UTILIZACIÓN DE MÉTRICAS RIEMANNIANAS EN ANALISIS DE DATOS MULTIDIMENSIONALES Y SU APLICACIÓN A LA BIOLOGÍA

JOSE Mª OLLER SALA

BARCELONA, 25 de NOVIEMBRE de 1982.

Esta operación es el producto tensorial externo. Si $A^{\mu_1 \dots \mu_p}_{\lambda_1 \dots \lambda_q}$ y $B^{\alpha_1 \dots \alpha_r}_{\beta_1 \dots \beta_s}$ son las componentes de dos tensores mixtos de ordenes p+q y r+s respectivamente con p y r índices contravariantes y q y s índices covariantes respectivamente, entonces $C^{\mu_1 \dots \mu_p \alpha_1 \dots \alpha_r}_{\lambda_1 \dots \lambda_q \beta_1 \dots \beta_s}$ definido por:

$$C^{\mu_1 \dots \mu_p \alpha_1 \dots \alpha_r}_{\lambda_1 \dots \lambda_q \beta_1 \dots \beta_s} = A^{\mu_1 \dots \mu_p}_{\lambda_1 \dots \lambda_q} \; B^{\alpha_1 \dots \alpha_r}_{\beta_1 \dots \beta_s} \tag{14}$$

$$\mu_1 \dots \mu_p, \lambda_1, \dots, \lambda_q, \alpha_1, \dots, \alpha_r, \beta_1, \dots, \beta_s = 1, \dots, n$$

son las componentes de un tensor de orden p+r+q+s con p+r índices contravariantes y q+s índices covariantes.

Otra de las operaciones a tener en cuenta es la contracción tensorial. Si $A^{\mu_1 \dots \mu_p}_{\lambda_1 \dots \lambda_q}$ son las componentes de un tensor mixto, con $p \geq 1$ índices contravariantes y $q \geq 1$ índices covariantes, podemos igualar un índice contravariante con otro covariante y sumar, respecto ambos, desde 1 hasta n, (tensor definido en variedad n-dimensional) obteniendo si efectuamos la operación respecto el j-ésimo índice contravariante y el k-ésimo índice covariante, el conjunto de n^{p+q-2} funciones, $B^{\mu_1 \dots \mu_{j-1} \mu_{j+1} \dots \mu_p}_{\lambda_1 \dots \lambda_{k-1} \lambda_{k+1} \dots \lambda_q}$ definidas (según convenio de sumación de índices repetidos) por:

$$B^{\mu_1 \dots \mu_{j-1} \mu_{j+1} \dots \mu_p}_{\lambda_1 \dots \lambda_{k-1} \lambda_{k+1} \dots \lambda_q} = A^{\mu_1 \dots \mu_{j-1} \alpha \mu_{j+1} \dots \mu_p}_{\lambda_1 \dots \lambda_{k-1} \alpha \lambda_{k+1} \dots \lambda_q} \tag{15}$$

$$\mu_1 \dots \mu_{j-1}, \mu_{j+1} \dots \mu_p, \dots, \lambda_1 \dots {}_{k-1} \lambda_{k+1} \dots \lambda_q = 1, \dots, n$$

son las componentes de un tensor mixto de orden p+q-2, con p-1 índices contravariantes y q-1 índices covariantes.

Si es posible efectuar un producto tensorial externo y seguidamente una contracción tensorial, al resultado combinado se le denomina producto interno (de tensores). Por ejemplo, si A_{ij} y B^k son las componentes de un tensor covariante de segundo orden y de un tensor contravariante de primer orden, podemos efectuar un producto externo y una contracción respecto el primer índice covariante de A_{ij}, obteniendo C_j:

$$C_j = A_{ij}\ B^i \qquad j = 1,\dots,n \tag{16}$$

donde C_j son las componentes de un tensor covariante de primer orden.

Nótese que la imagen de un vector por una forma lineal es de hecho la contracción tensorial de un tensor mixto de segundo orden, producto tensorial externo de dos tensores de primer orden, uno covariante y otro contravariante, es pues un ejemplo de producto interno de tensores:

$$A^{\mu}B_{\mu} = \text{invariante} \tag{17}$$

Un resultado frecuentemente utilizado es el siguiente, conocido con el nombre de ley del cociente. Sea $A(i_1,\dots,i_r)$ $i_1,\dots,i_r=1,\dots,n$, un conjunto de n^r funciones definidas en una variedad n-dimensional y sean B^{α} las componentes de un tensor contravariante. Si se cumple que:

$$C^{j_1\ldots j_q}_{k_1\ldots k_p} = A(\alpha, i_2, \ldots, i_r)B^{\alpha} \tag{18}$$

donde $C^{j_1\ldots j_q}_{k_1\ldots k_p}$ son las componentes de un tensor de orden p+q=r-1, con q índices contravariantes y p índices covariantes, entonces $A(i_1,\ldots,i_r)$ son las componentes de un tensor $A^{j_1\ldots j_q}_{\alpha k_1\ldots k_p}$ de orden r, con q índices contravariantes y p+1 índices covariantes. Análogamente, si sustituimos en (18) B^{α} por B_{α}, entonces $A(i_1\ldots i_r)$ serian las componentes de un tensor $A^{\alpha j_1\ldots j_q}_{k_1\ldots k_p}$, de orden r, con q+1 índices contravariantes y p índices covariantes.

Cuando al intercambiar dos índices de las componentes de un tensor no queda alterado el valor de éstas, por ejemplo: $A_{ij} = A_{ji}$, diremos que dicho tensor es simétrico respecto estos dos índices. Cuando al intercambiar los índices cambia el signo de las componentes, $A_{ij}=-A_{ji}$ diremos que el tensor es antisimétrico respecto estos dos índices. Nótese que la simetría y antisimetría no dependen del sistema de coordenadas con que trabajemos.

Es posible considerar conjuntos de n^{p+q} funciones definidas en una variedad n-dimensional, $A^{\mu_1\ldots\mu_p}_{\lambda_1\ldots\lambda_q}(x^1,\ldots,x^n)$, que al efectuar la transformación (5) se convierten en $\bar{A}^{\mu_1\ldots\mu_p}_{\lambda_1\ldots\lambda_q}(\bar{x}^1,\ldots,\bar{x}^n)$, de acuerdo con:

$$\bar{A}^{\mu_1\ldots\mu_p}_{\lambda_1\ldots\lambda_q} = \left|\frac{\partial x^i}{\partial \bar{x}^j}\right|^w \frac{\partial \bar{x}^{\mu_1}}{\partial x^{\alpha_1}}\cdot\ldots\cdot\frac{\partial \bar{x}^{\mu_p}}{\partial x^{\alpha_p}}\cdot\frac{\partial x^{\beta_1}}{\partial \bar{x}^{\lambda_1}}\cdot\ldots\cdot\frac{\partial x^{\beta_q}}{\partial \bar{x}^{\lambda_q}} A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q} \tag{19}$$

$$\mu_1,\ldots,\mu_p,\lambda_1,\ldots,\lambda_q = 1,\ldots,n$$

donde $|\partial x^i/\partial \bar{x}^j|$ es el determinante jacobiano de la transformación (5).

En estas condiciones diremos que las $A^{\alpha_1\ldots\alpha_p}_{\beta_1\ldots\beta_q}$ son las componentes de un tensor relativo de peso w. Nótese que si $w=0$, obtenemos un tensor mixto, tal como lo hemos definido en (11). En el caso particular que $w=1$, $p=q=0$,

$$\bar{A} = A \left| \frac{\partial x^i}{\partial \bar{x}^j} \right| \tag{20}$$

llamaremos a A densidad escalar. Ejemplo de densidad escalar, son las funciones de densidad de las variables aleatorias absolutamente continuas. Es posible definir operaciones, análogas a las que hemos visto para tensores mixtos (tensores relativos de peso cero), para tensores relativos de peso w cualquiera.

2.4. GEOMETRIA EN UNA VARIEDAD

A continuación vamos a intentar aplicar lo visto hasta ahora al estudio de las propiedades geométricas de una variedad.

Si consideramos un espacio euclídeo referido a un sistema de coordenadas cartesianas ortonormales, la distancia, al cuadrado, entre dos puntos $P(x^1,\ldots,x^n)$ y $P'(x^1+dx^1,\ldots,x^n+dx^n)$ viene dada por:

$$ds^2 = dx^i dx^i \tag{21}$$

donde llamaremos a ds elemento de longitud, elemento de arco o elemento de línea. Si efectuamos el cambio de coordenadas (5), teniendo en cuenta que:

$$dx^i = \frac{\partial x^i}{\partial \bar{x}^j}\, d\bar{x}^j \tag{22}$$

el cuadrado del elemento de arco será:

$$ds^2 = \frac{\partial x^i}{\partial \bar{x}^j}\frac{\partial x^i}{\partial \bar{x}^k}\, d\bar{x}^j\, d\bar{x}^k \tag{23}$$

si llamamos:

$$g_{jk} = \frac{\partial x^i}{\partial \bar{x}^j}\frac{\partial x^i}{\partial \bar{x}^k} \tag{24}$$

puesto que ds^2 es un invariante, ya que la distancia entre dos puntos de una variedad no debe depender del sistema de coordenadas, podemos demostrar que g_{jk} son las componentes de un tensor covariante de segundo orden, simétrico, que denominaremos tensor métrico o tensor fundamental covariante. Hasta ahora hemos introducido el tensor métrico para poder calcular el elemento de arco y por integración del mismo la distancia entre dos puntos de un espacio euclídeo, bajo un sistema de coordenadas arbitrario. Nótese que, en estas condiciones existirá, dado g_{jk}, una cierta transformación de coordenadas que reduzca al tensor métrico al tensor:

$$\delta_{ij} = \begin{cases} 1 & \Leftrightarrow \quad i=j \\ 0 & \Leftrightarrow \quad i\neq j \end{cases} \tag{25}$$

Riemann generalizó estas ideas, definiendo en una variedad cualquiera, el elemento de arco, a través de la forma cuadrática diferencial:

$$ds^2 = g_{\mu\nu} \; dx^\mu \; dx^\nu \tag{26}$$

siendo $g_{\mu\nu}$ (funciones de las coordenadas) las componentes de un tensor covariante, simétrico, definido positivo y arbitrario. En cada punto el tensor $g_{\mu\nu}$ define el producto escalar del espacio tangente a dicho punto de la variedad. La arbitrariedad de las $g_{\mu\nu}$ impide garantizar la existencia de una transformación de coordenadas que reduzca al tensor métrico a un tensor de la forma (25). Si $g_{\mu\nu}$ es definido positivo en toda la variedad, diremos que ésta es una variedad Riemanniana. Caso de ser semidefinido positivo, la denominaremos semiRiemanniana.

