

TEORIA DE CONFIGURACIONES

MANUEL ANTONIO TORRES GOMAR

TESIS

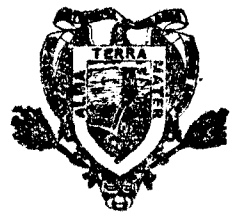
PRESENTADA COMO REQUISITO PARCIAL

PARA OBTENER EL GRADO DE:

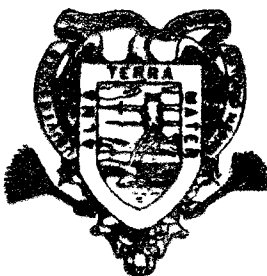
MAESTRO EN CIENCIAS

EN ESTADISTICA EXPERIMENTAL

Universidad Autónoma Agraria
"ANTONIO NARRO"



BIBLIOTECA



Universidad Autónoma Agraria
"Antonio Narro"

PROGRAMA DE GRADUADOS

Buenvista, Saltillo, Coah.

SEPTIEMBRE DEL 2000

Universidad Autónoma Agraria Antonio Narro

Subdirección de Postgrado

TEORÍA DE CONFIGURACIONES

TESIS

Por:

MANUEL ANTONIO TORRES GOMAR

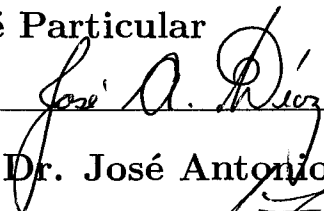
Elaborada bajo la supervisión del comité particular de asesoría y aprobada como
requisito parcial, para optar al grado de:

MAESTRO EN CIENCIAS

EN ESTADÍSTICA EXPERIMENTAL

Comité Particular

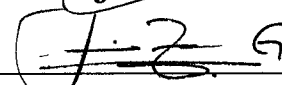
Asesor principal:


Dr. José Antonio Díaz García

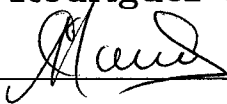
Asesor:


Dr. Alejandro Zarate Lupercio

Asesor:


M.C. Luis Rodríguez Gutiérrez

Asesor:


M.C. Roberto Coronado Niño


Dr. Ramiro López Trujillo
Subdirector de Postgrado

Buenavista, Saltillo, Coahuila. Septiembre de 2000.

AGRADECIMIENTOS

Al Dr. José Antonio Díaz García, asesor y amigo, por su tiempo, paciencia y guía a lo largo de este trabajo.

A mis sinodales Dr. Alejandro Zárate Lupercio, M.C. Luis Rodríguez Gutiérrez y M.C. Roberto Coronado Niño, por su apoyo en la realización de esta tesis.

Se agradece al Profesor Juan Aguirre Alvarez por su apoyo en la redacción de este trabajo y a mi compañero de maestría Miguel Angel Mendoza Zamora quién me apoyo a lo largo de estos estudios.

A los profesores de la maestría en Estadística de la UAAAN ya que con su entusiasmo y profesionalismo lograron compartir sus conocimientos, en especial a: M.C. Emilio Padrón Corral quién mi apoyo desde antes de entrar en esta. Dr. Víctor Leiva quien supo influir en mis estudios. Al Dr. Rolando Cavazos Cadena de quien aprendí no sólo conocimientos, sino también a elaborar mejores clases. Al M.C. Félix de Jesús Sánchez Pérez por su amistad a lo largo de esta estancia.

Se agradece a la Escuela de Matemáticas de la U.A. de C. por el apoyo brindado durante mis estudios, en particular a la dirección. Al MC José Reyes Valdés investigador del Centro de Investigación en Matemáticas de la U.A. de C., quien me apoyó con respuestas a dudas sobre los temas que estudiaba.

Se agradece el apoyo brindado por mi familia quienes siempre me impulsaron a terminar la tesis y nunca dejaron de recordarme que así lo hiciera.

DEDICATORIA

Para Rosy y Adriana con todo amor.

Por el tiempo dedicado y compartido

al realizar este trabajo.

COMPENDIO

Teoría de Configuraciones

POR

MANUEL ANTONIO TORRES GOMAR

MAESTRÍA

ESTADÍSTICA EXPERIMENTAL

UNIVERSIDAD AUTÓNOMA AGRARIA "ANTONIO NARRO"

BUENAVISTA, SALTILLO, COAHUILA. SEPTIEMBRE DE 2000

Dr. José Antonio Díaz García -Asesor-

Palabras clave: Configuración, Escalamiento, Rotación, Traslación, Polinomio Zonal, Variedad de Stiefel.

En el presente trabajo se plantea inicialmente la importancia de la teoría de configuraciones en diversas ramas de la ciencia y la tecnología; los requisitos matemáticos y las bases estadísticas para trabajar en esta disciplina. Como tema central se propone una técnica alternativa para el análisis de configuraciones basada en la factorización polar.

ABSTRACT

Shape Theory

BY

MANUEL ANTONIO TORRES GOMAR

MASTER OF SCIENCE

EXPERIMENTAL STATISTICS

UNIVERSIDAD AUTÓNOMA AGRARIA "ANTONIO NARRO"

BUENAVISTA, SALTILLO, COAHUILA. SEPTEMBER 2000

Dr. José Antonio Díaz García -Advisor-

Key Words: Rotation, Shape, Size, Stiefel Manifold, Translate, Zonal Polynomial.

This essay illustrates initially the relevance of shape theory in different branches of science and technology and the mathematical and statistical foundation necessary to work on the subject. As a primary topic, an optional technique is suggested for the analysis of shapes based upon polar factorization.

ÍNDICE DE CONTENIDO

	Página
ÍNDICE DE FIGURAS	ix
INTRODUCCIÓN	1
PREREQUISITOS	10
Álgebra Lineal	10
Transformaciones asociadas e interpretación geométrica	15
Producto de Kronecker y vectorización de una matriz	24
Factorización QR	27
ÁREAS DE APLICACIÓN	32
Configuración de una ciudad	32
Geometría Computacional	34
Manufactura	35
Medicina	35
Morfometría	36
Sistemas de información	37
Reconocimiento de caras por medio de fotografías	38
Reconocimiento de firmas	39
Topografía	40
DISTRIBUCIONES ASOCIADAS A LA TEORÍA DE CONFIGURACIONES	42

Distribución de tamaño-y-configuración	43
Distribución de la configuración	47
Distribución excluyendo reflexión	50
Distribución de tamaño-y-configuración mediante la factorización polar	51
LITERATURA CITADA	53
APÉNDICES	56
A. Producto Exterior	57
B. Medidas Invariantes	84
C. Polinomios Zonales	92
D. Estadística	104

ÍNDICE DE FIGURAS

Figura No.		Página
1.1	Configuraciones entre triángulos	3
2.1	Puntos en el plano	19
2.2	Puntos trasladados	20
2.3	Rotación de vectores	22
2.4	Puntos rotados-y-trasladados	23
2.5	Reflexión de vectores	24
3.1	Contornos de la ciudad separados por año	33
3.2	Contornos de la ciudad acoplados	34

INTRODUCCIÓN

Una cualidad del ser humano es: la comparación; ocurre desde el momento en que puede identificar a sus padres, pues el hecho de crecer con ellos le permite aprender a distinguir las diferencias entre estos y otras personas.

Dicha cualidad se va refinando con el paso del tiempo llegando a ser de gran importancia para el desarrollo del individuo, ya que la necesidad de comparar objetos se presenta a cada momento, por ejemplo: ir a una frutería requiere comparar frutas con el único objetivo de seleccionar las "mejores".

Cabe señalar que el objeto de estudio puede ser un objeto simple, o un sistema compuesto de varios de ellos. Como en el caso de la configuración del corazón o en la forma del contorno de una ciudad.

Por otra parte, la figura de un objeto no varía si éste se mueve y cambia su posición original, se pueden citar como movimientos de este tipo, una posible traslación y/o rotación del mismo. A la vez es común decir que dos objetos tienen la misma figura, cuando la forma asociada sea la misma y varíe tan solo por el tamaño. En este caso se dice que difieren por un posible escalamiento. Cuando se presente alguna de las transformaciones mencionadas anteriormente, se dirá que estas diferencias no son significativas y por tal motivo las figuras serán equivalentes.

Este tipo de comparaciones son realizadas de manera innata, y varía de

individuo a individuo. Como en el caso del sabor de una fruta, donde no a todos les sabe igual, aún y cuando se trate del mismo producto.

En la ciencia y la tecnología no basta el realizar una comparación intuitiva entre dos objetos, más bien se requiere dar un procedimiento que permita establecer si en realidad dos objetos son equivalentes en figura o en el sentido antes mencionado, con el fin de probar cierta hipótesis relacionada con ellos.

Debido a la necesidad de poder cuantificar la diferencia entre dos figuras, de manera tal, que esta cuantificación sea precisa, es indispensable usar las herramientas matemáticas, dado que esta área de la ciencia permite medir diversos tipos de cuantificaciones de manera eficaz. Estas mediciones se pueden realizar en cuanto al peso, color, sabor, aroma, forma, etc.

Algunos métodos matemáticos para realizar la comparación están bien establecidos, como en el caso de comparar pesos, sin embargo el método matemático requerido para realizar otro tipo de comparaciones, no existe o no es fácil de aplicar, como en el caso del sabor de una fruta.

