$

1.3 Producto tensorial

Los componentes de w con respecto a la nueva base son:

$$\bar{w}^{rs} = L_i^r L_j^{'s} w^{ij}$$

donde L es la matriz inversa de K y $L^{'}$ la matriz inversa de $K^{'}$. Esta es la regla de transformación tensorial para los componentes de un vector contravariante en $V_1 \otimes V_2$. En la mayoría de las aplicaciones, sólo se ocupa un espacio V, en el cual se desarrollan sus productos tensoriales $V^{\otimes k}$. Además $V_1 = V_2$, $m = n$, $K = K^{'}$ y $L = L^{'}$.

1.4 Multiplicación tensorial y contracción

Operativamente, un tensor es un arreglo de número o funcionales que obedece una regla de transformación multi-lineal particular bajo un cambio de base en el espacio trabajado. Suponer que se trabaja en un solo espacio V y que los vectores de cambio de base están dados por

$$\bar{e}_i = a_i^r e_r \quad o \quad e_r = b_r^i \bar{e}_i$$

donde b_r^i es la matriz inversa de a_i^r. Estas matrices pertenecen a un grupo relevante al problema, usualmente al grupo lineal general, pero en circunstancias apropiadas al grupo ortogonal o al simétrico. Un arreglo $w = w^{ijk}$ cuyos valores en el nuevo sistema de coordenadas está dado por

$$\bar{w}^{ijk} = b_r^i b_s^j b_t^k w^{rst}$$

que se dice que es un vector contravariante de orden tres. Técnicamente w^{ijk} son los componentes de un punto o vector en $V^{\otimes 3}$. De igual manera, un arreglo γ_{ij} que se transforma por la regla

$$\gamma_{ij} = a_i^r a_j^s \gamma_{rs}$$

se dice que es covariante de orden dos. El producto directo de dos tensores es un tensor, por ejemplo,

$$\Psi_{rs}^{ijk} = w^{ijk}\gamma_{rs}$$

es un tensor de orden covariante dos y orden contravariante tres. Sigue la regla de transformación para un vector mixto.
La multiplicación directa resulta en un tensor de orden aumentado. En general, los índices pueden tener diferentes rangos y pueden referirse a espacios no relacionados. Sin embargo, cuando un superíndice y un subíndice se refieren a components en un espacio y su dual, necesariamente tienen el mismo rango. En esos casos, es válido sumar sobre los índices, a este proceso se le conoce como contracción, resultando en un tensor de orden menor. En el ejemplo anterior,

$$\Phi^i = \Psi_{rs}^{irs} = w^{irs}\gamma_{rs}$$

es un vector contravariante, o el vector componente de un punto en V.

El cálculo tensorial es la simple aplicación de estas reglas de multiplicación y contracción. Uno de los casos más importantes de contracción tensorial ocurre cuando resulta en un tensor de orden cero, conocido también como escalar o invariante.Esas cifras son importantes porque su valor es el mismo en todo sistema de coordenadas del grupo.

1.5 Transformación Lineal de espacios vectoriales

El espacio de transformaciones lineales de U de dimensión n a V de dimensión m es el conjunto de funciones g en U con rango en V que satisfacen:

$$g(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 g(x_1) + \alpha_2 g(x_2)$$

Consecuentemente, el conjunto de esas transformaciones constituye un espacio vectorial de dimensión mn. Los componentes de g con respecto a los vectores base de U y V forman una matrix $\{g_r^i\}$ de $m\ x\ n$: los componentes de $y = g(x)$ son dados por

$$y^i = g_r^i x^r$$

Así, g_r^ison los componentes del tensor mixto g en el espacio $U \otimes V^*$.

2 Teoría elemental de *acumulantes*

Nota: Un *acumulante* es un conjunto de cantidades que sirven como alternativa para calcular los momentos de alguna distribución, éstos son de gran ayuda debido a su simplicidad.