La longitud entre dos puntos de la variedad medida sobre una curva σ, viene dada por:

$$\ell_{\sigma_{ab}} = \int_{t_a}^{t_b} \sqrt{g_{\mu\nu} \frac{dx^\mu}{dt} \frac{dx^\nu}{dt}} \; dt \tag{27}$$

siendo t un parámetro de la curva σ. La distancia entre dos puntos P y Q se define como el ínfimo del conjunto de longitudes entre P y Q medidas sobre todas las posibles curvas que unen ambos puntos. Se denominan geodésicas a las curvas de longitud mínima que unen dos puntos. Las geodésicas siempre existen localmente, y es posible asegurar que si existe una curva de longitud mínima ésta es geodésica. Las condiciones de existencia de curvas de longitud mínima pueden verse en Hicks (1974).

Si existe una transformación de coordenadas tal que el tensor métrico quede reducido a un tensor constante (y por tanto con una transformación lineal podamos reducir al tensor métrico a la forma (25)), diremos que el espacio es euclídeo. En estas condiciones, bajo un sistema de coordenadas conveniente, localmente, podremos calcular la distancia entre dos puntos de coordenadas $(x_A^1,\dots,x_A^n)$ y $(x_B^1,\dots,x_B^n)$ mediante la conocida:

$$d_{AB} = \sqrt{(x_A^i - x_B^i)(x_A^i - x_B^i)} \tag{28}$$

posteriormente indicaremos una caracterización de la euclidianidad del espacio.

En cuanto al elemento de volumen, que en un espacio euclídeo referido a un sistema de coordenadas cartesianas ortonormales viene dado por:

$$dV = dx^1 \cdot \ldots \cdot dx^n \tag{29}$$

en una variedad de Riemann puede demostrarse que viene dada por:

$$dV = \sqrt{|g_{\mu\nu}|}\; dx^1 \cdot \ldots \cdot dx^n \tag{30}$$

donde $|g_{\mu\nu}|$ es el determinante del tensor métrico.

A partir del tensor métrico, $g_{\mu\nu}$, con $|g_{\mu\nu}| \neq 0$, podemos definir un tensor contravariante, de segundo orden y simétrico, a partir de la ecuación tensorial:

$$g_{\mu\nu}\, g^{\nu\alpha} = \delta_\mu^\alpha \tag{31}$$

siendo δ^{α}_{μ} un tensor mixto constante, definido por:

$$\delta^{\alpha}_{\mu} = \begin{cases} 1 & \Leftrightarrow \quad \alpha = \mu \\ 0 & \Leftrightarrow \quad \alpha \neq \mu \end{cases} \tag{32}$$

Al tensor $g^{\nu\alpha}$ así definido se le denomina el recíproco del tensor métrico, o tensor fundamental contravariante. Ambos tensores, $g_{\mu\nu}$, $g^{\mu\nu}$ se les denomina tensores fundamentales. Nótese que:

$$g^{\mu\nu} = \frac{G^{\nu\mu}}{|g_{\mu\nu}|} \tag{33}$$

donde $G^{\nu\mu}$ es el adjunto del elemento $g_{\nu\mu}$ y $|g_{\mu\nu}|$ el determinante de $(g_{\mu\nu})$.

Un tensor obtenido por el proceso de multiplicación interna de un cierto tensor $A^{\mu_1 \ldots \mu_p}_{\lambda_1 \ldots \lambda_q}$ por $g_{\mu\nu}$ ó $g^{\mu\nu}$ se denomina tensor asociado al tensor dado. Por ejemplo, si A_{ij} son las componentes de un tensor covariante de segundo orden, resulta:

$$g^{\alpha\beta} A_{\beta\gamma} = A^{\alpha}_{\cdot\gamma} \tag{34}$$

análogamente si B^{ij} son las componentes de un tensor contravariante de segundo orden:

$$g_{\alpha\beta} B^{\beta\gamma} = B_{\alpha}^{\cdot\gamma} \tag{35}$$

$A^{\alpha}_{\cdot\gamma}$ es un tensor mixto asociado a $A_{\alpha\gamma}$ y $B_{\alpha}^{\cdot\gamma}$ es también un tensor mixto asociado a $B^{\alpha\gamma}$. Observemos que el proceso permite "subir" y "bajar" índices a un tensor cualquiera, según utilicemos $g^{\alpha\beta}$ ó $g_{\alpha\beta}$ respectivamente.

2.5. SIMBOLOS DE CRISTOFFEL. DERIVADA COVARIANTE.

Algunas combinaciones de derivadas parciales del tensor fundamental, $g_{\mu\nu}$, juegan un papel importante en el desarrollo del cálculo tensorial, son conocidas como símbolos de Christoffel.

Al conjunto de funciones:

$$[i\ j,k] = \frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j} + \frac{\partial g_{jk}}{\partial x^i} - \frac{\partial g_{ij}}{\partial x^k}\right) \qquad (36)$$

$$i,j,k = 1,\dots,n$$

las denominaremos símbolos de Christoffel de primera especie, mientras que al conjunto de funciones definidas a través de:

$$\Gamma^k_{ij} = g^{k\alpha}\,[ij,\alpha] \qquad (37)$$

$$i,j,k = 1,\dots,n$$

se las denominan símbolos de Christoffel de segunda especie. Dada la simetría de $g_{\mu\nu}$ hay $n^2(n+1)/2$ símbolos de Christoffel independientes de cada especie. A partir de la definición se sigue:

$$\begin{aligned} [ij,k] &= [ji,k] \\ \Gamma^k_{ij} &= \Gamma^k_{ji} \end{aligned} \qquad i,j,k=1,\dots,n \qquad (38)$$

Puede demostrarse que los símbolos de Christoffel se anulan (en todo su dominio) si el tensor métrico es un tensor constante (en todo su dominio).

Hay que resaltar que los símbolos de Christoffel no tienen caracter tensorial. En efecto, si designamos por ${}_{x}[ij,k]$ a los símbolos de Christoffel de primera especie referidos a un sistema de coordenadas x, y ${}_{\bar{x}}[ij,k]$ a los mismos referidos al sistema de coordenadas $\bar{x}$, las fórmulas de transformación, al efectuar el cambio (5), vienen dadas por:

$$ {}_{\bar{x}}[ij,k] = \frac{\partial x^{\alpha}}{\partial \bar{x}^{i}} \frac{\partial x^{\beta}}{\partial \bar{x}^{j}} \frac{\partial x^{\gamma}}{\partial \bar{x}^{k}} \, {}_{x}[\alpha\beta,\gamma] + \frac{\partial^{2} x^{\alpha}}{\partial \bar{x}^{i} \partial \bar{x}^{j}} \frac{\partial x^{\beta}}{\partial \bar{x}^{k}} g_{\alpha\beta} \qquad (39) $$

$$ i,j,k=1,\ldots,n $$

que nos muestran que $[ij,k]$ no se transforman como un tensor, salvo que la transformación de coordenadas sea afin, y por tanto, se anule el segundo término de (39).

En cuanto a los símbolos de Christoffel de segunda especie, si los designamos como ${}_{x}\Gamma^{k}_{ij}$ y ${}_{\bar{x}}\Gamma^{k}_{ij}$ según los refiramos al sistema x ó $\bar{x}$, al efectuar el cambio (5), las fórmulas de transformación vienen dadas por:

$$ {}_{\bar{x}}\Gamma^{k}_{ij} = \frac{\partial \bar{x}^{k}}{\partial x^{\gamma}} \frac{\partial x^{\alpha}}{\partial \bar{x}^{i}} \frac{\partial x^{\beta}}{\partial \bar{x}^{j}} \, {}_{x}\Gamma^{\gamma}_{\alpha\beta} + \frac{\partial^{2} x^{\alpha}}{\partial \bar{x}^{i} \partial \bar{x}^{j}} \frac{\partial \bar{x}^{k}}{\partial x^{\alpha}} \qquad (40) $$

$$ k,i,j = 1,\ldots,n $$

por tanto no se transforman como un tensor, salvo que la transformación de coordenadas sea afin, ya que entonces se anula la segunda parte de (40).

El sistema (40) puede resolverse respecto $\dfrac{\partial^2 x^\alpha}{\partial \bar{x}^i \partial \bar{x}^j}$, obteniéndose:

$$\frac{\partial^2 x^m}{\partial \bar{x}^i \partial \bar{x}^j} = {}_{\bar{x}}\Gamma^\gamma_{ij} \frac{\partial x^m}{\partial \bar{x}^\gamma} - {}_{x}\Gamma^m_{\alpha\beta} \frac{\partial x^\alpha}{\partial \bar{x}^i} \frac{\partial x^\beta}{\partial \bar{x}^j} \tag{41}$$

$$i,j,m = 1,\ldots,n$$

Anteriormente hemos indicado que las derivadas parciales de una función escalar (invariante) son las componentes de un tensor covariante de primer orden, ya que:

$$\frac{\partial f}{\partial \bar{x}^i} = \frac{\partial f}{\partial x^\alpha} \frac{\partial x^\alpha}{\partial \bar{x}^i} \tag{42}$$

Examinemos si las derivadas parciales de las componentes de un tensor covariante tienen caracter tensorial. Sean A_i y $\bar{A}_i$ las compo-- nentes de un tensor covariante de primer orden referidos al sistema x y al $\bar{x}$ respectivamente, derivando parcialmente obtenemos:

$$\frac{\partial \bar{A}_i}{\partial \bar{x}^j} = \frac{\partial}{\partial \bar{x}^j} \left(A_\alpha \frac{\partial x^\alpha}{\partial \bar{x}^i}\right) = \frac{\partial A_\alpha}{\partial x^\beta} \frac{\partial x^\beta}{\partial \bar{x}^j} \frac{\partial x^\alpha}{\partial \bar{x}^i} + A_\alpha \frac{\partial^2 x^\alpha}{\partial \bar{x}^j \partial \bar{x}^i} \tag{43}$$

El segundo término de la expresión impide que $\dfrac{\partial A_\alpha}{\partial x^\beta}$ sean las componentes de un tensor covariante de segundo orden. Podemos expresar el segundo término de (43) a partir de (41), obteniendo:

$$\frac{\partial \bar{A}_i}{\partial \bar{x}^j} = \frac{\partial x^\alpha}{\partial \bar{x}^i}\frac{\partial x^\beta}{\partial \bar{x}^j}\frac{\partial A_\alpha}{\partial x^\beta} + {}_{\bar{x}}\Gamma^\gamma_{ij}\frac{\partial x^\alpha}{\partial \bar{x}^\gamma}A_\alpha - {}_x\Gamma^\gamma_{\alpha\beta}\frac{\partial x^\alpha}{\partial \bar{x}^i}\frac{\partial x^\beta}{\partial \bar{x}^j}A_\gamma, \tag{44}$$

pero como:

$$\bar{A}_\gamma = \frac{\partial x^\alpha}{\partial \bar{x}^\gamma}A_\alpha \tag{45}$$

resulta:

$$\frac{\partial \bar{A}_i}{\partial \bar{x}^j} - {}_{\bar{x}}\Gamma^\gamma_{ij}\bar{A}_\gamma = \frac{\partial x^\alpha}{\partial \bar{x}^i}\frac{\partial x^\beta}{\partial \bar{x}^j}\left(\frac{\partial A_\alpha}{\partial x^\beta} - {}_x\Gamma^\gamma_{\alpha\beta}A_\gamma\right) \tag{46}$$

El sistema de n^2 funciones:

$$\frac{\partial A_\alpha}{\partial x^\beta} - \Gamma^\gamma_{\alpha\beta}A_\gamma \qquad \alpha,\beta = 1,\dots,n \tag{47}$$

se transforman según la ley de transformación de un tensor covariante de segundo orden.