Dentro de estas herramientas matemáticas existe una de ellas llamada teoría de configuraciones (traducido del inglés "Shape Theory"), surge con el objetivo de cuantificar la figura de un objeto y poderla comparar. Kendall (1984), da nombre a esta disciplina.

Por la configuración de un objeto se entiende a la información que se almacena en un arreglo matricial de cierto número de puntos, una vez que estos han sido

trasladados, rotados y/o escalados. Este arreglo matricial contiene la localización de cada punto en el espacio donde éste está definido.

Con esta metodología el objeto es identificado por su configuración, y no por su figura en sí, de manera tal que se ha realizado una abstracción de la figura. Por otra parte la teoría de configuraciones permite determinar si dos objetos son equivalentes, siempre y cuando estos difieran tan sólo por una posible traslación, rotación y/o escalamiento.

En un arreglo matricial las diferencias debidas se pueden eliminar al aplicar ciertas transformaciones lineales que permitan invarianza, de manera tal, que el arreglo resultante sea el mismo.

En el caso de un triángulo, la configuración se puede guardar en un arreglo de coordenadas que contenga sus tres vértices; se puede decir que dos triángulos tienen la misma configuración si sólo difieren por una de las posibles transformaciones mencionadas anteriormente, tal como se ilustra en la Figura (1.1).

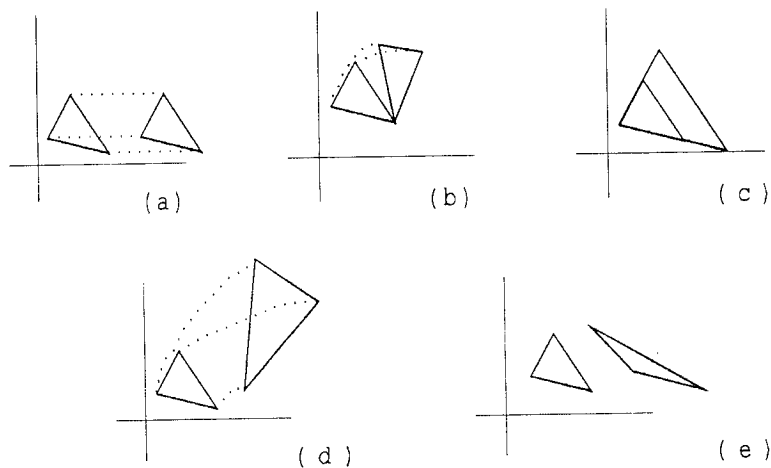


Figura 1.1 Configuraciones entre triángulos

Se puede observar que los dos triángulos que se describen. En la Figura (1.1.a) poseen la misma configuración, basta aplicar una traslación. En las figuras (1.1.b) y (1.1.c), los triángulos poseen la misma configuración dado que al aplicar una rotación o escalamiento respectivamente se llega al mismo triángulo. En la Figura (1.1.d) los triángulos también poseen la misma configuración, aunque en este caso hay que realizar las tres transformaciones a la vez. Finalmente, en la Figura (1.1.e) los dos triángulos no poseen la misma configuración, porque no es posible, que por medio de alguna de las transformaciones anteriores, obtener la misma configuración.

En el ejemplo de los triángulos no fue difícil precisar en que momento los objetos presentaban semejanza en cuanto a su configuración, pero a medida que los sistemas son más complejos, al tratar de determinar dicha igualdad se hace necesario hablar de técnicas más precisas que permitan realizar esta comparación.

La estadística, dentro de la teoría de configuraciones juega un papel importante, dado que permite clasificar la configuración de un objeto, bajo el supuesto de normalidad de los puntos y así se puedan realizar comparaciones entre diferentes configuraciones, por medio de pruebas de hipótesis.

Los pioneros de este enfoque fueron Kendall (1984) y Bookstein (1986).

Aunque anteriormente, Gower (1975) adaptó el procedimiento de procrustes a la teoría de configuraciones.

Small (1981), encontró la densidad asociada para una configuración en el plano, bajo el supuesto de media cero, la densidad es del tipo uniforme con respecto a la métrica de Reimann.

Por su parte. Kendall (1984), clasificó el espacio de la teoría de configuraciones, además de las propiedades geométricas asociadas.

Posteriormente Bookstein (1986), encontró la distribución asociada a un conjunto de datos en dos dimensiones con respecto al tamaño de la configuración.

Las distribuciones asociadas a la configuración de los puntos, bajo el supuesto de que estos provienen de una distribución normal; Tales distribuciones fueron determinadas por Dryden y Mardia (1991), se refieren a las posibles operaciones básicas de invarianza, es decir hallaron las distribuciones de traslación, tamaño y/o rotación.

Goodall (1991), estimó los parámetros de aproximación, aplicando los métodos procrustes, por medio de máxima verosimilitud.

La factorización QR fue introducida en la teoría de configuraciones, por Goodall y Mardia (1991, 1992).

Kent (1994), encontró la distribución compleja de Bingham, utilizando la teoría de configuraciones y dando con ello una aportación relevante a la estadística multidimensional no centrada, y también trabajó en la aproximación de una configuración de un objeto por medio de una imagen, utilizando el método procrusters.

Siguiendo la notación en Goodall y Mardia (1991), suponga ahora que $X \in \mathbb{R}^{N \times K}$ es una matriz que representa la figura geométrica, compuesta por N puntos (también llamados en inglés "landmarks") en \mathbb{R}^K , y asumiendo que la matriz X posee una distribución normal con media μ_X , es decir,

$$X \sim N(\mu_X, \sigma^2 I_N \otimes I_K),$$

X es la configuración de los puntos, una vez que a la figura se le ha aplicado una posible traslación, rotación y/o escalamiento. Las coordenadas de la configuración U de X , se construyen en varios pasos, resumidos en la siguiente expresión,

$$LX = Y = TH = rWH = rW(U)H. \quad (1.1)$$

El primer miembro que aparece en la ecuación (1.1) permite realizar una traslación a los puntos originales de la matriz X , al multiplicar esta matriz por la izquierda por una matriz L , donde L puede ser la submatriz de Helmert, la cual está en el espacio $\mathbb{R}^{(N-1) \times K}$ y es ortogonal por renglones al vector $\frac{\mathbf{1}}{n} = \frac{1}{n}(1, \dots, 1)^t$, si $\mu = L\mu_X$ se tiene que,

$$Y \sim N(\mu, \sigma^2 I_{N-1} \times I_K), \quad (1.2)$$

la matriz $Y \in \mathbb{R}^{(N-1) \times K}$ es invariante bajo la traslación de la figura X , es decir la matriz Y contiene coordenadas de rotación-tamaño-y-configuración (del inglés rotation-size-and-shape).

En el segundo y tercer miembro de la ecuación (1.1), se da una factorización de la matriz $Y = TH$, como el producto de dos matrices, la matriz $T \in \mathbb{R}^{(N-1) \times n}$ es una matriz quasi-triangular inferior, donde $n = \min(N-1, K)$, con elementos en la diagonal no negativos $[t_{ii} \geq 0, i = 1, \dots, \min(n, K-1)]$. La matriz $H \in \mathbb{R}^{n \times K}$ es una matriz semiortogonal, donde la factorización de la matriz Y es la factorización QR .

Dado que existen dos condiciones para el parámetro n , cada una de ellas

implica alternativas diferentes. En la primera de ellas se puede considerar que $n = N - 1$, lo cual implica que $N - 1 < K$, de manera tal que $T \in \mathbb{R}^{(N-1) \times (N-1)}$ es una matriz triangular inferior con $t_{ii} \geq 0, i = 1, \dots, N-1$, y la matriz $H \in \mathbb{R}^{(N-1) \times K}$ pertenece al conjunto $V_{N-1,K} = \{H \in \mathbb{R}^{(N-1) \times K} : HH^t = I_{N-1}\}$.

Por otra parte si $n = K$, se tiene que $N - 1 \leq K$, en este caso la matriz $T \in \mathbb{R}^{(N-1) \times K}$ es una matriz quasi-triangular inferior, y la matriz $H \in \mathbb{R}^{K \times K}$ pertenece a $O(n)$, (el conjunto de matrices ortogonales).

La matriz T es una matriz invariante bajo traslación y rotación de la matriz X , y sus elementos son llamados coordenadas de tamaño-y-configuración de X .

En el caso de que $N - 1 \geq K$ la matriz $H \in \mathbb{R}^{K \times K}$ es una matriz cuadrada y se distinguen dos casos:

Caso 1 La matriz H incluye reflexión, $H \in O(K)$, $|H| = \pm 1, t_{kk} \geq 0$, y por definición T se escribe como T^R , ésta contiene coordenadas de reflexión-tamaño-y-configuración.

Caso 2 En este caso H excluye reflexión, $H \in O^+(K)$, $|H| = 1, t_{kk}$ no se restringen y por definición la matriz T se escribe como T^{NR} .

Finalmente cuando $n < K$, se trabaja con la estadística usual multivariada, donde $n = \min(N - 1, k) = N - 1 < k$.