Si la variable aleatoria X tiene una función de densidad $f_X(x)$ definida en $-\infty < x < \infty$, entonces la esperanza de la función $g(x)$ es

$$E\{g(X)\} = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

En el caso discreto, la integral se susituye por una suma en los valores admisibles de la función.

2.1 Definiciones

Sea X v.a. cuyos componentes son $X^1, ..., X^p$.

Considerar la serie infinita

$$M_X(\xi) = 1 + \xi_i \kappa^i + \xi_i \xi_j \kappa^{ij}/2! + \xi_i \xi_j \xi_k \kappa^{ijk}/3! + ...$$

La cual se asume que converge para toda $|\xi|$ suficientemente pequeña, esta suma también se puede escribir de la siguiente forma

$$M_X(\xi) = E\{exp(\xi_i X^i)\}$$

y los momentos son las derivadas parciales de $M_X(\xi)$ evaluadas en $\xi = 0$.

Los *acumulantes* son definidos a través de su función generadora,

$$K_X(\xi) = log M_X(\xi)$$

la cual tiene una expansión

$$K_X(\xi) = \xi_i \kappa^i + \xi_i \xi_j \kappa^{i,j}/2! + \xi_i \xi_j \xi_k \kappa^{i,j,k}/3! + ...$$

Esta expansión define implícitamente a todos los *acumulantes*, denotados como $\kappa_i, \kappa^{i,j}, \kappa^{i,j,k}...$ en términos de sus momentos correspondientes. La forma en la que se van a diferenciar los *acumulantes* de los momentos en por sus superíndices, en los primeros éstos aparecen separados con comas, mientras que en los últimos no, también hay que resaltar que en el caso de el primer momento y el primer *acumulante*, ambos son iguales.

2.2 *Acumulantes* y momentos

Los *acumulantes* y los momentos están fuertemente relacionados, la demostración de relación termina en la fórmula recursiva que se describa de la siguiente manera

$$\kappa_n = \mu_n^{'} - \sum_{m=1}^{n-1} \binom{n-1}{m-1} \kappa_m \mu_{n-m}^{'}$$

2.3 Transformaciones lineales y afines

Si Y es una función lineal de X, se puede escribir

$$Y^r = a_i^r X^i$$

donde a_i^r es un arreglo de constantes. Se puede observar que los momentos de Y son

$$a_i^r, \quad a_i^r a_j^s \kappa^{ij}, \quad a_i^r a_j^s a_k^t \kappa^{ijk}, ...$$

Mientras que los *acumulantes* son

$$a_i^r, \quad a_i^r a_j^s \kappa^{i,j}, \quad a_i^r a_j^s a_k^t \kappa^{i,j,k}, ...$$

También se observar que bajo transformaciones lineales, tanto los momentos como los *acumulantes* se transforman como tensores contravariantes. Hay que señalar que la matriz a_i^r no necesita ser de rango completo.
Las transformaciones afines conllevan un cambio de origen de acuerdo a la ecuación

$$Y^r = a^r + a_i^r X^i$$

Los *acumulantes* de Y, que se probarán más adelante, son

$$a^r + a_i^r \kappa^i \quad a_i^r a_j^s \kappa^{i,j} \quad a_i^r a_j^s a_k^t \kappa^{i,j,k} ...$$

El cambio de origen afecta únicamente al primer *acumulante*, también llamado vector media. Por esta razón, los *acumulantes* son llamados a veces semi-invariantes. Por otro lado, los momentos de Y son

$$a^r + a_i^r \kappa^i,$$
$$a^r a^s + a^r a_i^s \kappa^i[2] + a_i^r a_j^s \kappa^{ij},$$
$$a^r a^s a^t + a^r a^s a_i^t \kappa^i[3] + a^r a_i^s a_j^t \kappa^{ij}[3] + a_i^r a_j^s a_k^t \kappa^{ijk}$$

y así sucesivamente, donde $a^r a_i^s \kappa^i[2] = a^r a_i^s \kappa^i + a^s a_i^r \kappa^i$. Así, a diferencia de los *acumulantes*, los momentos no se transforman de una forma sencilla bajo transformaciones afines de coordenadas.