Por definición, denominamos derivada covariante x^ν de la componente A_μ, respecto del tensor fundamental $g_{\mu\nu}$, y la representaremos por $A_{\mu,\nu}$ a:

$$A_{\mu,\nu} = \frac{\partial A_\mu}{\partial x^\nu} - \Gamma^\gamma_{\mu\nu}A_\gamma \tag{48}$$

$A_{\mu,\nu}$ son las componentes de un tensor covariante de segundo orden, que denominaremos derivada covariante del tensor A_μ.

Análogamente se define la derivada covariante x^ν de la componente A^μ, respecto el tensor fundamental $g_{\mu\nu}$, y la representaremos por $A^\mu{}_{,\nu}$ a:

$$A^\mu{}_{,\nu} = \frac{\partial A^\mu}{\partial x^\nu} + \Gamma^\mu_{\gamma\nu} A^\gamma \qquad (49)$$

y es posible demostrar que los $A^\mu{}_{,\nu}$ son los componentes de un tensor mixto de segundo orden, que denominaremos derivada covariante del tensor A^μ.

La derivada covariante de tensores de ordenes superiores se define de forma parecida:

$$A^{\mu_1\ldots\mu_p}_{\mu_{p+1}\ldots\mu_{r,\nu}} = \frac{\partial A^{\mu_1\ldots\mu_p}_{\mu_{p+1}\ldots\mu_r}}{\partial x^\nu} + \Gamma^{\mu_1}_{\alpha\nu} A^{\alpha\mu_2\ldots\mu_p}_{\mu_{p+1}\ldots\mu_r} + \ldots + \Gamma^{\mu_p}_{\alpha\nu} A^{\mu_1\ldots\mu_{p-1}\alpha}_{\mu_{p+1}\ldots\mu_r} -$$

$$- \Gamma^\alpha_{\mu_{p+1}\nu} A^{\mu_1\ldots\mu_p}_{\alpha\mu_{p+2}\ldots\mu_r} - \ldots - \Gamma^\alpha_{\mu_r\nu} A^{\mu_1\ldots\mu_p}_{\mu_{p+1}\ldots\mu_{r-1}\alpha} \qquad (50)$$

y es posible demostrar que $A^{\mu_1\ldots\mu_p}_{\mu_{p+1}\ldots\mu_r,\nu}$ son las componentes de un tensor mixto de orden r+1, con p índices contravariantes y r-p+1 índices covariantes, denominado derivada covariante de $A^{\mu_1\ldots\mu_p}_{\mu_{p+1}\ldots\mu_r}$.

Dado un campo tensorial, la variación de las componentes de dicho campo a lo largo de una línea coordenada no refleja la variación del tensor, puesto que aún cuando éste fuese constante en toda la variedad, las derivadas parciales de las componentes no tendrían porque anularse si el sistema de coordenadas fuese curvilineo. La variación del tensor respecto una dirección coordenada viene reflejada por la derivada covariante de las componentes del mismo, respecto a esta dirección, no por las derivadas parciales correspondientes. Si el sistema de coordenadas es tal que $\partial g_{\mu\nu}/\partial x^{\alpha} = 0$ entonces las derivadas parciales de las componentes coinciden con las derivadas covariantes de las mismas.

Cabe destacar que la derivada covariante de los tensores fundamentales es nula, resultado que se conoce con el nombre de teorema de Ricci.

Dada una curva σ de clase c^1 y parámetro t:

$$\sigma : x^i = x^i(t) \qquad t_1 \leq t \leq t_2 \qquad i=1,\dots,n \tag{51}$$

se denomina derivada absoluta o intrínseca de $A^{\mu_1\dots\mu_p}_{\mu_{p+1},\dots,\mu_r}$ respecto el parámetro t al producto interno:

$$\frac{\delta A^{\mu_1\dots\mu_p}_{\mu_{p+1}\dots\mu_r}}{\delta t} = A^{\mu_1\dots\mu_p}_{\mu_{p+1}\dots\mu_r,\beta}\frac{dx^{\beta}}{dt} \tag{52}$$

que desarrollado se obtiene:

$$\frac{\delta A^{\mu_1\dots\mu_p}_{\mu_{p+1}\dots\mu_r}}{\delta t} = \frac{dA^{\mu_1\dots\mu_p}_{\mu_{p+1}\dots\mu_r}}{dt} + \Gamma^{\mu_1}_{\alpha\beta} A^{\alpha\mu_2\dots\mu_p}_{\mu_{p+1}\dots\mu_r}\frac{dx^{\beta}}{dt} + \dots + \tag{53}$$

$$+ \Gamma^{\mu_p}_{\alpha\beta} A^{\mu_1\dots\mu_{p-1}\alpha}_{\mu_{p+1}\dots\mu_r}\frac{dx^{\beta}}{dt} - \Gamma^{\alpha}_{\mu_{p+1}\beta} A^{\mu_1\dots\mu_p}_{\alpha\mu_{p+2}\dots\mu_r}\frac{dx^{\beta}}{dt} - \dots - \Gamma^{\alpha}_{\mu_r\beta} A^{\mu_1\dots\mu_p}_{\mu_{p+1}\dots\mu_{r-1}\alpha}\frac{dx^{\beta}}{dt}$$

Dado un campo tensorial, la derivada absoluta refleja la variación del mismo a lo largo de la curva. Si el campo es constante a lo largo de toda una curva σ, lo denominaremos campo de tensores paralelo (a la curva σ) y analíticamente verifica:

$$\frac{\delta A^{\mu_1 \ldots \mu_p}{}_{\mu_{p+1} \ldots \mu_r}}{\delta t} = 0 \tag{54}$$

Mientras que el orden con que derivamos parcialmente dos veces una función, de clase C^2, es indiferente, cara al resultado, no ocurre así en el caso de derivar covariantemente dos veces un tensor. Para que el orden de derivación covariante sea indiferente, debe anularse un cierto tensor de cuarto orden, que sólo depende de los $g_{\mu\nu}$, conocido como tensor de Riemann-Christoffel. En efecto, si A_i son las componentes de un tensor covariante arbitrario, si derivamos covariantemente dos veces, obtenemos:

$$A_{i,jk} = \frac{\partial A_{i,j}}{\partial x^k} - \Gamma^{\alpha}_{ik} A_{\alpha,j} - \Gamma^{\alpha}_{jk} A_{i,\alpha} \tag{55}$$

y puesto que:

$$A_{i,j} = \frac{\partial A_i}{\partial x^j} - \Gamma^{\alpha}_{ij} A_{\alpha} \tag{56}$$

resulta:

$$A_{i,jk} = \frac{\partial^2 A_i}{\partial x^k \partial x^j} - \frac{\partial \Gamma^{\alpha}_{ij}}{\partial x^k} A_{\alpha} - \Gamma^{\alpha}_{ij} \frac{\partial A_{\alpha}}{\partial x^k} - \Gamma^{\alpha}_{ik} \frac{\partial A_{\alpha}}{\partial x^j} + \Gamma^{\alpha}_{ik} \Gamma^{\beta}_{\alpha j} A_{\beta} - \Gamma^{\alpha}_{jk} \frac{\partial A_i}{\partial x^{\alpha}} + $$

$$+ \Gamma^{\alpha}_{jk} \Gamma^{\beta}_{i\alpha} A_{\beta} \tag{57}$$

y por tanto:

$$A_{i,jk} - A_{i,kj} = \left(\frac{\partial \Gamma^{\alpha}_{ik}}{\partial x^{j}} - \frac{\partial \Gamma^{\alpha}_{ij}}{\partial x^{k}} + \Gamma^{\beta}_{ik}\, \Gamma^{\alpha}_{\beta j} - \Gamma^{\beta}_{ij}\, \Gamma^{\alpha}_{\beta k} \right) A_{\alpha} \tag{58}$$

por la ley del cociente el término entre paréntesis de (58) es un tensor de cuarto orden, con un índice contravariante y tres covariantes, que denominaremos tensor de Riemann-Christoffel de segunda especie:

$$R^{\alpha}_{ijk} = \frac{\partial \Gamma^{\alpha}_{ik}}{\partial x^{j}} - \frac{\partial \Gamma^{\alpha}_{ij}}{\partial x^{k}} + \Gamma^{\beta}_{ik}\, \Gamma^{\alpha}_{\beta j} - \Gamma^{\beta}_{ij}\, \Gamma^{\alpha}_{\beta k} \tag{59}$$

Además por (58) observamos que el orden de derivación covariante es indiferente si y sólo si $R^{\alpha}_{ijk} = 0$.

Se define el tensor de Riemann-Christoffel de primera especie como:

$$R_{ijkl} = g_{i\alpha}\, R^{\alpha}_{jkl} \tag{60}$$

2.6. CARACTERIZACION DE LA EUCLIDIANIDAD. CURVATURA RIEMANNIANA

Pasemos a examinar seguidamente el problema de si la variedad de Riemann es euclidea o no. Se reduce a determinar si existe una transformación de coordenadas T: $x \rightarrow \bar{x}$ tal que reduzca al tensor métrico a un tensor constante. En estas condiciones $_{\bar{x}}\Gamma^{k}_{ij} = 0$ y a partir de las ecuaciones de transformación (41) cambiando las x^{i} por $\bar{x}^{i}$, obtenemos el siguiente sistema de ecuaciones diferenciales, en derivadas parciales, que debe ser satisfecho:

$$\frac{\partial^2 \bar{x}^m}{\partial x^i \partial x^j} - \Gamma^{\gamma}_{ij} \frac{\partial \bar{x}^m}{\partial x^{\gamma}} = 0 \qquad i,j,m = 1,\ldots,n \tag{61}$$

Es posible demostrar que este sistema de ecuaciones admite solución si y sólo si el tensor de Riemann-Christoffel es un tensor nulo, Sokolnikoff (1971). Dadas las componentes del tensor métrico, $g_{\mu\nu}$, para determinar si la variedad de Riemann es euclídea o no, se procede a calcular el tensor de Riemann-Christoffel y comprobar si es nulo o no. Si es nulo, el siguiente paso consiste en resolver (61), y trabajar con un sistema de coordenadas para el que el tensor métrico sea constante, en particular $g_{\mu\nu} = \delta_{\mu\nu}$, y poder calcular así, al menos a nivel local, mediante (28), la distancia entre dos puntos de la variedad. En caso contrario, para hallar la distancia entre dos puntos de la variedad hay que integrar el elemento de linea a lo largo de una curva geodésica.