En referencia al cuarto término de la ecuación (1.1), con respecto a la norma de Frobenious se normaliza la matriz Y . La forma de la matriz de configuración W se da al dividir la matriz T por su norma,

$$r = \|T\| = \sqrt{\text{tr } T^t T} = \|Y\|.$$

Alternativamente a utilizar la factorización QR en el análisis de configuraciones, Le, Kendall y Goodall, proponen utilizar la descomposición en valores singulares de la matriz X . Otra alternativa que aquí se propone, es realizar la factorización polar de la matriz X (vea apéndice A). Así en forma análoga (secuencia bajo QR) se tienen los siguientes pasos,

$$LX = Y = RH_1 = rWH_1 = rW(u)H_1,$$

en donde la interpretación de Y, R y $W(u)$ es análoga a la dada en (1.1) y se resume a continuación; la matriz Y contiene las coordenadas de rotación-tamaño-y-configuración, R contiene las coordenadas de tamaño-y-configuración y W contiene las coordenadas de configuración.

En el presente trabajo se plantea inicialmente la importancia de la teoría de configuraciones en diversas ramas de la ciencia y la tecnología; los requisitos matemáticos y las bases estadísticas para trabajar en esta disciplina. Como tema central se propone una técnica alternativa para el análisis de configuraciones basada en la factorización polar.

En el segundo capítulo se inicia con el estudio de algunos conceptos básicos del Álgebra lineal como son: Las transformaciones de traslación, rotación y escalamiento. Posteriormente se estudia la factorización QR , su interpretación geométrica y finalmente el producto de Kronecker.

En el tercer capítulo se describen varios reportes de investigación de la aplicación de la teoría de configuraciones a diversas áreas como son: medicina, biología, agricultura, morfometría, manufactura, topografía y geometría computacional.

En el último capítulo se determinan las distribuciones asociadas a la configuración de un objeto con respecto al tamaño, rotación y traslación. En la parte final se da el resultado más importante de este trabajo con respecto a la parte técnica ya que se determina la distribución asociada al tamaño y forma de la configuración utilizando la factorización polar.

En la parte final de este trabajo se agregan cuatro apéndices que complementan la descripción de los tópicos matemáticos necesarios para el estudio de la teoría de configuraciones, estos temas son el producto exterior, las medidas invariantes, los polinomios zonales y estadística.

PREREQUISITOS

Álgebra lineal

Las definiciones y teoremas presentados en este capítulo se encuentran con mayor detalle en los siguientes textos: Mood, Graybill and Boes (1974), Hoffman y Kunze (1984), Meyer (1992) y Muirhead (1982).

Se inicia con algunos conceptos del álgebra lineal, material indispensable para el desarrollo de este trabajo.

Definición 2.1 . (*Matriz.*) Una matriz es un arreglo de números reales y/o complejos con la siguiente estructura

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

donde a_{ij} denota el ij -ésimo elemento de la matriz.

En adelante se denotarán con letras mayúsculas a las matrices y con minúsculas a sus elementos.

Definición 2.2 . (*Matriz transpuesta.*) Dada una matriz $A \in \mathbb{R}^{r \times s}$, $A = (a_{ij})$ se define la matriz transpuesta de A como una matriz en $\mathbb{R}^{s \times r}$, donde sus elementos son de la forma (a_{ji}) y se denota por medio del símbolo A^t .

Definición 2.3 . (*Traza de una matriz*) La traza de una matriz $A \in \mathbb{R}^{r \times r}$, se define como la suma de los elementos de su diagonal y se denota por medio de $\text{tr } A = \sum a_{ii}$, donde $i = 1, 2, \dots, r$.

Definición 2.4 . (*Matriz identidad*) La matriz identidad se define como la matriz cuadrada cuyos elementos en la diagonal son números unos, y el resto son ceros.

$$I_r = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Definición 2.5 . (*Matriz inversa*) Si A es una matriz cuadrada en $\mathbb{R}^{r \times r}$, se dice que la matriz posee inversa, si existe una matriz $A^{-1} \in \mathbb{R}^{r \times r}$, tal que $AA^{-1} = A^{-1}A = I_r$. Si A^{-1} existe entonces se dice que A es una matriz no singular.

Definición 2.6 . (*Matriz simétrica*) Si A es una matriz cuadrada en $\mathbb{R}^{r \times r}$, se dice que es una matriz simétrica, si la matriz A^t es igual a la matriz A , es decir $A^t = A$.

Definición 2.7 . (*Matriz anti-simétrica*) Si A es una matriz cuadrada en $\mathbb{R}^{r \times r}$, se dice que es una matriz anti-simétrica si $-A^t = A$.

Nota 2.1 . Note que una matriz es anti-simétrica, si los elementos de la diagonal son ceros.

$$A = \begin{pmatrix} 0 & a_{12} & a_{13} & \cdots & a_{1r} \\ a_{12} & 0 & a_{23} & \cdots & a_{2r} \\ a_{13} & a_{23} & 0 & \cdots & a_{3r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{1r} & \cdots & \cdots & \cdots & 0 \end{pmatrix}.$$

Definición 2.8 . (*Matriz triangular*) Una matriz cuadrada $A \in \mathbb{R}^{r \times r}$, se dice que es triangular superior (inferior) si los elementos debajo de la diagonal principal son ceros. (si los elementos sobre la diagonal principal son ceros).

$$a_{ij} = 0, \text{ para } i > j. \quad (a_{ij} = 0, \text{ para } i < j).$$

Definición 2.9 . (*Matriz quasi-triangular*) Una matriz $A \in \mathbb{R}^{r \times s}$, es una matriz quasi-triangular superior (inferior), si

$$a_{ij} = 0, \text{ para } i > j. \quad (a_{ij} = 0, \text{ para } i < j).$$

Ejemplo 2.1 . Para $r = 4$ y $s = 5$, B es una matriz quasi-triangular superior.

$$B = \begin{bmatrix} * & * & * & * & * \\ 0 & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & * & * \end{bmatrix}.$$

Para $r = 4$ y $s = 2$, C también es una matriz quasi-triangular superior.

$$C = \begin{bmatrix} * & * \\ 0 & * \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Definición 2.10 . (Matriz ortogonal) Una matriz cuadrada $A \in \mathbb{R}^{r \times r}$, no singular, se dice que es ortogonal, si $A^t A = A A^t = I_r$.

Proposición 2.1 . El determinante de una matriz ortogonal es igual a ± 1 .

Definición 2.11 . (Variedad de Stiefel) Si $H_1 \in \mathbb{R}^{r \times s}$ con $r \leq s$, es una matriz con columnas ortogonales, es decir $H_1^t H_1 = I_s$. El conjunto de todas las matrices H_1 es llamado la variedad de Stiefel. Se denota por $V_{s,r}$.

$$V_{s,r} = \{H_1 \in \mathbb{R}^{r \times s} : H_1^t H_1 = I_s\}.$$

Nota 2.2 .

1.- Aún cuando la anterior definición de la variedad de Stiefel puede ser utilizada en el contexto de la teoría de configuraciones, la matriz H_1 actúa en general por el lado derecho en las diferentes factorizaciones, con lo cual se obtiene la siguiente definición alternativa, pero equivalente, de la variedad de Stiefel,

$$V_{s,r} = \{Q_1 \in \mathbb{R}^{r \times s} : Q_1 Q_1^t = I_s\}.$$

En donde en este caso las filas de Q_1 son ortogonales. Tanto la matriz Q_1 (filas ortogonales), como la matriz H_1 (columnas ortogonales), son también conocidas en la literatura como matrices semiortogonales. Observe finalmente que $Q_1 = H_1^t$.

2.- Cuando $s = r$, este conjunto es llamado el conjunto ortogonal, y se denota por $O(r)$, a la vez se tiene la siguiente definición $O(r)^+ = \{H \in O(r) : |H| = 1\}$.

Definición 2.12 . (*Matriz definida positiva*) Se dice que una matriz $A \in \mathbb{R}^{r \times r}$ es definida positiva, si la función $f(\mathbf{h}) = \mathbf{h}^t A \mathbf{h} > 0, \forall \mathbf{h}$, donde \mathbf{h} es un vector en $\mathbb{R}^{r \times r}$. (Se considerará de aquí en adelante que $A > 0$, denota una matriz definida positiva).

Definición 2.13 . (*Eigenvalores de una matriz*) Si $A \in \mathbb{R}^{r \times r}$ es una matriz cuadrada, los eigenvalores de la matriz A , se calculan al determinar las raíces del polinomio característico, definido por el determinante de la matriz $\lambda(I_r - A)$, donde λ es un número complejo.

Ejemplo 2.2 . Los eigenvalores de la matriz,

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 4 \\ 0 & 1 & 0 \end{pmatrix},$$

se determinan al resolver la ecuación $\lambda(1 - \lambda)(\lambda - 2) = 0$, siendo las soluciones de dicha ecuación los valores de $\lambda = 0, 1$ y 2 .

Definición 2.14 . (*Norma de Frobenius*) Dada una matriz $A \in \mathbb{R}^{r \times s}$, la norma de Frobenius de la matriz A , se define como la raíz cuadrada positiva de las matrices A^t por A , esto se escribe y se denota como,

$$\|A\|_F = \sqrt{\text{tr } A^t A} = \sqrt{\sum_{i=1}^r \sum_{j=1}^s a_{ij}^2}.$$

Transformaciones asociadas e interpretación geométrica

La siguiente matriz permite definir la traslación de una figura, por medio de una transformación ortogonal por renglones, además se ilustrará cómo se efectúa dicha traslación, ya que no es la usual.