2.4 Interpretación de *Acumulantes*

Se va a tratar con *acumulantes* univariados y posteriormente se llevará la intuición de los multivariados.

El primer *acumulante* de X, κ_1 es la media de la distribución, mientras que el segundo, κ_2 es la varianza.
El tercer *acumulante* de X es una medida de asimetría en el mismo sentido que $\mu_3 = E(X-\mu_1)^3$ es cero si X está distribuida simétricamente. Por su parte, $\kappa_3 = 0$ no implica, por sí solo, simetría en la distribución: para garantizar simetría se necesita que todos los *acumulantes* nones desaparezcan y que la distribución esté determinada por sus momentos. A partir del tercero y todos los *acumulantes* estandarizados de orden mayor dados por $\rho_r = \kappa_r/\kappa_2^{r/2}$ pueden ser interpretados como medidas resumidas de alejamiento de normalidad en el sentido de que si X es normal, todos los *acumulantes* de orden tres o más son cero.

La interpretación más sencilla de *acumulantes* bivariados es en término de independencia. Si X^1 y X^2 son independientes, entonces todos los *acumulantes* mixtos que incluyen solo X^1 y X^2 son cero. Por lo tanto, $\kappa^{1,2} = \kappa^{1,1,2} = \kappa^{1,2,2} = ... = 0$ o siendo más concisos usando notación de potencia, $\kappa_{rs} = 0 \; \forall \; r, s \geq 1$. Dado que los momentos determinan la distribución conjunta, lo contrario también es verdadero, eso quiere decir que si $\kappa_{rs} = 0 \; \forall \; r, s \geq 1$, entonces X^1 y X^2 son independientes. La sugerencia es que si $\kappa_{rs} = 0$ para $r, s = 1, ...t$ entonces X^1 y X^2 son aproximadamente independientes en algún sentido.

2.5 *Acumulantes* condiconales

Suponer que son dados los *acumulantes* condicionales conjuntos de las variables aleatorias condicionales $X^1, ..., X^P$ para algún evento A. La forma de combinar los *acumulantes* condiconales para obtener los *acumulantes* conjuntos no-condiconales es

$$E(X^1X^2...) = E_A E(X^1X^2...|A)$$

En otras palabras, los momentos no-condicionales son un promedi de los momentos condicionales. Pero, no es difícil mostrar que la covarianza de X^i y X^j satisface

$$\kappa^{i,j} = E_A\{cov(X^i, X^j|A)\} + cov_A\{E(X^i|A), E(X^j|A)\}.$$

Para ver cómo esto se generaliza a *acumulantes* de cualquier orden, se denotan los *acumulantes* condicionales $\lambda^i, \lambda^{i,j}, \lambda^{i,,j,k}$ y se usa la identidad que conecta con la función generadora de momentos

$$M_X(\xi) = E_A M_{X|A}(\xi)$$

La expansión de esta identidad y la comparación de sus coeficientes da

$$\kappa^i = E_A\{\lambda^i\}$$

$$\kappa^{i,j} + \kappa^i \kappa^j = E_A\{\lambda^{ij} + \lambda^i \lambda^j\}$$

$$\kappa^{i,j,k} + \kappa^i \kappa^{j,k}[3] + \kappa^i \kappa^j \kappa^k = E_A\{\lambda^{i,j,k} + \lambda^i \lambda^{j,k}[3] + \lambda^i \lambda^j \lambda^k\}$$

La primera expresión de esta sección para la covarianza no-condicional se sigue de la segunda expresión de arriba. Usando este resultado en la tercera expresión, se tiene

$$\kappa^{i,j,k} = E(\lambda^{i,j,k}) + \kappa_2(\lambda^i, \lambda^{j,k})[3] + \kappa_3(\lambda^i, \lambda^j, \lambda^k).$$