Una medida de la euclidianidad de la variedad, lo constituye la curvatura Riemanniana, asociada a dos campos tensoriales contravariantes, A^j, B^k y viene dada, empleando el convenio de sumación de índices repetidos, por:

$$K = \frac{R_{ijkm} A^i A^k B^j B^m}{(g_{ik} g_{im} - g_{im} g_{ik}) A^i A^k B^j B^m} \tag{62}$$

es nula si el espacio es euclideo y constante si es isótropo (no depende de los A^j, B^k elegidos).

2.7. CALCULO DE GEODESICAS

Para hallar las geodésicas hay que resolver el siguiente problema variacional:

$$\delta \int_a^b ds = 0 \tag{63}$$

puesto que:

$$ds = \sqrt{g_{\mu\nu} dx^\mu dx^\nu} \tag{64}$$

podemos escribir, si llamamos:

$$\dot{x}^\mu = \frac{dx^\mu}{dt} \qquad \dot{x}^\nu = \frac{dx^\nu}{dt} \tag{65}$$

siendo t un parámetro cualquiera, la expresión:

$$\delta \int_{t_a}^{t_b} \sqrt{g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu} \; dt = 0 \tag{66}$$

equivalente a (63). A partir de las ecuaciones de Euler, obtenemos:

$$\frac{1}{2} (g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{-1/2} \frac{\partial g_{\mu\nu}}{\partial x^\alpha} \dot{x}^\mu\dot{x}^\nu - \frac{d}{dt} ((g_{\mu\nu}\dot{x}^\mu\dot{x}^\nu)^{-1/2} \; g_{\beta\alpha}\dot{x}^\beta) = 0 \tag{67}$$

$$\alpha=1,\ldots,n$$

Si tomamos como parámetro la longitud del arco s, entonces:

$$g_{\mu\nu} \; \dot{x}^\mu\dot{x}^\nu = 1 \tag{68}$$

por tanto:

$$\frac{1}{2}\frac{\partial g_{\mu\nu}}{\partial x^{\alpha}}\dot{x}^{\mu}\dot{x}^{\nu} - \frac{\partial g_{\beta\alpha}}{\partial x^{\gamma}}\dot{x}^{\gamma}\dot{x}^{\beta} - g_{\beta\alpha}\ddot{x}^{\beta} = 0 \qquad (69)$$

$$\alpha=1,\ldots,n$$

pero como:

$$\frac{\partial g_{\beta\alpha}}{\partial x^{\gamma}}\dot{x}^{\gamma}\dot{x}^{\beta} = \frac{1}{2}\frac{\partial g_{\beta\alpha}}{\partial x^{\gamma}}\dot{x}^{\gamma}\dot{x}^{\beta} + \frac{1}{2}\frac{\partial g_{\gamma\alpha}}{\partial x^{\beta}}\dot{x}^{\beta}\dot{x}^{\gamma} \qquad (70)$$

resulta:

$$g_{\beta\alpha}\ddot{x}^{\beta} + [\gamma\beta,\alpha]\ \dot{x}^{\gamma}\dot{x}^{\beta} = 0 \qquad (71)$$

$$\alpha = 1,\ldots,n$$

multiplicando por $g^{\gamma\beta}$ obtenemos:

$$\ddot{x}^{\gamma} + \Gamma^{\gamma}_{\alpha\beta}\dot{x}^{\alpha}\dot{x}^{\beta} = 0 \qquad \gamma=1,\ldots,n \qquad (72)$$

sistema que habrá que resolver para unas determinadas condiciones de contorno, para que la curva obtenida pase por dos puntos datos. Nótese que en general, dados dos puntos cualquiera, no tiene porque existir una geodésica que los una. Una de las propiedades de las geodésicas es que si consideramos el campo de vectores (tensores) tangentes unitarios a la geodésica en cada punto, éste es un campo de vectores paralelo (a la geodésica), en el sentido definido en (54).

En este capítulo hemos resumido brevemente el aparato matemático necesario para el desarrollo de distancias entre funciones de densidad paramétricas, que se van a ver en los siguientes capítulos.

3. UNA DISTANCIA RIEMANNIANA ENTRE FUNCIONES DE DENSIDAD. RESULTADOS GENERALES

Resumen:

En el presente capítulo se estudian algunas de las propiedades generales de una distancia definida, para funciones de densidad paramétricas, a través de la matriz de información de Fisher. Se aborda también el problema de la estimación de dicha distancia.

3.1. DEFINICION Y PROPIEDADES BASICAS.

3.2. UNA CONDICION SUFICIENTE DE EUCLIDIANIDAD.

3.3. CASO DE INDEPENDENCIA ESTOCASTICA.

3.4. RELACION CON LA DISTANCIA DE MAHALANOBIS.

3.5. ESTIMACION DE LA DISTANCIA Y DISTRIBUCION ASINTOTICA.

3.1. DEFINICION Y PROPIEDADES BASICAS.

Dadas k variables aleatorias $X_1,\ldots,X_k$, observables sobre distintas poblaciones, nos planteamos el problema de introducir una distancia que nos permita diferenciarlas. Vamos a admitir, como hipótesis, que dichas variables aleatorias admiten función de densidad conjunta, cuando nos restringimos a cada una de las poblaciones, y que toda la información utilizable para la construcción de una distancia está contenida en la función de densidad conjunta de las k variables aleatorias.

En general no será satisfactorio caracterizar a cada población por los valores medios de las variables en dicha población, puesto que éstos no determinan univocamente a la función de densidad conjunta. Una alternativa razonable es el admitir que la función de densidad conjunta de éstas k-variables aleatorias, en cualquiera de las poblaciones estudiadas, pertenece a una determinada clase de funciones de densidad paramétricas, $f(x^1,\ldots,x^k,\ \theta^1,\ldots,\theta^n)$. Así una población podrá venir caracterizada por una n-pla $(\theta^1,\ldots,\theta^n)$, coordenadas de un elemento de un espacio o variedad paramétrica E, definida por:

$$E = \{(\theta^1,\ldots,\theta^n)\in \mathbb{R}^n \mid f(x^1,\ldots,x^k,\ \theta^1,\ldots,\theta^n) \text{ es función de densidad}\} \tag{1}$$

de esta forma, el problema de distanciar poblaciones es equivalente al problema de introducir una distancia en la variedad paramétrica E, hecho que podemos lograr definiendo un campo tensorial covariante de segundo orden, simétrico y definido positivo, tomándolo como el tensor métrico de la variedad y utilizando la métrica Riemanniana definida (ver cap. 2). De esta forma, la distancia entre poblaciones resultará invariante frente a cualquier transformación admisible de los parámetros, puesto que coincidirá con la distancia entre puntos de la variedad E.

Los componentes de la matriz de información de Fisher, Rao (1948b), supuestas definidas, son las componentes de un tensor covariante de segundo orden, simétrico y generalmente definido positivo, que podemos tomar como el tensor métrico de la variedad E:

$$g_{\mu\nu} = E\left(\frac{1}{f^2} \frac{\partial f}{\partial \theta^\mu} \frac{\partial f}{\partial \theta^\nu} \right) \qquad \mu,\nu = 1,\dots,n \tag{2}$$

Las $g_{\mu\nu}$ así definidas son, en efecto, los componentes de un tensor covariante, de segundo orden ya que si efectuamos el cambio $(\theta^1,\dots,\theta^n) \rightarrow$ $\rightarrow (\bar{\theta}^1,\dots,\bar{\theta}^n)$, resultará, empleando el convenio de sumación de los índices repetidos:

$$\bar{g}_{\mu\nu} = E\left(\frac{1}{\bar{f}^2} \frac{\partial \bar{f}}{\partial \bar{\theta}^\mu} \frac{\partial \bar{f}}{\partial \bar{\theta}^\nu} \right) = E\left(\frac{1}{f^2} \frac{\partial f}{\partial \theta^\alpha} \frac{\partial \theta^\alpha}{\partial \bar{\theta}^\mu} \frac{\partial f}{\partial \theta^\beta} \frac{\partial \theta^\beta}{\partial \bar{\theta}^\nu} \right) = \frac{\partial \theta^\alpha}{\partial \bar{\theta}^\mu} \frac{\partial \theta^\beta}{\partial \bar{\theta}^\nu} E\left(\frac{1}{f^2} \frac{\partial f}{\partial \theta^\alpha} \frac{\partial f}{\partial \theta^\beta} \right) =$$

$$= \frac{\partial \theta^\alpha}{\partial \bar{\theta}^\mu} \frac{\partial \theta^\beta}{\partial \bar{\theta}^\nu} g_{\alpha\beta} . \qquad \mu,\nu = 1,\dots,n \tag{3}$$

ya que f es un invariante en E, $f=\bar{f}$.

Si ln f es de clase C^2 y el soporte de f no depende de los parámetros, podemos escribir:

$$g_{\mu\nu} = - E\left(\frac{\partial^2 \ln f}{\partial \theta^\mu \partial \theta^\nu} \right) \qquad \mu,\nu = 1,\dots,n \tag{4}$$

ya que para el caso absolutamente continuo tenemos:

$$\int_{\mathbb{R}^k} f = 1 \tag{5}$$

por tanto:

$$\frac{\partial}{\partial\theta^{\mu}}\int_{\mathbb{R}^k} f = \int_{\mathbb{R}^k}\frac{\partial f}{\partial\theta^{\mu}} = \int_S \frac{\partial \ln f}{\partial\theta^{\mu}} f = 0 \qquad \mu=1,\dots,n \tag{6}$$

(siendo S el soporte de f)

y derivando de nuevo:

$$\int_S \left(\frac{\partial^2 \ln f}{\partial\theta^{\mu}\partial\theta^{\nu}} f + \frac{\partial \ln f}{\partial\theta^{\mu}}\frac{\partial f}{\partial\theta^{\nu}} \right) = E\left(\frac{\partial^2 \ln f}{\partial\theta^{\nu}\partial\theta^{\mu}} \right) + g_{\mu\nu} = 0 \tag{7}$$

$$\mu,\nu = 1,\dots,n$$

y de (7) se sigue (4). Análogamente se prueba para el caso discreto, sustituyendo las integrales por sumatorios.

La expresión (4) la verifican muchas de las distribuciones usadas corrientemente, como la distribución normal, la exponencial, la Poisson,etc., mientras que no la verifican distribuciones como la uniforme.

La distancia así definida en el espacio paramétrico también tiene la propiedad de résultar invariante si efectuamos una transformación admsible de las variables aleatorias $X^1\dots X^k$, ya que no supone ningún cambio en la "información" que obtenemos de la población, puesto que es simplemente una forma distinta de tomar las medidas.