Definición 2.15 . (*Matriz de Helmert*) La matriz de Helmert es una matriz en $\mathbb{R}^{r \times r}$, cuyos renglones son ortogonales al primer renglón $\frac{1}{r} = \frac{1}{r}(1, 1, \dots, 1)$, y a la vez estos renglones son ortonormales. Se denota por H y está dada por,

$$\begin{pmatrix} \frac{1}{r} & \frac{1}{r} & \frac{1}{r} & \frac{1}{r} & \frac{1}{r} & \dots & \frac{1}{r} \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & \dots & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & 0 & \dots & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{\sqrt{(r-1)r}} & \frac{1}{\sqrt{(r-1)r}} & \frac{1}{\sqrt{(r-1)r}} & \frac{1}{\sqrt{(r-1)r}} & \frac{1}{\sqrt{(r-1)r}} & \dots & \frac{r-1}{\sqrt{(r-1)r}} \end{pmatrix}$$

Se dice que L es una sub-matriz de Helmert, si L se define a partir de la matriz de Helmert, omitiendo el renglón $\frac{1}{r}$. Es decir L pertenece a $\mathbb{R}^{(r-1) \times r}$.

Esta matriz permite definir la traslación de un conjunto de puntos en el plano, de una forma diferente a la tradicional. Por otra parte se da la invarianza bajo dicha transformación, como se ejemplifica a continuación.

$$\text{Sean } A = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ \vdots & \vdots \\ x_r & y_r \end{pmatrix} \text{ y } B = \begin{pmatrix} x_1 + c & y_1 + d \\ x_2 + c & y_2 + d \\ x_3 + c & y_3 + d \\ \vdots & \vdots \\ x_r + c & y_r + d \end{pmatrix},$$

dos matrices, que representan r puntos en el plano. Como se puede apreciar B representa la matriz A trasladada, c unidades en el eje X , y d unidades en el eje Y .

Ahora al operar a la matriz B por la izquierda, utilizando la sub-matriz de Helmert en $\mathbb{R}^{(r-1) \times 2}$, se concluye que $LB = LA$, de manera que esta matriz permite la invarianza bajo la transformación de traslación.

Para ilustrar lo anterior, operemos sobre el r -ésimo renglón de la matriz L y las columnas de la matrices A y B , generando la componente (la_{r1}) y (lb_{r1}) .

Entonces,

$$la_{r1} = \left(\frac{1}{\sqrt{(r-1)r}} \right) [x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r].$$

Por otra parte,

$$\begin{aligned} lb_{r1} &= \left(\frac{1}{\sqrt{(r-1)r}} \right) [(x_1 + c) + (x_2 + c) + \cdots + (x_{r-1} + c) - (r-1)(x_r + c)], \\ &= [(x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r] + [(r-1)c - (r-1)c], \\ &= \left(\frac{1}{\sqrt{(r-1)r}} \right) [x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r] + [0], \\ &= la_{r1}. \end{aligned}$$

Como se mencionó anteriormente, dicha traslación no es la tradicional. Hecho que también se puede observar en el siguiente ejemplo, donde se le aplica la transformación de la sub-matriz de Helmert a dos conjuntos de puntos que difieren por una traslación.

Ahora al operar a la matriz B por la izquierda, utilizando la sub-matriz de Helmert en $\mathbb{R}^{(r-1) \times 2}$, se concluye que $LB = LA$, de manera que esta matriz permite la invarianza bajo la transformación de traslación.

Para ilustrar lo anterior, operemos sobre el r -ésimo renglón de la matriz L y las columnas de la matrices A y B , generando la componente (la_{r1}) y (lb_{r1}) . Entonces,

$$la_{r1} = \left(\frac{1}{\sqrt{(r-1)r}} \right) [x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r].$$

Por otra parte,

$$\begin{aligned} lb_{r1} &= \left(\frac{1}{\sqrt{(r-1)r}} \right) [(x_1 + c) + (x_2 + c) + \cdots + (x_{r-1} + c) - (r-1)(x_r + c)], \\ &= [(x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r] + [(r-1)c - (r-1)c], \\ &= \left(\frac{1}{\sqrt{(r-1)r}} \right) [x_1 + x_2 + \cdots + x_{r-1} - (r-1)x_r] + [0], \\ &= la_{r1}. \end{aligned}$$

Como se mencionó anteriormente, dicha traslación no es la tradicional. Hecho que también se puede observar en el siguiente ejemplo, donde se le aplica la transformación de la sub-matriz de Helmert a dos conjuntos de puntos que difieren por una traslación.

Ejemplo 2.3 . Dados los siguientes conjuntos de puntos en el plano,

$$\{(-2, 1), (-1, -1), (-1, 2), (1, 1), (2, 3), (3, -1), (3, 2)\},$$

y

$$\{(-3, 1), (-2, -1), (-2, 2), (0, 1), (1, 3), (2, -1), (2, 2)\}.$$

Estos se puede representar por medio de los siguientes arreglos matriciales en $\mathbb{R}^{2 \times 7}$.

$$A = \begin{pmatrix} -2 & 1 \\ -1 & -1 \\ -1 & 2 \\ 1 & 1 \\ 2 & 3 \\ 3 & -1 \\ 3 & 2 \end{pmatrix}, \quad \text{y} \quad B = \begin{pmatrix} -3 & 1 \\ -2 & -1 \\ -2 & 2 \\ 0 & 1 \\ 1 & 3 \\ 2 & -1 \\ 2 & 2 \end{pmatrix}.$$

Como se muestra en la Figura (2.1) estos puntos difieren tan sólo por una traslación en el eje X de una unidad, es decir, los puntos del arreglo B , se localizan una unidad hacia la izquierda con respecto a los del arreglo A .

Ahora al aplicar la submatriz de Helmert L ,

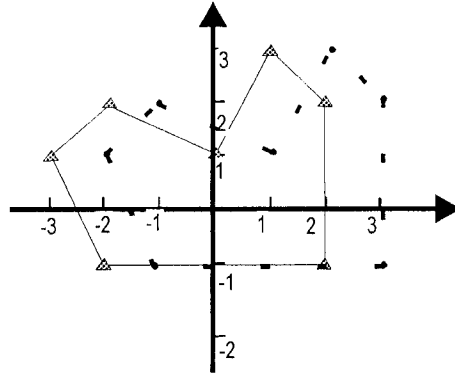


Figura 2.1 Puntos en el plano

$$L = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 & 0 & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{-3}{\sqrt{12}} & 0 & 0 \\ \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{1}{\sqrt{20}} & \frac{-4}{\sqrt{20}} & 0 \\ \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{30}} & \frac{1}{\sqrt{30}} & \frac{-5}{\sqrt{30}} \end{pmatrix},$$

se obtiene el siguiente resultado

$$C = LA = LB = \begin{pmatrix} -0.7071 & 1.4142 \\ -0.4082 & -1.6330 \\ -2.0207 & -0.2887 \\ -2.4597 & -2.0125 \\ -2.9212 & 2.0083 \\ -2.4689 & -1.0801 \end{pmatrix}$$

En la Figura (2.2) se pueden observar dos cosas, una de ellas es que el número de puntos disminuye, esto se debe a que cuando se trasladan los puntos, se pierde uno al considerar a uno de ellos como el punto central. Por otra parte la forma original de la figura se modifica y esto es lo que se había mencionado anteriormente, de que no es la traslación usual.

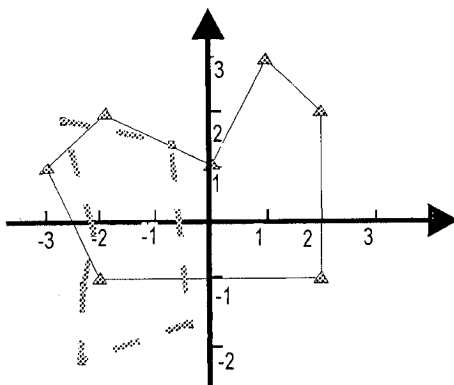


Figura 2.2 *Puntos trasladados*

A continuación se dan las definiciones de matriz de rotación y traslación, junto con algunos ejemplos que ilustran la forma de trabajar.

Se exponen dos transformaciones lineales relacionadas con el análisis de

configuraciones. Una de ellas es la de rotación de la configuración original y la otra con respecto a una posible reflexión.

Primeramente se expondrá la forma de rotar la configuración del sistema en el plano bidimensional. Esto se define a partir de una matriz ortogonal en donde el ángulo de rotación queda implícita en la matriz de rotación.

En \mathbb{R}^2 la matriz de rotación esta dada por,

$$P(\theta) = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ \operatorname{sen} \theta & \cos \theta \end{bmatrix},$$

y P es ortogonal.

El siguiente ejemplo muestra como se aplica esta transformación a dos parejas ordenadas de puntos en el plano, denotando a estas parejas como vectores.

Ejemplo 2.4 . (Rotación)

Sean $V_1 = (1, -2)^t$ y $V_2 = (1, 1)^t$ vectores en el plano, al aplicar la matriz de rotación con $\theta = 30$ grados. Entonces se generan dos vectores W_1 y W_2 ,

$$W_1 = R(30)V_1 = \begin{bmatrix} \cos(30) & -\operatorname{sen}(30) \\ \operatorname{sen}(30) & \cos(30) \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} .86 & -0.50 \\ 0.50 & 0.86 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$W_2 = R(30)V_2 = \begin{bmatrix} \cos(30) & -\operatorname{sen}(30) \\ \operatorname{sen}(30) & \cos(30) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} .86 & -0.50 \\ 0.50 & 0.86 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

De aquí,

$$W_1 = \begin{bmatrix} 1.86 \\ -1.23 \end{bmatrix} \text{ y } W_2 = \begin{bmatrix} 0.36 \\ 1.36 \end{bmatrix}.$$

La Figura (2.3) muestra la rotación de los vectores V_1 y V_2 en los vectores W_1 y W_2 .

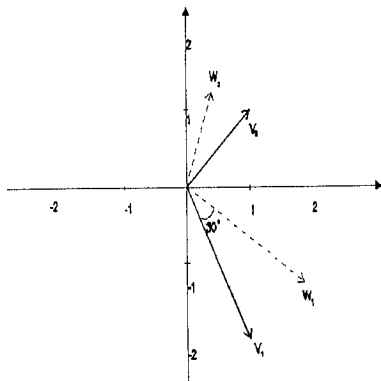


Figura 2.3 Rotación de vectores

Continuando con el ejemplo (2.3) y al aplicar la rotación a la matriz C , con un ángulo de 90 grados, se obtiene la siguiente matriz

$$D = \begin{pmatrix} -1.4142 & -0.7071 \\ 1.6330 & -0.4082 \\ 0.2887 & -2.0207 \\ 2.0125 & -2.4597 \\ -2.0083 & -2.9212 \\ 1.0801 & -2.4689 \end{pmatrix}.$$

La Figura (2.4) muestra la rotación de las componentes de la matriz C y de la nueva composición alcanzada por la matriz D .

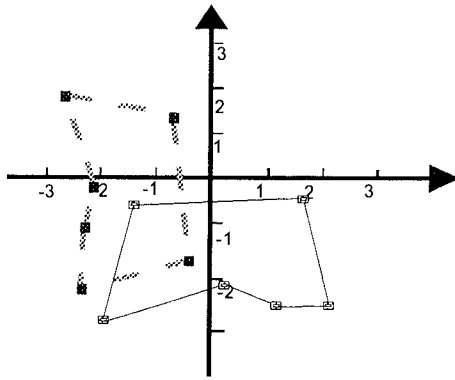


Figura 2.4 *Puntos rotados-y-trasladados*

Otra transformación importante se da por medio de las reflexiones alrededor de una línea recta, donde por reflexión se entiende proyectar perpendicularmente los puntos de una figura con respecto a una línea, como sucede con un espejo, de manera que la imagen aparece del otro lado pero en forma invertida. La transformación es expresada en función de la matriz de Householder.

El siguiente ejemplo muestra como se aplica esta transformación.

Ejemplo 2.5 . En este ejemplo se aplica la transformación de reflexión al vector $V_1 = (1, -2)^t$ con respecto a la recta identidad. Para esto se utiliza al vector $V = (-1, 1)^t$ que pertenece a la recta y de esta forma la matriz P (de Householder) esta bien definida. La matriz P se define de la siguiente forma

$$P = I - \frac{2VV^t}{V^tV} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

de manera tal que el vector de transformación W_1 es igual a,

$$W_1 = PV_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

La Figura (2.5) muestra la reflexión del vector V_1 con respecto a la recta l y se tiene como resultado al vector W_1 .

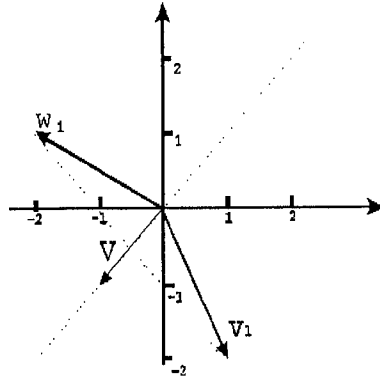


Figura 2.5 Reflexión de vectores

Producto de Kronecker y vectorización de una matriz

Los siguientes dos conceptos son los del producto de Kronecker entre dos matrices y el de vectorización de una matriz.

Definición 2.16 . (Producto Kronecker) Si $A = (a_{ij})$ es una matriz en $\mathbb{R}^{p \times q}$ y $B = (b_{ij})$ es una matriz en $\mathbb{R}^{r \times s}$. El producto de Kronecker de A y B se denota y define como,

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1q}B \\ a_{21}B & a_{22}B & \dots & a_{2q}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}B & a_{p2}B & \dots & a_{pq}B \end{bmatrix}.$$

donde $A \otimes B$ es una matriz en $\mathbb{R}^{pr \times qs}$. A este producto también se le llama el producto directo.

De la definición anterior se genera la siguiente expresión,

$$I_p \otimes B = \begin{bmatrix} B & 0 & \dots & 0 \\ 0 & B & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B \end{bmatrix},$$

donde se multiplica la matriz I_p por la matriz $B \in \mathbb{R}^{r \times s}$.

Definición 2.17 (*Vectorización*) Si $T = [t_1 \ t_2 \ \dots \ t_q]$ es una matriz en $\mathbb{R}^{p \times q}$ cuyas columnas son vectores en \mathbb{R}^p , entonces la vectorización de la matriz T se define como;

$$\text{vec}(T) = \begin{pmatrix} t_1^t \\ t_2^t \\ \vdots \\ t_q^t \end{pmatrix} \in \mathbb{R}^{pq \times 1},$$

donde cada t_i es un vector en \mathbb{R}^p é $i = 1, \dots, q$.

A continuación se citan algunas propiedades asociadas a estas definiciones las cuales son utilizadas al calcular algunas distribuciones multivariadas no centradas.

Proposición 2.2 . (Propiedades) Si A, B y C son matrices, α y β son constantes, entonces se tienen las siguientes propiedades...

$$(a) \quad (\alpha A) \otimes (\beta B) = \alpha\beta(A \otimes B).$$

$$(b) \quad (A + B) \otimes C = A \otimes C + B \otimes C.$$

$$(c) \quad (A \otimes B) \otimes C = A \otimes (B \otimes C).$$

$$(d) \quad (A \otimes B)^t = A^t \otimes B^t.$$

$$(e) \quad \text{tr}(A \otimes B) = (\text{tr } A)(\text{tr } B).$$

$$(f) \quad \det(A \otimes B) = (\det A)^n (\det B)^m.$$

$$(g) \quad \text{vec}(BC) = (I \otimes B) \text{vec } C = (C^t \otimes I) \text{vec } B.$$

Se demostrará únicamente el inciso (a).

Demostración:

$$(\alpha A) \otimes (\beta B) = \begin{vmatrix} (\alpha a_{11})(\beta B) & (\alpha a_{12})(\beta B) & \cdots & (\alpha a_{1q})(\beta B) \\ (\alpha a_{21})(\beta B) & (\alpha a_{22})(\beta B) & \cdots & (\alpha a_{2q})(\beta B) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha a_{p1})(\beta B) & (\alpha a_{p2})(\beta B) & \cdots & (\alpha a_{pq})(\beta B) \end{vmatrix},$$

$$\begin{aligned}
&= (\alpha\beta) \begin{vmatrix} a_{11}B & a_{12}B & \cdots & a_{1q}B \\ a_{21}B & a_{21}B & \cdots & a_{21}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}B & a_{p1}B & \cdots & a_{pq}B \end{vmatrix}, \\
&= \alpha\beta(A \otimes B).
\end{aligned}$$

■

Factorización QR

En esta sección se enuncia el teorema de la factorización QR por medio del cual es posible factorizar una matriz X como el producto de dos matrices, una ortogonal y la otra quasi-triangular superior.

La importancia de dicha factorización radica en poder realizar cálculos rápidos de eigenvalores, eigenvectores y en este trabajo de calcular jacobianos, por medio del producto exterior.

Existen diferentes métodos de calcular la factorización QR entre los cuales se encuentran los siguientes: Transformaciones de Householder, Bloques de Householder, Método de Givens, Transformaciones rápidas de Givens y Proceso de Ortogonalización de Gram-Schmidt(G-S).

Aunque cada uno de los métodos produce la misma factorización, la diferencia radica en la rapidez de llegar a la factorización requerida. En este trabajo se utilizará el proceso de ortogonalización G-S.

Teorema 2.1 . (Factorización QR) Si $X \in \mathbb{R}^{r \times s}$ es una matriz de rango s ($s \leq r$), entonces X se puede factorizar en forma única como, $X = QR$, donde $Q \in \mathbb{R}^{r \times s}$ es una matriz semiortogonal (columnas ortogonales) y $R \in \mathbb{R}^{s \times s}$ es una matriz triangular superior, con elementos en la diagonal positivos $t_{ii} > 0$.