En efecto, para el caso absolutamente continuo, por ser la función de densidad, una densidad escalar, en $\mathbb{R}^k$, al efectuar la transformación:

$$(x^1\dots x^k) \longrightarrow (\bar{x}^1\dots \bar{x}^k) \tag{8}$$

la función de densidad se transforma en:

$$\bar{f} = f \cdot \left| \frac{\partial x^i}{\partial \bar{x}^j} \right| \tag{9}$$

luego resulta:

$$E\left(\frac{1}{\bar{f}^2}\frac{\partial \bar{f}}{\partial \theta^\mu}\frac{\partial \bar{f}}{\partial \theta^\nu}\right) = E\left(\frac{1}{f^2\left|\frac{\partial x^i}{\partial \bar{x}^j}\right|^2}\frac{\partial f}{\partial \theta^\mu}\left|\frac{\partial x^i}{\partial \bar{x}^j}\right|\frac{\partial f}{\partial \theta^\nu}\left|\frac{\partial x^i}{\partial \bar{x}^j}\right|\right) = E\left(\frac{1}{f^2}\frac{\partial f}{\partial \theta^\mu}\frac{\partial f}{\partial \theta^\nu}\right) \tag{10}$$

$$\mu, \nu = 1,\ldots,n$$

Para el caso discreto, la función de densidad es un invariante $\bar{f}(\bar{x}) = f(x)$, por tanto:

$$E\left(\frac{1}{\bar{f}^2}\frac{\partial \bar{f}}{\partial \theta^\mu}\frac{\partial \bar{f}}{\partial \theta^\nu}\right) = E\left(\frac{1}{f^2}\frac{\partial f}{\partial \theta^\mu}\frac{\partial f}{\partial \theta^\nu}\right) \tag{11}$$

A partir de (10) y (11) llegamos al resultado:

3.1.1. El tensor métrico de la variedad paramétrica E, y por tanto la distancia entre dos puntos de E, permanece inalterado al efectuar una trans-formación admisible de las variables aleatorias $x^1,\ldots,x^k$.

Es posible establecer también una relación entre el tensor métrico así definido y el concepto de información. Si definimos la información asociada a un resultado $(x^1,\ldots,x^k)$ como:

$$I = -\ln f(x^1,\ldots,x^k, \theta^1,\ldots,\theta^n) \tag{12}$$

entonces I es un invariante en E, una vez hemos fijado $(x^1,\ldots,x^k)$. En estas condiciones se cumple:

$$I,_{\mu\nu} = -\frac{\partial^2 \ln f}{\partial \theta^\mu \partial \theta^\nu} + \Gamma^\alpha_{\mu\nu}\frac{\partial \ln f}{\partial \theta^\alpha} \tag{13}$$

y al tomar la esperanza:

$$E(I,_{\mu\nu}) = -E\left(\frac{\partial^2 \ln f}{\partial \theta^\mu \partial \theta^\nu}\right) + \Gamma^\alpha_{\mu\nu}E\left(\frac{\partial \ln f}{\partial \theta^\alpha}\right) \tag{14}$$

pero $E(\frac{\partial \ln f}{\partial \theta^\alpha}) = 0$, por (6), por tanto:

$$g_{\mu\nu} = E(I,_{\mu\nu}) \qquad \mu,\nu = 1,\dots,n \tag{15}$$

con lo que obtenemos el resultado:

3.1.2. Independientemente del tensor métrico que definamos en E, se cumple que la esperanza de la segunda derivada covariante de I es igual al tensor definido por la matriz de información de Fisher.

Nótese que dicho resultado podría ser utilizado para definir en E al tensor fundamental. De alguna forma 3.1.2. refleja la íntima relación que hay con el cambio de información medio y la distancia, ya que al ser:

$$E(I,_{\mu\nu}) = E(I,_\mu I,_\nu) \tag{16}$$

podemos escribir:

$$ds^2 = E(I,_\mu I,_\nu)\, d\theta^\mu d\theta^\nu = E(I,_\mu I,_\nu\, d\theta^\mu d\theta^\nu) = E(dI^2) \tag{17}$$

por lo que obtenemos el resultado siguiente:

3.1.3 El cuadrado del elemento de línea es el promedio del cuadrado del cambio, diferencial, de información.

3.2. UNA CONDICION SUFICIENTE DE EUCLIDIANIDAD.

Vamos a intentar deducir a continuación alguna condición que deba cumplir la función de densidad conjunta, $f(x^1,\dots,x^k, \theta^1,\dots,\theta^n)$, suficiente para poder asegurar que la variedad paramétrica es euclídea.

Si la variedad paramétrica es euclídea, existirá algún sistema de coordenadas $(\theta^1,\ldots,\theta^n)$ para el que:

$$g_{\mu\nu} = \delta_{\mu\nu} \tag{18}$$

por tanto podemos escribir:

$$- E(\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu}) = \delta_{\mu\nu} \qquad \mu,\nu = 1,\ldots,n \tag{19}$$

Es suficiente para que se cumpla (19):

$$\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu} = - \delta_{\mu\nu} \tag{20}$$

Integrando respecto θ^1 resulta:

$$\frac{\partial \ln f}{\partial\theta^\nu} = -\delta_{1\nu}\ \theta^1 + C_1 \qquad \nu = 1,\ldots,n \tag{21}$$

donde C_1 es una función de $x^1,\ldots,x^k$ y $\theta^2,\ldots,\theta^n$. Integrando (21) para el caso $\nu = 1$, resulta:

$$\ln f = - \frac{(\theta^1)^2}{2} + C_1\theta^1 + D_1 \tag{22}$$

donde D_1 es una función de $\theta^2,\ldots,\theta^n$ y $x^1,\ldots,x^k$. Si derivamos parcialmente (22) respecto θ^1 y θ^ν con $\nu \neq 1$, resultará:

$$\frac{\partial^2 \ln f}{\partial\theta^\nu \partial\theta^1} = \frac{\partial}{\partial\theta^\nu}(-\theta^1 + C_1) = \frac{\partial C_1}{\partial\theta^\nu} = 0 \qquad \nu = 2,\ldots,n \tag{23}$$

donde la última igualdad se cumple por (20), por tanto C_1 no puede depender de ninguna θ^ν, es pues C_1 una función sólo de $x^1,\ldots,x^k$. Por otra parte si derivamos (22) respecto θ^2 dos veces, obtenemos:

$$\frac{\partial^2 \ln f}{\partial\theta^2 \partial\theta^2} = \frac{\partial^2 D_1}{\partial\theta^2 \partial\theta^2} = -1 \tag{24}$$

por lo que obtenemos:

$$D_1 = -\frac{(\theta^2)^2}{2} + C_2\theta^2 + D_2 \tag{25}$$

donde, si reiteramos el proceso, obtendremos que C_2 es sólo función de las $x^1,\ldots,x^k$, y D_2 es una función de $\theta^3,\ldots,\theta^n$ y $x^1,\ldots,x^k$, que verifica:

$$\frac{\partial^2 D_2}{\partial\theta^3 \partial\theta^3} = -1 \tag{26}$$

y por recurrencia obtenemos finalmente:

$$\ln f = \sum_{\mu=1}^{n} \left(-\frac{(\theta^\mu)^2}{2} + C_\mu\, \theta^\mu\right) + A \tag{27}$$

donde los C_μ y A son funciones de $x^1,\ldots,x^k$, y despejando f, resulta:

3.2.1. Si la función de densidad conjunta es de la forma:

$$f(x^1,\ldots,x^k,\theta^1,\ldots,\theta^n) = k(x^1,\ldots,x^k)\ e^{-\frac{1}{2}\sum_{\mu=1}^{n}(\theta^\mu)^2 + \sum_{\mu=1}^{n} C_\mu(x^1,\ldots,x^k)\,\theta^\mu} \tag{28}$$

(salvo transformaciones admisibles de los parámetros), podemos asegurar que el espacio paramétrico es euclídeo.

Nótese que la función de densidad definida en (28) es de la familia exponencial. Esta no es, sin embargo, una condición necesaria.

3.3. CASO DE INDEPENDENCIA ESTOCASTICA.

Veamos a continuación el caso de que las variables aleatorias $X^1,\dots,X^k$ sean estocásticamente independientes, con la suposición extra de que cada función de densidad marginal sea uniparamétrica:

$$f(x^1,\dots,x^k,\theta^1,\dots,\theta^k) = f_1(x^1,\theta^1)\dots f_k(x^k,\theta^k) \tag{29}$$

En estas condiciones, si se cumple (4), el tensor métrico verificará:

$$\begin{aligned} g_{\mu\nu} &= -E\left(\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu}\right) = -E\left(\frac{\partial^2}{\partial\theta^\mu\partial\theta^\nu}\sum_{i=1}^{k} \ln f_i\,(x^i,\theta^i)\right) = \\ &= \sum_{i=1}^{k} E\left(\frac{\partial^2 \ln f_i(x^i,\theta^i)}{\partial\theta^\mu\,\partial\theta^\nu}\right) = -E\left(\frac{\partial}{\partial\theta^\mu}\left[\frac{\partial \ln f_\nu(x^\nu,\theta^\nu)}{\partial\theta^\nu}\right]\right) = \\ &= -\,\delta_{\mu\nu} E\left(\frac{\partial^2 \ln f_\nu(x^\nu,\theta^\nu)}{\partial\theta^\nu\,\partial\theta^\nu}\right) \end{aligned} \tag{30}$$

En otras palabras, el tensor métrico es de forma diagonal, y verifica:

$$g_{\mu\nu} = \begin{cases} h_\mu(\theta^\mu) & \Leftrightarrow \quad \mu = \nu \\ 0 & \Leftrightarrow \quad \mu \neq \nu \end{cases} \tag{31}$$

es decir el elemento $g_{\mu\mu}$ es una cierta función de θ^μ, sólo depende de la μ-ésima coordenada.Por tanto, los símbolos de Christoffel se anularan salvo para el caso que los tres índices sean iguales, en este caso:

$$[\mu\mu,\mu] = \frac{1}{2}\frac{\partial h_\mu}{\partial\theta^\mu} \qquad \mu = 1,\dots,n \tag{32}$$

En cuanto a los de segunda especie:

$$\Gamma^\mu_{\alpha\beta} = g^{\mu\nu}[\alpha\beta,\nu] = g^{\mu\alpha}[\alpha\beta,\alpha] = \frac{\delta_{\mu\alpha}}{h_\mu}\,\frac{1}{2}\,\delta_{\alpha\beta}\,\frac{\partial h_\alpha}{\partial\theta^\alpha} \tag{33}$$

$$\mu,\alpha,\beta = 1,\dots,n$$

es decir, los símbolos de Christoffel de segunda especie no se anularan únicamente en el caso de que los tres índices sean iguales $\Gamma^{\alpha}_{\alpha\alpha}$, y entonces:

$$\Gamma^{\alpha}_{\alpha\alpha} = \frac{\partial \ln \sqrt{h_{\alpha}}}{\partial \theta^{\alpha}} \qquad \alpha = 1,\ldots,n \tag{34}$$

por tanto el tensor de Riemann-Christoffel es:

$$R^{a}_{ijk} = \frac{\partial \Gamma^{a}_{ik}}{\partial \theta^{j}} - \frac{\partial \Gamma^{a}_{ij}}{\partial \theta^{k}} + \Gamma^{\beta}_{ik}\,\Gamma^{a}_{\beta j} - \Gamma^{\beta}_{ij}\,\Gamma^{a}_{\beta k} =$$

$$= \frac{\partial \Gamma^{a}_{ik}}{\partial \theta^{j}} - \frac{\partial \Gamma^{a}_{ij}}{\partial \theta^{k}} + \Gamma^{k}_{ik}\,\Gamma^{a}_{kj} - \Gamma^{j}_{ij}\,\Gamma^{a}_{jk} = 0 \tag{35}$$

ya que $\frac{\partial \Gamma^{a}_{ik}}{\partial \theta^{j}}$ sólo será distinto de cero si $a=i=k=j$. Debido a que se anula el tensor de Riemann-Christoffel de segunda especie podemos enunciar:

<u>3.3.1</u>. Si las variables aleatorias $X^{1},\ldots,X^{k}$ son estocásticamente independientes, con marginales uniparamétricas, entonces la variedad paramétrica es euclídea.