Demostración:

Expresa la matriz X en base a sus vectores columnas x_1, x_2, \dots, x_s que la generan, por el proceso de ortogonalización G-S, la nueva base de vectores ortonormalizados y_1, y_2, \dots, y_s , quedan expresados por medio de;

$$\begin{aligned} y_1 &= x_1. \\ y_2 &= x_2 - t_{12}y_1. \\ y_3 &= x_3 - t_{13}y_1 - t_{23}y_2. \\ &\vdots = \vdots \\ y_s &= x_{s-1} - \sum_{i=1}^{s-1} t_{ij}y_i. \end{aligned}$$

Reescribiendo a los vectores x en función de los vectores y , se tiene que;

$$\begin{aligned} x_1 &= y_1. \\ x_2 &= t_{12}y_1 + y_2. \\ x_3 &= t_{13}y_1 + t_{23}y_2 + y_3. \\ &\vdots = \vdots \\ x_s &= \sum_{i=1}^{s-1} t_{ij}y_i + y_{s-1}. \end{aligned}$$

Esto se puede escribir como;

$$[x_1, x_2, x_3, \dots, x_s] = [y_1, y_2, y_3, \dots, y_s] \begin{bmatrix} 1 & t_{12} & t_{13} & \cdots & t_{1,r-1} \\ 0 & 1 & t_{23} & \cdots & t_{2,r-1} \\ 0 & 0 & 1 & \cdots & t_{3,r-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Sea $Q = [y_1, y_2, \dots, y_s]$ y R la matriz de la derecha, obteniendo el resultado deseado. ■

Si la matriz X es de rango $g \leq s$, eso significa que $s-g$ columnas de la matriz Y en el proceso de ortogonalización de G-S son cero, de manera tal que eliminando estas columnas y sus correspondientes renglones en la matriz T , el teorema sigue siendo válido, sólo que $Q \in \mathbb{R}^{r \times g}$ y $R \in \mathbb{R}^{g \times s}$ ($t_{ii} \leq 0$). En este caso $Q \in V_{g,r}$.

Si la matriz X es de rango $g = r$ eso significa que $g = r$, de manera tal que el rango sigue siendo s y la matriz $Q \in \mathbb{R}^{r \times r}$ es cuadrada. Aquí $Q \in O(r)$ y se distinguen dos casos:

(i) Si $Q \in O(r)$, $|H| = \pm 1$, $t_{ii} \leq 0$ y T se escribe como T^R .

(ii) Si $Q \in O^+(r)$, $|H| = 1$, t_{ii} no están restringidos y T se escribe como T^{NR} .

Nota 2.3 . Observe que si $Z = X^t$ en el teorema 2.1. se obtiene la siguiente forma alternativa de la factorización QR ; $Z = TH_1$, en donde T es una matriz triangular inferior y H_1 es una matriz semiortogonal (filas ortogonales).

Ejemplo 2.6 . Continuando con los ejemplos (2.3) y (2.5) y aplicando la factorización QR a las matrices C y D se obtienen las siguientes factorizaciones.

$$C = \begin{pmatrix} -0.7071 & 1.4142 \\ -0.4082 & -1.6330 \\ -2.0207 & -0.2887 \\ -2.4597 & -2.0125 \\ -2.9212 & 2.0083 \\ -2.4689 & -1.0801 \end{pmatrix}$$

$$= \begin{pmatrix} 1.5811 & 0 \\ -1.2780 & -1.0954 \\ 0.6455 & -1.9365 \\ -0.7000 & -3.1000 \\ 3.1027 & -1.7146 \\ 0.1380 & -2.6913 \end{pmatrix} \begin{pmatrix} -0.4472 & 0.8944 \\ 0.8944 & 0.4472 \end{pmatrix} = R_1 Q_1.$$

Por otra parte, la factorización de la matriz D esta dada por,

$$\begin{aligned}
 D &= \begin{pmatrix} -1.4142 & -0.7071 \\ 1.6330 & -0.4082 \\ 0.2887 & -2.0207 \\ 2.0125 & -2.4597 \\ -2.0083 & -2.9212 \\ 1.0801 & -2.4689 \end{pmatrix} \\
 &= \begin{pmatrix} 1.5811 & 0 \\ -1.2780 & -1.0954 \\ 0.6455 & -1.9365 \\ -0.7000 & -3.1000 \\ 3.1027 & -1.7146 \\ 0.1380 & -2.6913 \end{pmatrix} \begin{pmatrix} -0.8944 & -0.4472 \\ -0.4472 & 0.8944 \end{pmatrix} = R_2 Q_2.
 \end{aligned}$$

Se puede observar, que $R_2 = R_1$, de tal forma que se puede afirmar que la transformación de rotación es invariante sobre los datos al obtener la misma matriz triangular. Se dice que la matriz D contiene coordenadas de traslación y rotación, debido a las propiedades de invarianza.

ÁREAS DE APLICACIÓN

A continuación se describen algunas investigaciones de la teoría de configuraciones; en cada una de ellas se ejemplifica la forma de trabajo.

Configuración de una ciudad

Existe el postulado que afirma que la configuración de la ciudad no cambia con el paso del tiempo y las diferencias que se presentan no son significativas.

Para una persona que vive en la ciudad es importante llegar a conocer las tendencias de crecimiento de la ciudad, ya que esto le representa beneficios económicos en cuanto a la plusvalía del lugar, áreas comerciales más abundantes o un aumento considerable de servicios públicos como: bancos, lugares de diversión, áreas verdes, ect. Por otra parte a los comerciantes les interesa saber hacia dónde crece una ciudad dado que les permite establecer puntos estratégicos de posibles lugares comerciales a futuro.

Bajo las circunstancias anteriores es factible utilizar la teoría de configuraciones para analizar la configuración de una ciudad a través del tiempo.

Brown, Foote y Holz (1997), realizaron una investigación con respecto a la ciudad de Austin Texas, con el objetivo de determinar si el postulado era falso.

Ellos trabajaron con fotografías aéreas de alta resolución, y se contó con tan sólo cuatro de ellas debido a la dificultad de conseguirlas, los años que se incluyen son los de 1964, 1973, 1980 y 1990. A la vez definieron la configuración de la ciudad en base al avance urbano y no a los límites geopolíticos que la delimitan.

En la Figura (3.1) se muestra la ciudad de Austin en los años de 1964, 1973, 1980 y 1990. Por otra parte en la Figura (3.2) se muestran todas las figuras juntas. La figura de 1964, se expandió con el objetivo de compararla con la de 1990, más sin embargo no fue posible establecer diferencias significativas, aún y cuando se trató de que cubrieran áreas iguales. En relación a las diferencias que aparecieron, se hallaron dos factores: el medio de transporte y las condiciones geológicas del lugar.

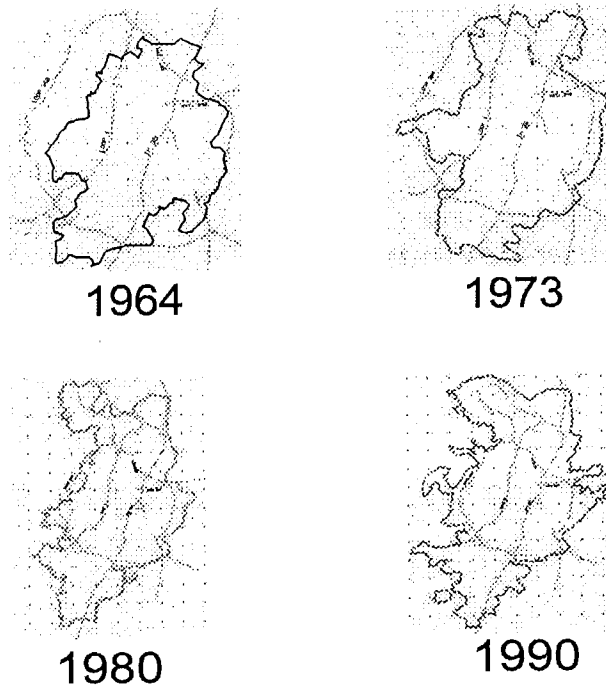


Figura 3.1 *Contornos de la ciudad separados por año*

Finalmente no pudieron concluir si el postulado es falso debido a la difi-

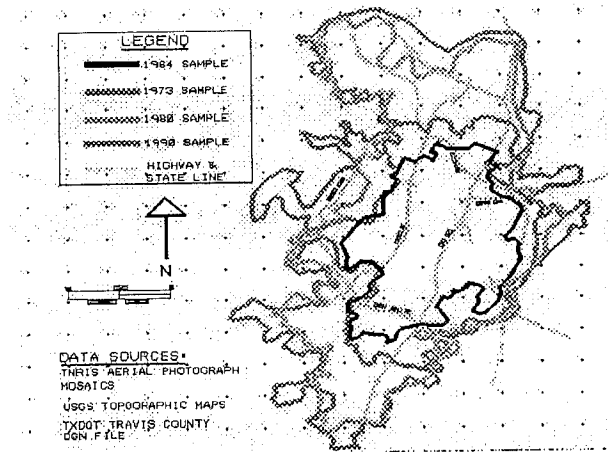


Figura 3.2 *Contornos de la ciudad acoplados*

cultad al momento de cuantificar la figura, de aquí que se proponga el uso de la teoría de configuraciones para determinar si dos configuraciones de una ciudad son equivalentes.