En estas condiciones existirá una transformación de coordenadas tal que reducirá al tensor métrico a un tensor constante. Veamos a continuación como obtenerla.

Habremos de efectuar un cambio $(\theta^{1},\ldots,\theta^{n}) \rightarrow (\alpha^{1},\ldots,\alpha^{n})$ que verifique el sistema (ver cap. 2, (61)):

$$\frac{\partial^{2}\alpha^{m}}{\partial\theta^{i}\partial\theta^{j}} - \Gamma^{\gamma}_{ij}\,\frac{\partial\alpha^{m}}{\partial\theta^{\gamma}} = 0 \qquad i,j,m=1,\ldots,n \tag{36}$$

Podemos intentar encontrar soluciones particulares de forma que α^i sólo sea función de θ^i. En estas condiciones podemos reescribir (36) como:

$$\frac{d^2\alpha^i}{d(\theta^i)^2} - \Gamma^i_{ii}\frac{\partial\alpha^i}{\partial\theta^i} = 0 \qquad (37)$$

$$i=1,\ldots,n$$

y teniendo en cuenta que:

$$\Gamma^i_{ii} = g^{i\beta}[ii,\beta] = g^{ii}[ii,i] = \frac{1}{2g_{ii}}\frac{dg_{ii}}{d\theta^i} = \frac{d\ \ln\sqrt{g_{ii}}}{d\theta^i} \qquad (38)$$

podemos reescribir (37) como:

$$\frac{d^2\alpha^i}{d(\theta^i)^2} - \frac{d\ \ln\sqrt{g_{ii}}}{d\theta^i}\frac{d\alpha^i}{d\theta^i} = 0 \qquad (39)$$

$$i=1,\ldots,n$$

Sistema de ecuaciones diferenciales ordinario, que puede integrarse cada una de ellas, por separado, obteniendo:

$$\frac{1}{\frac{d\alpha^i}{d\theta^i}}\frac{d^2\alpha^i}{d(\theta^i)^2} = \frac{d\ \ln\sqrt{g_{ii}}}{d\theta^i} \qquad i=1,\ldots,n \qquad (40)$$

equivalente a:

$$\frac{d}{d\theta^i}\left(\ln\frac{d\alpha^i}{d\theta^i}\right) = \frac{d\ \ln\sqrt{g_{ii}}}{d\theta^i} \qquad i=1,\ldots,n \qquad (41)$$

integrando obtenemos:

$$\frac{d\alpha^i}{d\theta^i} = k_i\sqrt{g_{ii}} \qquad i=1,\ldots,n \qquad (42)$$

y finalmente podemos escribir el resultado:

3.3.2. Bajo las condiciones de 3.3.1., las transformaciones de coordenadas definidas por:

$$\alpha^i = k_i \int \sqrt{g_{ii}}\, d\theta^i + C_i \qquad i=1,\dots,n \tag{43}$$

donde k_i y C_i son constantes, reduce al tensor métrico a un tensor constante.

Bajo ciertas condiciones es posible establecer una curiosa relación entre la transformación (43) y las transformaciones normalizantes de una variable aleatoria. En efecto, si bajo las hipótesis supuestas en 3.2.2., existe una transformación admisible de parámetros:

$$\theta^i \xrightarrow{M_i} \mu^i \qquad i=1,\dots,n \tag{44}$$

tal que:

$$\mu^i = E(X^i) \qquad i=1,\dots,n \tag{45}$$

y si las funciones de densidad marginales f_i son de la forma:

$$f_i(x^i,\mu^i) = A_i(x^i)\, e^{B_i'(\mu^i)(x^i-\mu^i)+B_i(\mu^i)} \tag{46}$$

$$i=1,\dots,n$$

entonces si consideramos una transformación

$$\theta^i \xrightarrow{T_i} \alpha^i \qquad i=1,\dots,n \tag{47}$$

que reduzca al tensor métrico a un tensor constante, a partir de (43) podemos escribir:

$$\alpha^i = k_i \int \sqrt{g_{ii}}\, d\theta^i \qquad i=1,\dots,n \tag{48}$$

expresión que es un invariante frente a transformaciones admisibles de parámetros tal que su jacobiano sea de forma diagonal:

$$\alpha^i = k_i \int \sqrt{\bar{g}_{ii}}\, d\bar{\theta}^i = k \int \sqrt{\frac{\partial\theta^\alpha}{\partial\bar{\theta}^i}\frac{\partial\theta^\beta}{\partial\bar{\theta}^i}\, g_{\alpha\beta}}\ \frac{\partial\bar{\theta}^i}{\partial\theta^\gamma}\, d\theta^\gamma = k \int \sqrt{\left(\frac{\partial\theta^i}{\partial\bar{\theta}^i}\right)^2 g_{ii}}\ \frac{\partial\bar{\theta}^i}{\partial\theta^i}\, d\theta^i =$$

$$= k \int \sqrt{g_{ii}}\ d\theta^i \qquad (49)$$

luego resulta que:

$$(T_i \circ M_i^{-1})(\mu^i) = T_i(\theta^i) = k_i \int \sqrt{g_{ii}}\ d\mu^i \qquad (50)$$

estando en esta última expresión g_{ii} referido al sistema $(\mu^1,\ldots,\mu^n)$. Por otra parte X^i es un estadístico insesgado de μ^i, por (45), y

$$\mathrm{var}(X^i) = \sigma_i^2 = \frac{1}{E\left(\left(\frac{\partial \ln f_i}{\partial\mu^i}\right)^2\right)} = \frac{1}{g_{ii}} \qquad (51)$$

se cumple si y sólo si:

$$\frac{\partial \ln f_i}{\partial \mu^i} = c_i(x^i-\mu^i) \qquad (52)$$

para una cierta función c_i de μ^i. La expresión (52) es equivalente a:

$$\ln f_i = x^i \int c_i\, d\mu^i - \int c_i \mu^i d\mu^i = x^i \int c_i d\mu^i - \mu^i \int c_i d\mu^i + \iint c_i d\mu^i + k_i \qquad (53)$$

y si llamamos:

$$B_i(\mu^i) = \iint c_i\ d\mu^i \qquad (54)$$

obtenemos:

$$f_i = A_i(x^i)\ e^{B_i'\,(x^i-\mu^i)+B_i} \qquad i=1,\ldots,n \qquad (55)$$

como por hipótesis (46), se cumple (55), resulta:

$$\sigma_i = \frac{1}{\sqrt{g_{ii}}} \tag{56}$$

y por tanto podemos enunciar el resultado:

3.3.3. Bajo las hpótesis de 3.3.1, (44), (45) y (47), existe una cierta transformación admisible T_i tal que

$$(T_i \circ M_i^{-1})(\mu^i) = \int \frac{k_i}{\sigma_i}\, d\mu^i \tag{57}$$

es decir, se trata de una transformación funcionalmente equivalente a una transformación normalizante.

El resultado 3.3.1. nos relaciona la independencia estocástica con la euclidianidad en el caso de que las funciones de densidad marginales sean uniparamétricas. Vamos a examinar ahora el caso de que, aún suponiendo independencia, éstas dependan de varios parámetros. En estas condiciones

$$f(x^1,\ldots,x^k,\theta_1^1,\ldots,\theta_1^{n_1},\ldots,\theta_k^1,\ldots,\theta_k^{n_k}) = f_1(x^1,\theta_1^1\ldots\theta_1^{n_1}) \ldots$$

$$\ldots f_k(x^k,\theta_k^1\ldots\theta_k^{n_k}) \tag{58}$$

Podemos definir:

$$M_i = \{(\theta_i^1\ldots\theta_i^{n_i}) \in \mathbb{R}^{n_i} / f_i(x^i,\theta_i^1\ldots\theta_i^{n_i}) \text{ es función de densidad}\} \tag{59}$$

$i=1,\ldots,k$

y

$$M = \{(\theta_1^1 \ldots \theta_1^{n_1}, \ldots, \theta_k^1 \ldots \theta_k^{n_k}) \in \mathbb{R}^{n_1+\ldots+n_k} / \; f(x^1 \ldots x^k, \theta_1^1 \ldots \theta_k^{n_k}) \text{ es función de densidad}\} \tag{60}$$

con los tensores métricos definidos en E como:

$$g_{\mu\nu} = E\left(\frac{1}{f^2}\frac{\partial f}{\partial \theta_j^i}\frac{\partial f}{\partial \theta_r^m}\right) \tag{61}$$

$$\nu = (n_1+\ldots+n_{r-1}) + m$$

$$\mu = (n_1+\ldots+n_{j-1}) + i$$

y en M_i como:

$$_i g_{\mu\nu} = E\left(\frac{1}{f_i^2}\frac{\partial f_i}{\partial \theta_i^\mu}\frac{\partial f_i}{\partial \theta_i^\nu}\right) \qquad i=1\ldots k \tag{62}$$

entonces resulta:

$$g_{\mu\nu} = -E\left(\frac{\partial^2 \ln f}{\partial\theta_j^i \partial\theta_r^m}\right) = -E\left(\frac{\partial^2}{\partial\theta_j^i \partial\theta_r^m}\left(\sum_{h=1}^{k} \ln f_h(x^h, \theta_h^1 \ldots \theta_h^{n_h})\right)\right) =$$

$$= -E\left(\frac{\partial}{\partial\theta_j^i}\left(\frac{\partial \ln f_r}{\partial\theta_r^m}\right)\right) = \begin{cases} 0 & \text{si } j \neq r \\ -E\left(\dfrac{\partial^2 \ln f_j}{\partial\theta_j^i \partial\theta_j^m}\right) & \text{si } j=r \end{cases} \tag{63}$$

$\mu = (n_1+\ldots+n_{j-1})+i$

$\nu = (n_1+\ldots+n_{r-1})+m$

además:

$$_j g_{im} = -E\left(\frac{\partial^2 \ln f_j}{\partial\theta_j^i \partial\theta_j^m}\right) \qquad i,m = 1,\ldots,n_j \tag{64}$$

por lo que si llamamos $G_i = ({}_i g_{\mu\nu})$ $i=1,\dots,n$ y $G = (g_{\mu\nu})$ obtenemos el resultado siguiente:

3.3.4. Bajo la hipótesis (58), entonces:

$$G = \sum_{i=1}^{k}{}^{+} \; G_i \tag{65}$$

donde Σ^+ representa la suma directa de matrices.