Geometría Computacional

Cootes y Taylor (1996), muestran que el estudio de configuraciones está ligado a las técnicas de la Geometría Computacional. Una aplicación de esta técnica se presenta en el reconocimiento de huellas digitales, donde el objetivo es poder reconocer la huella digital de una persona, ya que la huella dactilar identifica a cada persona de manera precisa, desde el momento en que ésta nace.

En el caso de un crimen, las huellas digitales del asesino pueden quedar impresas en algún objeto y de ahí se puede obtener la huella del criminal, para posteriormente compararla con un banco de huellas e identificarla. Otra aplicación se da al momento de que una persona muere y es necesario reconocerla.

Hoy en día esta técnica es usada con más frecuencia debido a la modernidad de los sistemas de cómputo, donde es posible digitalizar la huella y de esta forma definir la configuración de la misma.

Por último es importante recalcar que en muchas de las ocasiones la huella a reconocer no está completa, se encuentra rotada, está expandida, etc., la teoría de configuraciones toma en cuenta estas modificaciones, con lo cual este modelo es de gran ayuda para el investigador.

Manufactura

Goodall y Mardia (1993), exponen la importancia de esta teoría en el proceso de manufactura de sillas, en el ejemplo que presenta, delimita la configuración de la silla espacio tridimensional en base a 10 puntos que pertenecen a las patas y al respaldo de la misma.

La importancia de esta aplicación radica en poder determinar si las sillas cumplen con un patrón de calidad con respecto a los puntos señalados sobre la silla, sin importar en un momento dado el tamaño de la silla, pero sí que mantengan el mismo patrón.

Medicina

Goodall (1991), menciona la importancia del uso de la teoría de configuraciones dentro del área de la medicina para el seguimiento de un tratamiento.

Por ejemplo si el objeto de interés es un tumor que apareció en un individuo, el médico está interesado en llegar a saber si al aplicar cierto tratamiento al tumor éste reduzca, crezca o mantenga su tamaño. De manera tal que le permita saber si el paciente no corre algún peligro y continuar aplicando dicho tratamiento a otros pacientes.

Pero a la vez se corre el riesgo de que este tumor cambie su posición, crezca o rote. En estos casos puede ocurrir que alguno de ellos por si solos no ocasione trastornos al sistema; pero a la vez puede ocurrir que uno solo de estos factores o varios de ellos combinados a la vez afecten al individuo y el tratamiento no dé el resultado esperado, y sea necesario extirpar el tumor por medio de una intervención quirúrgica o un método indirecto, como la radioterapia.

En el caso de una intervención quirúrgica, se corre el riesgo de que el tumor aún y cuando sólo se haya trasladado, se mueva a una parte delicada del organismo, ocasionando una operación riesgosa.

De manera tal que el poder determinar la configuración del tumor y su desenvolvimiento con el paso del tiempo, es de gran ayuda para poder interpretar el comportamiento del tumor y así tener mejor información para tomar una decisión oportuna.

Morfometría

Lewis, Kendall y Thomson (1992), comentan la importancia del uso de la teoría de configuraciones dentro de la Morfometría. Aquí se estudian las diferen-

cias que se presentan en cuanto a los aspectos métricos biológicos relacionados con la forma de un sujeto. Variaciones de este tipo pueden suceder en un ambiente prenatal, donde los factores genéticos desempeñan un papel importante para el desenvolvimiento de la forma del bebé y sus características particulares, tales como el tamaño, peso, forma de los ojos, etc.

Uno de los objetivos de la teoría de configuraciones sobre esta área es llegar a determinar si un objeto a evolucionado genéticamente o a mantenido su forma original con el paso del tiempo.

En el caso de un ser vivo, se espera que evolucione genéticamente, mientras que en el caso de un cristal, la expectativa radica en probar que mantiene su misma estructura bajo las mismas condiciones ambientales al momento de su formación.

Sistemas de información

El manejo de datos geográficos resulta ser un campo muy importante en la recopilación, almacenamiento y organización de datos. A la vez a resultado ser un campo rico de investigación. Teniendo un amplio desarrollo en los últimos años, debido a la gran cantidad de información que proviene de diversos medios, tales como: sistemas computacionales, bases de datos, censos, etc., con el objeto de tomar decisiones oportunas.

Wentz (1994), describe la importancia de entender los factores que afectan los procesos espaciales haciendo uso de los sistemas de información geográfica (Geographics Information Systems, GIS). Estos almacenan y recopilan la información

de manera eficiente, con el fin de generar bases de datos confiables. Por otra parte, debido al crecimiento continuo de las comunidades se requiere de técnicas analíticas más avanzadas, con el fin de hacer uso de ellos de manera más rápida y oportuna.

Aquí se pretenden medir las diferencias en el medio: como el área, perímetro o la forma de una configuración, con el fin de poder realizar comparaciones. Hasta el momento, con el sistema GIS, se ha podido medir el área y perímetro de una configuración sin mucha dificultad. Sin embargo al tratar de cuantificar la forma del mismo, esto no se ha podido debido a la complejidad del concepto. Es aquí donde se propone incorporar la teoría de configuraciones en el manejo de datos ya que esta disciplina permite medir la configuración del objeto.

Reconocimiento de caras por medio de fotografías

Aykroyd, Bowie y Mardia (1997), propusieron un estudio relacionado con la configuración de un rostro humano en base a fotografías faciales.

El objetivo del trabajo fue llegar a reconocer una cara por medio de métodos estadísticos. El camino a seguir consistió en definir la configuración del rostro a partir de tres fotografías de una persona bajo condiciones controladas y se consideraron cierto número de puntos sobresalientes de cada rostro humano.

Esta propuesta de trabajo abre una línea de investigación muy interesante que se puede relacionar con la identificación de cadáveres, personas accidentadas, etc. Debido a que en muchas de estas circunstancias resulta difícil identificar a la persona.

Reconocimiento de firmas

Otro rasgo donde se hace patente la necesidad de comparar dos figuras, se presenta al tratar de identificar la firma de una persona, con el objetivo de determinar si ésta es auténtica. La comparación se realiza con respecto a cierto patrón preestablecido. Motivos sobran para tratar de realizar dicha comparación como: el caso del cajero al cambiar un cheque, legalizar un documento oficial, etc.

Goodall (1991), da un ejemplo de una investigación realizada en los laboratorios AT & T Bell Laboratories, en donde desarrollaron un modelo estadístico para el reconocimiento de firmas. Este se basa en tomar un promedio de firmas de un individuo en particular, con el objetivo de obtener un modelo patrón de comparación.

Ellos obtuvieron su modelo a partir de una computadora, haciendo que una persona firmará sobre un pad, el cual permite registrar puntualmente la trayectoria seguida por la firma de la persona, ésta se observa y se guarda en la computadora con respecto a las coordenadas observadas en la pantalla. Se le pidió a dicha persona que firmara cierto número de veces y en base a los rasgos se encontró al modelo patrón.

La idea propuesta por ellos para definir la configuración de la firma no está lejos de la realidad considerando , que aún cuando la persona realice la firma en dos ocasiones, ésta sufre cierta modificación, cuyas diferencias no son significativas y ambas se pueden considerar auténticas.

En el artículo no se dan conclusiones sólo se especifican más detalles técnicos

de la investigación, sin embargo dicho modelo coincide con fundamentos de la teoría de configuraciones.

Topografía

Otro campo de investigación relacionado con la teoría de configuraciones se da cuando se analiza cierta región geográfica con un objetivo específico, llámese del tipo forestal o urbano. Un ejemplo urbano ya se mencionó anteriormente y se explicó detalladamente.

Al igual que en el problema de urbanización el objetivo es poder determinar si cierta región ha cambiado significativamente con el paso del tiempo, por ejemplo si cierta región vegetal se ha deteriorado o se ha conservado.

Otra técnica diferente consiste en guardar de una fotografía sus rasgos característicos punto a punto. Esto se realiza al cuadricular una fotografía del lugar y guardar la información en un arreglo matricial con respecto a los colores, donde a cada color se le asigna un número en particular. Una vez registrada la información en este arreglo matricial, se pueden utilizar las técnicas de la teoría de configuraciones sin realizar grandes cambios y con el supuesto de normalidad de los datos.

Como las fotografías son aéreas, es imposible igualar la misma toma aún y cuando se tomen muchas precauciones. Las diferencias se deben a la modernización de la tecnología, por otra parte el avión pudo llegar a volar más alto o más bajo, dando con esto cambios en la escala de la fotografía o con respecto al punto de referencia y sufriendo esta toma una traslación. Finalmente puede suceder que

la fotografía aparezca un poco rotada, de manera tal que es de interés para el investigador el poder pasar por alto estas tres condiciones.

DENSIDADES ASOCIADAS A LA TEORÍA DE CONFIGURACIONES

En este capítulo se determinan las densidades asociadas a las transformaciones asociadas de traslación, rotación y escalamiento.

En relación al desarrollo expuesto en el Capítulo 1, la matriz W puede incluir reflexión o excluir reflexión, W^{NR} .

Dado que $\|W\| = 1$, los elementos de W son vectores de dirección para la configuración y \mathbf{U} contiene $m = (N - 1)k - nk + \frac{n(n + 1)}{2} - 1$ coordenadas polares generalizadas.