En estas condiciones es posible expresar la distancia entre dos puntos de E, de coordenadas $(a_1^1 \dots a_1^{n_1}, \dots, a_k^1 \dots a_k^{n_k})$ y $(b_1^1 \dots b_1^{n_1}, \dots b_k^1 \dots b_k^{n_k})$, s, en función de las distancias obtenidas entre los puntos $(a_i^1 \dots a_i^{n_i})$ y $(b_i^1 \dots b_i^{n_i})$ calculadas en cada una de las variedades M_i, s_i.

En efecto, las ecuaciones de las geodésicas en E son:

$$\frac{d^2 x^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0 \tag{66}$$

$$\gamma = 1, \dots, n_1 + \dots + n_k$$

Obsérvese que $\Gamma^\gamma_{\alpha\beta} = 0$ si los índices γ, α, β no referencian a parámetros asociados a la misma función de densidad marginal, por tanto (66) puede escribirse como

$$\frac{d^2 \theta_i^\gamma}{ds^2} + {}_i\Gamma^\gamma_{\alpha\beta} \frac{d\theta_i^\alpha}{ds} \frac{d\theta_i^\beta}{ds} = 0 \tag{67}$$

$$\gamma = 1, \dots, n_i$$
$$i = 1, \dots, k$$

donde los ${}_i\Gamma^{\gamma}_{\alpha\beta}$ son los símbolos de Christoffel en M_i, y (67) el conjunto de las ecuaciones de las geodésicas en cada M_i.

Nótese que para obtener s_i hay que integrar, fijando la i, el sistema (67), con la condición extra:

$$ {}_ig_{\mu\nu} \frac{d\theta_i^{\mu}}{ds} \frac{d\theta_i^{\nu}}{ds} = 1 \tag{68} $$

condición de que el vector tangente a la geodésica en M_i sea unitario. Para obtener s, la distancia en E, debe imponerse:

$$ \sum_{i=1}^{k} {}_ig_{\mu\nu} \frac{d\theta_i^{\mu}}{ds} \frac{d\theta_i^{\nu}}{ds} = 1 \tag{69} $$

Si resolvemos, fijada i, (67) con:

$$ {}_ig_{\mu\nu} \frac{d\theta_i^{\mu}}{ds} \frac{d\theta_i^{\nu}}{ds} = \lambda_i \tag{70} $$

la condición (69) puede escribirse como:

$$ \sum_{i=1}^{k} \lambda_i = 1 \tag{71} $$

Debido a que la solución del sistema (67) con la condición (71), y las condiciones de contorno adecuadas para que para s=0 la geodésica pase por el punto $(a_1^1 \ldots a_1^{n_1} \ldots a_k^1 \ldots a_k^{n_k})$ y, para un cierto s, pase por el punto $(b_1^1 \ldots b_1^{n_1} \ldots b_k^1 \ldots b_k^{n_k})$, es también solución de los sistemas (69) con las condiciones (70), deberá verificarse:

$$s = \frac{s_i}{\sqrt{\lambda_i}} \qquad i=1...k \tag{72}$$

junto con (71) obtenemos

$$\lambda_i = \frac{s_i^2}{s_1^2+...+s_k^2} \qquad i=1...k \tag{73}$$

y finalmente el resultado:

3.3.5. Bajo la hipótesis (58), la distancia en E puede calcularse a partir de la distancia en cada M_i, por la relación:

$$s = \sqrt{s_1^2+...+s_k^2} \tag{74}$$

Bajo la hipótesis (58), los símbolos de Christoffel de segunda especie, en E, $\Gamma^{\gamma}_{\alpha\beta}$, se anulan si los índices no referencian a parámetros asociados a la misma función de densidad, por tanto el tensor de Riemann-Christoffel, R^{a}_{ijk}, no se anulara, a lo sumo, si los cuatro índices referencian a parámetros asociados a la misma función de densidad. De aquí, se sigue el resultado:

3.3.6. En las condiciones anteriormente expresadas, E sera euclídeo, si y sólo si cada una de las M_i, con el tensor métrico definido en (62), lo es.

Si bajo la hipótesis (58), consideramos fijos a todos los paráme-- tros asociados a cada marginal $f_h(x^h, \theta_h^1 \ldots \theta_h^{n_h})$, excepto para un cier- to $\theta_h^{k_h}$, es decir, $\theta_h^i = a_h^i \quad i=1,\ldots,\hat{k}_h,\ldots n_h$, entonces si considera-- mos la subvariedad incluida en E:

$$V = \{(x_1^1 \ldots x_1^{n_1} \ldots x_k^1 \ldots x_k^{n_k}) \in \mathbb{R}^{n_1+\ldots+n_k} / \; x_h^i = a_h^i \;\; i=1\ldots\hat{k}_h\ldots n_h \;\; h=1\ldots k\} \tag{75}$$

el tensor métrico inducido en V, dado el tensor métrico definido en E, viene dado por:

$$g_{\mu\nu} = -\sum_{i=1}^{k} E\left(\frac{\partial^2 \ln f_i}{\partial\theta_\mu^{k_\mu} \partial\theta_\nu^{k_\nu}} \right) = -E\left(\frac{\partial^2 \ln f_\mu}{\partial\theta_\mu^{k_\mu} \partial\theta_\nu^{k_\nu}} \right) \quad \mu,\nu=1\ldots n \tag{76}$$

por tanto el tensor métrico inducido en V es diagonal, y $g_{\mu\mu}$ sólo de-- pende de $\theta_\mu^{k_\mu}$, de aquí se sigue al resultado:

3.3.7. Bajo las hipótesis citadas, V es euclídeo.

Hemos obtenido pues, en este apartado otras condiciones suficien-- tes para asegurar que el espacio paramétrico es euclídeo y en particu-- lar establecida ciertas relaciones entre la independencia estocástica con la euclidianidad del espacio, viendo que no siempre aquella implica ésta.

3.4. RELACION CON LA DISTANCIA DE MAHALANOBIS.

Consideremos a continuación la matriz de varianzas y covarianzas supuesta existente, $\Sigma = (\sigma_{ij})$, con:

$$\sigma_{ij} = E((X^i - E(X^i))(X^j - E(X^j))) \qquad i,j=1\ldots k \tag{77}$$

asociada a $f(x^1,\ldots,x^k,\theta^1\ldots\theta^n)$. Cada una de las componentes de dicha matriz, σ_{ij}, es una función de los parámetros, $\theta^1\ldots\theta^n$, invariante frente a transformaciones admisibles de los mismos en E. Lo mismo ocurrirá por tanto con la matriz inversa de Σ, supuesta existente también, $\Sigma^{-1} = (\sigma^{ij})$. Nótese que en ambos casos los índices carecen de caracter tensorial. Por tanto estas matrices no nos sirven para definir en E una métrica Riemanniana, independiente del sistema de coordenadas, por no ser las σ_{ij} ó σ^{ij} las componentes de un tensor covariante. Sin embargo, fijado un sistema de coordenadas particular, es posible definir un tensor covariante de segundo orden, simétrico, $h_{\mu\nu}$, que verifique:

$$h_{\mu\nu} = \sigma^{\mu\nu} \qquad \mu,\nu = 1\ldots k \tag{78}$$

es decir, bajo este sistema de coordenadas particular, las componentes de dicho tensor coincidan con las componentes de la matriz inversa a la matriz de varianzas y covarianzas.

Supongamos que $n \geq k$, es decir hay al menos tantos parámetros como variables, y admitamos además que existe un sistema de coordenadas tal que:

$$E(X^i) = \theta^i \qquad i=1\ldots k \tag{79}$$

Entonces si definimos una variedad V a partir de:

$$V = \{(\theta^1\ldots\theta^n) \in E \ / \ \theta^{k+1} = \alpha_1,\ldots,\theta^n = \alpha_{n-k}\} \tag{80}$$

donde las α_i son constantes, podemos considerar, en ella al tensor métrico inducido por el tensor métrico en E, y con él formar la matriz $(k \times k)$ $G = (g_{\mu\nu})$ que supondremos definida positiva. Definamos ahora en este sistema de coordenadas, un tensor covariante a partir de (78) y formemos $H = (h_{\mu\nu})$. Entonces para cada punto de V, se cumple que $G-H$ es semidefinida positiva. En efecto, $(X^1,\ldots,X^k)$ es un estimador insesgado de $\theta^1,\ldots,\theta^k$, (79), entonces bajo las hipótesis necesarias para que se cumpla la desigualdad de Cramer-Rao, $H^{-1}-G^{-1}$ es semidefinida positiva, por tanto para un vector k dimensional Z cualquiera, se cumple:

$$Q(Z) = Z^t(H^{-1}-G^{-1})Z \geq 0 \tag{81}$$

por ser H^{-1} simétrica y G^{-1} simétrica y definida positiva, existe una transformación lineal no singular tal que:

$$\begin{aligned} P^t\, G^{-1} P &= I \\ P^t\, H^{-1} P &= D \end{aligned} \tag{82}$$

con D una matriz diagonal, de valores propios, solución de

$$\left|H^{-1} - \lambda G^{-1}\right| = 0 \tag{83}$$

Si efectuamos la transformación:

$$Y = P^t X \tag{84}$$

resulta:

$$Q(X) = Q(Y) = Y^t P^t (H^{-1}-G^{-1}) P Y = Y^t (D-I) Y = \sum_{i=1}^{k} (\lambda_i - 1) y_i^2 \geq 0 \tag{85}$$

puesto que esto se verifica para todo Y, ello implica que $\lambda_i \geq 1$ $i=1\ldots k$.

Consideremos ahora:

$$|H - \mu G| = 0 \tag{86}$$

las soluciones son:

$$\mu_i = \frac{1}{\lambda_i} \qquad i = 1,\ldots n \tag{87}$$

luego si consideramos la forma cuadrática:

$$R(Z) = Z^t (H-G)Z \tag{88}$$

y efectuamos una transformación ortogonal tal que:

$$\begin{aligned} T^t G\ T &= I \\ T^t H\ T &= D^{-1} \end{aligned} \tag{89}$$

obtendremos, si

$$Y = T^t Z \tag{90}$$

la siguiente expresión:

$$R(Z) = R(Y) = Y^t T^t (H-G) TY = Y^t (D^{-1} - I) Y = \sum_{i=1}^{k} (\frac{1}{\lambda_i} - 1)\ y_i^2 \tag{91}$$

pero al ser $\lambda_i \geq 1$ resulta:

$$R(Z) \leq 0 \tag{92}$$

por lo que R(Z) es semidefinida negativa, o equivalentemente G-H es semidefinida positiva.

Bajo estas condiciones, la longitud, medida a lo largo de una curva C incluida en V, usando el elemento de linea obtenido a través de $h_{\mu\nu}$, es inferior a la misma longitud obtenida a partir de la restricción de $g_{\mu\nu}$ a V. En efecto:

$$\int_C \sqrt{h_{\mu\nu} \frac{d\theta^{\mu}}{dt} \frac{d\theta^{\nu}}{dt}} \, dt \leq \int_C \sqrt{g_{\mu\nu} \frac{d\theta^{\mu}}{dt} \frac{d\theta^{\nu}}{dt}} \, dt \qquad (93)$$

debido a que:

$$(g_{\mu\nu} - h_{\mu\nu}) \frac{d\theta^{\mu}}{dt} \frac{d\theta^{\nu}}{dt} \geq 0 \qquad (94)$$

como la distancia entre dos puntos A y B se define como el ínfimo del conjunto de las longitudes de todas las curvas que unen A y B, éste, en el caso de usar $h_{\mu\nu}$ como tensor métrico, no puede ser superior al que se obtiene usando $g_{\mu\nu}$, debido a (93), por lo que obtenemos el resultado:

3.4.1. Bajo ciertas hipótesis generales, la distancia entre dos puntos de V obtenida a partir del tensor métrico inducido en V, por el tensor métrico de E, es superior o igual a la distancia obtenida definiendo en V, como tensor métrico, $h_{\mu\nu}$.

La igualdad entre el tensor $h_{\mu\nu}$ y $g_{\mu\nu}$ (restringido a V) se obtiene si la función de densidad es una función lineal de:

$$\frac{1}{f} \frac{\partial f}{\partial \theta^1}, \; \ldots, \; \frac{1}{f} \frac{\partial f}{\partial \theta^k} \qquad (95)$$

Mardia, (1979). Ello impone una cierta forma general a f, al resolver la ecuación diferencial involucrada.

Con ello hemos relacionado la distancia obtenida a través de la matriz de información de Fisher con distancias análogas a la distancia de Mahalanobis, ya que en un sistema de coordenadas particular, el tensor métrico $h_{\mu\nu}$ coincidirá con la matriz inversa de varianzas y covarianzas.

3.5. ESTIMACION DE LA DISTANCIA Y DISTRIBUCION ASINTOTICA

Hasta aquí hemos supuesto conocidos a los parámetros poblacionales, en la práctica desconocidos, pero pueden ser estimados por máxima verosimilitud. Si llamamos β^2 a la distancia al cuadrado entre el punto $(\theta^1, \ldots, \theta^n) \in E$ real, y el estimado, $(\hat{\theta}^1, \ldots, \hat{\theta}^n) \in E$ a partir de una muestra aleatoria simple de tamaño N, para muestras grandes, es conocido que la forma cuadrática:

$$\chi^2 = -N\, E\left(\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu} \right) (\hat{\theta}^\mu - \theta^\mu)(\hat{\theta}^\nu - \theta^\nu) \tag{96}$$

se distribuye aproximadamente según una distribución ji-cuadrado con n grados de libertad, Mood & Graybill (1963). Por tanto para tamaños muestrales elevados y para un cierto a convenientemente elegido, podremos escribir:

$$\text{Prob}\left[d \geq \frac{a}{\sqrt{N}} \right] \leq \varepsilon \tag{97}$$

donde ε es una constante positiva arbitrariamente pequeña y d viene definida por:

$$d^2 = -E\left(\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu} \right) (\hat{\theta}^\mu - \theta^\mu)(\hat{\theta}^\nu - \theta^\nu) \tag{98}$$

y además como se verifica:

$$\lim_{d\to 0} \frac{|d-\beta|}{d} = 0 \tag{99}$$

resulta que podemos, para tamaños muestrales elevados, escribir:

$$\beta^2 \simeq E\left(\frac{\partial^2 \ln f}{\partial\theta^\mu \partial\theta^\nu}\right)(\hat{\theta}^\mu-\theta^\mu)(\hat{\theta}^\nu-\theta^\nu) \tag{100}$$

y por tanto:

$$\chi^2 \simeq N\,\beta^2 \tag{101}$$

debido a que $N\beta^2$ difiere en tan poco como queramos (con una probabilidad tan grande como se quiera) de Nd^2, eligiendo un tamaño muestral suficientemente grande. Por tanto, podemos escribir:

3.5.1. La distribución asintótica de $N\beta^2$ es una distribución ji-cuadrado con n grados de libertad. Los intervalos confidenciales seran esferas geodésicas, en el espacio paramétrico.

Consideremos a continuación dos poblaciones estadísticas, representables como puntos del espacio paramétrico E, de coordenadas reales $(\theta_A^1,\ldots,\theta_A^n)$ y $(\theta_B^1,\ldots\theta_B^n)$ respectivamente, y sean $\hat{\theta}_A^1,\ldots,\hat{\theta}_A^n$ y $\hat{\theta}_B^1,\ldots,\hat{\theta}_B^n$ los estimadores máximo-verosímiles de los parámetros reales, contruidos a partir de dos muestras aleatorias simples de tamaños N_A y N_B respectivamente. Si los tamaños muestrales son elevados, es bien sabido que $(\hat{\theta}_A^1,\ldots\hat{\theta}_A^n)$ y $(\hat{\theta}_B^1\ldots\hat{\theta}_B^n)$ siguen ambas aproximadamente una distribución normal multivariante, y teniendo en cuenta (100), podemos escribir:

$$
\begin{aligned}
f(\hat{\theta}^1_A \ldots \hat{\theta}^n_A, \theta^1_A \ldots \theta^n_A) &= k_A\, e^{-\frac{1}{2} N_A \beta^2_A} \\
f(\hat{\theta}^1_B \ldots \hat{\theta}^n_B, \theta^1_B \ldots \theta^n_B) &= k_B\, e^{-\frac{1}{2} N_B \beta^2_B}
\end{aligned}
\tag{102}
$$

si llamamos β^2_A y β^2_B a las distancias al cuadrado entre el punto poblacional real y el estimado en las poblaciones A y B respectivamente. Por ser las $\hat{\theta}^i_A$ y $\hat{\theta}^j_B$ independientes, la función de densidad conjunta vendrá dada por:

$$
f(\hat{\theta}^1_A \ldots \hat{\theta}^n_B, \theta^1_A, \ldots \theta^n_B) = C\, e^{-\frac{1}{2}(N_A \beta^2_A + N_B \beta^2_B)} \tag{103}
$$

Supongamos existente, en E, la curva de longitud mínima que una $(\theta^1_A \ldots \theta^n_A)$ con $(\theta^1_B \ldots \theta^n_B)$. Siempre existirá un sistema de coordenadas $(x^1 \ldots x^n)$ tal que las derivadas parciales del tensor métrico se anularan en toda la curva, Sokolnikov (1971), y por tanto el tensor métrico permanecerá constante a lo largo de la misma, y en particular puede ser la identidad. Por tanto, bajo el nuevo sistema de coordenadas, cuando consideremos un par de puntos próximos a dicha geodésica, podemos calcular su distancia, de forma aproximada, como si se tratara de un espacio euclídeo con una referencia cartesiana. Una vez efectuado el cambio, si llamamos $x^1_A \ldots x^n_A$ y $x^1_B \ldots$ $\ldots x^n_B$ a las coordenadas poblacionales estimadas y $a^1_A \ldots a^n_A$ y $a^1_B \ldots a^n_B$ a las coordenadas poblaciones reales, podemos escribir:

$$
f(x^1_A, \ldots, x^n_A, x^1_B, \ldots, x^n_B) = (2\pi)^{-n} (N_A N_B)^{n/2} e^{-\frac{1}{2}\left(N_A \sum_{i=1}^{n} (x^i_A - a^i_A)^2 + N_B \sum_{i=1}^{n} (x^i_B - a^i_B)^2\right)} \tag{104}
$$

debido a que, para tamaño muestrales elevados, los puntos estimados son próximos a los puntos reales, en el sentido dado por (97). A partir de esta última expresión es posible deducir la función de densidad conjunta de $z^i = x_A^i - x_B^i$ $i=1,\ldots,n$, igual a:

$$f(z^1,\ldots,z^n) = (2\pi)^{-n/2} \left(\frac{N_A N_B}{N_A+N_B}\right)^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}\sum_{i=1}^{n}(z^i - a_A^i + a_B^i)^2} \tag{105}$$

equivalente a:

$$f(z^1,\ldots,z^n) = (2\pi)^{-n/2} \left(\frac{N_A N_B}{N_A+N_B}\right)^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}\left(\sum_{i=1}^{n}(z^i)^2 - 2\sum_{i=1}^{n} z^i (a_A^i - a_B^i) + \sum_{i=1}^{n}(a_A^i - a_B^i)^2\right)} \tag{106}$$

Si llamamos D a la distancia estimada y Δ a la distancia real, podemos escribir:

$$f(z^1,\ldots,z^n) = (2\pi)^{-n/2} \left(\frac{N_A N_B}{N_A+N_B}\right)^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}\left(D^2 - 2\sum_{i=1}^{n} z^i (a_A^i - a_B^i) + \Delta^2\right)} \tag{107}$$

llamando ahora θ al menor angulo entre los vectores $(z^1,\ldots,z^n)$ y $(a_A^1 - a_B^1, \ldots, a_A^n - a_B^n)$ resulta:

$$f(D,\theta) = (2\pi)^{-n/2} \left(\frac{N_A N_B}{N_A+N_B}\right)^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}(D^2 - 2D\Delta\cos\theta + \Delta^2)} \cdot \frac{2\pi^{\frac{n-1}{2}}}{\Gamma\left(\frac{n-1}{2}\right)} (D \operatorname{sen}\theta)^{n-2} D \qquad D \geq 0 \qquad \theta \in [0,\pi] \tag{108}$$

y para hallar la función de densidad de D, bastará integrar θ entre 0 y π, debido a las propiedades de las integrales esféricas, Schwartz (1969). Teniendo en cuenta que (108) puede escribirse como:

$$f(D,\theta) = \frac{N_A N_B}{N_A+N_B} D^{n/2} e^{-\frac{1}{2}\frac{N_A N_B}{N_A+N_B}(D^2+\Delta^2)} \frac{1}{\Delta^{\frac{n-2}{2}}} \left(\frac{D\Delta N_A N_B}{2(N_A+N_B)}\right)^{(n-2)/2} \cdot$$

$$\cdot \frac{1}{\sqrt{\pi}\,\Gamma\left(\frac{n-1}{2}\right)} \cdot \operatorname{sen}^{n-2}\theta \; e^{(N_A N_B / N_A+N_B)\, D\,\Delta\cos\theta} \qquad D \geq 0 \quad \theta \in [0,\pi] \tag{109}$$