Como se mencionó anteriormente, cuando $n = N - 1 \leq K$, es un caso del análisis multivariado y se tiene que la matriz $B = YY^t \in \mathbb{R}^{n \times n}$ posee una distribución Whisart no centrada, a saber

$$B \sim W_n(K, \sigma^2 I; \mu\mu^t),$$

la cual es invariante bajo orientación y reflexión, ésta distribución hipergeométrica es del tipo Bessel ${}_0F_1$ y caracteriza a la distribución de tamaño-y-configuración.

Cuando $N - 1 \geq K$, la distribución Whisart no centrada no existe (formalmente la densidad no existe con respecto a la medida de Lebesgue, definida sobre el conjunto de las matrices definidas positivas), de manera que no se puede utilizar

el análisis multivariado y esta es una de las razones para hacer uso del análisis de configuraciones.

La distribución de rotación-tamaño-y-configuración esta dada por (1.2). A continuación se encuentra la distribución de tamaño-y-configuración, bajo el supuesto de que el número de puntos sea mayor a la dimensión ($N - 1 > K$).

Distribución de tamaño-y-configuración

Teorema 4.1 . Si $Y = TH$, donde TH es la factorización QR de la matriz Y , aquí $Y \in \mathbb{R}^{(N-1) \times K}$, $T \in \mathbb{R}^{(N-1) \times n}$ es una matriz semitriangular inferior con elementos en la diagonal no negativos ($t_{ii} \geq 0, i = 1, \dots, \min(n, K - 1)$), donde $n = \min(N-1, K)$ y la matriz $H \in \mathbb{R}^{n \times K}$ es una matriz semiortogonal. Entonces la distribución asociada es del tipo hipergeométrica ${}_0F_1$ con densidad dado por,

$$f_T(T) = \frac{\prod_{i=1}^n t_{ii}^{K-i}}{\sigma^M 2^{\frac{M-2n}{2}} \pi^{\frac{M-nK}{2}} \Gamma_n\left(\frac{K}{2}\right)} \exp\left(-\frac{1}{2\sigma^2} \text{tr}(TT^t - \mu^t \mu)\right) {}_0F_1\left(\frac{K}{2}; A\right), \quad (4.1)$$

donde $M = (N - 1)K$ y $A = \frac{\mu^t T T^t \mu}{4\sigma^4}$.

Demostración

Dado que $Y \sim N(\mu_X, \sigma^2 I_N \otimes I_K)$ se sigue que,

$$dF_Y(Y) = \frac{1}{(2\pi)^{\frac{M}{2}} \sigma^M} \exp\left[-\frac{1}{2\sigma^2} \text{tr}(Y - \mu)^t (Y - \mu)\right] (dY). \quad (4.2)$$

Dada la factorización QR de $Y = TH$, $T \in \mathbb{R}^{(N-1) \times n}$ y $H \in V_{n,k}$, donde $n = \min(N-1, K)$, se tiene que

$$(dY) = \prod_{i=1}^n t_{ii}^{K-i} (dT)(HdH^t). \quad (4.3)$$

Por otra parte la densidad conjunta de (T, H) esta dada por,

$$dF_{T,H}(T, H) = dF_{Y=TH}(TH)J(Y \rightarrow TH)(dY). \quad (4.4)$$

Sustituyendo (4.2) y (4.3) en (4.4) se llega a que,

$$dF_{T,H}(T, H) = \frac{1}{(2\pi)^{\frac{M}{2}} \sigma^M} \exp\left(-\frac{1}{2\sigma^2} \text{tr}(TH - \mu)^t(TH - \mu)\right) \prod_{i=1}^n t_{ii}^{K-i} (dT)(HdH^t). \quad (4.5)$$

Ahora desarrollando el argumento de la función exponencial en la expresión (4.5),

$$\begin{aligned} -\frac{1}{2\sigma^2} \text{tr}(TH - \mu)^t(TH - \mu) &= -\frac{1}{2\sigma^2} \text{tr}(H^t T^t TH - H^t T \mu - \mu^t TH + \mu^t \mu), \\ &= -\frac{1}{2\sigma^2} \text{tr}(T^t TH H^t - 2\mu^t TH + \mu^t \mu), \\ &= \text{tr}\left(-\frac{T^t T}{2\sigma^2} - \frac{\mu^t \mu}{2\sigma^2} + \frac{\mu^t TH}{\sigma^2}\right). \end{aligned}$$

De manera tal que la densidad conjunta de T y H está dada por,

$$dF_{T,H}(T, H) = \frac{1}{(2\pi)^{\frac{M}{2}} \sigma^M} \exp\left[-\frac{1}{2\sigma^2} \text{tr}(T^t T + \mu^t \mu)\right] \exp\left[\text{tr}\left(\frac{\mu^t TH}{\sigma^2}\right)\right] \prod_{i=1}^n t_{ii}^{K-i} (dT)(HdH^t).$$

Para encontrar la marginal de T , se integra con respecto a H y del lema

C.1. se tiene que,

$$\begin{aligned}
 \int_{V_{n,K}} \exp\left(\operatorname{tr} \frac{\mu^t T H}{\sigma^2}\right) (H dH^t) &= \frac{\Gamma_{K-n} \left[\frac{K-n}{2} \right]}{2^{K-n} \pi^{\frac{(K-n)^2}{2}}} \int_{H \in V_{n,K}} \int_{K \in O(K-n)} \exp\left[\operatorname{tr} \left(\frac{\mu^t T H}{\sigma^2} \right)\right] (K dK^t) (H dH^t), \\
 &= \frac{\Gamma_{K-n} \left[\frac{K-n}{2} \right]}{2^{K-n} \pi^{\frac{(K-n)^2}{2}}} \int_{H \in O(K)} \exp\left[\operatorname{tr} \left(\frac{\mu^t T}{\sigma^2} \right)\right] (H dH),
 \end{aligned} \tag{4.6}$$

y finalmente al normalizar la medida en el último término de la ecuación (4.6) se concluye que,

$$\int_{V_{n,K}} \exp\left[\operatorname{tr} \left(\frac{\mu^t T H}{\sigma^2} \right)\right] (H dH^t) = \frac{2^n \pi^{\frac{nK}{2}}}{\Gamma_n \left(\frac{K}{2} \right)} \int_{O(K)} \exp\left[\operatorname{tr} \left(\frac{\mu^t T H}{\sigma^2} \right)\right] (dH).$$

Así utilizando el teorema C.3.,

$$\begin{aligned}
 \int_{V_{n,K}} \exp\left[\operatorname{tr} \left(\frac{\mu^t T H}{\sigma^2} \right)\right] (H dH^t) &= \frac{2^n \pi^{\frac{nK}{2}}}{\Gamma_n \left(\frac{K}{2} \right)} \int_{O(K)} \exp\left[\operatorname{tr} \left(\frac{\mu^t T H}{\sigma^2} \right)\right] (dH), \\
 &= \frac{2^n \pi^{\frac{nK}{2}}}{\Gamma_n \left(\frac{K}{2} \right)} {}_0F_1 \left(\frac{1}{2} K, \frac{\mu^t T T^t \mu}{4\sigma^4} \right).
 \end{aligned}$$

Luego,

$$f_T(T) = \frac{\prod_{i=1}^n t_{ii}^{K-i}}{\sigma^M 2^{\frac{(M-2n)}{2}} \pi^{\frac{(M-nK)}{2}} \Gamma_n\left(\frac{K}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \text{tr}(TT^t + \mu\mu^t)\right] {}_0F_1\left(\frac{1}{2K}, \frac{\mu^t T T^t \mu}{4\sigma^4}\right)$$

■

Note que si $N - 1 \leq k$, T es no singular, $B = TT^t > 0$. Más aún, del teorema A.5., se calcula el producto exterior de B , siendo este igual a,

$$(dB) = 2^n \prod_{i=1}^n t_{ii}^{n+1-i} (dT), \quad \text{donde} \quad |B| = \prod_{i=1}^n t_{ii}^2.$$

se deduce que el producto exterior de dT es,

$$(dT) = 2^{-n} \prod_{i=1}^n t_{ii}^{-(n+1-i)} (dB),$$

donde $B = YY^t = (TH)(H^t T^t) = TT^t$ y

$$\prod_{i=1}^n t_{ii}^{K-i} \prod_{i=1}^n t_{ii}^{-(n+1-i)} = \prod_{i=1}^n t_{ii}^{K-n-1} = \prod_{i=1}^n t_{ii}^{\frac{2(K-n-1)}{2}} = |B|^{\frac{(K-n-1)}{2}}.$$

De aquí se concluye que,

$$f_B(B) = \frac{|B|^{\frac{(K-n-1)}{2}}}{\sigma^m 2^{\frac{m}{2}} \Gamma_n\left(\frac{K}{2}\right)} \exp\left(-\frac{1}{2\sigma^2} \text{tr}(B + \mu\mu^t)\right) {}_0F_1\left(\frac{K}{2}; \frac{\mu^t B \mu}{4\sigma^2}\right),$$

donde $N - 1 \leq K$, $n = N - 1$, y $\pi^{\frac{(M-nK)}{2}} = \pi^{\frac{[(N-1)K - (N-1)K]}{2}} = 1$.

■

Los elementos de la matriz T son llamadas coordenadas de tamaño-y-configuración. La factorización QR ya había sido utilizada anteriormente y los elementos de T también son llamadas coordenadas rectangulares, Rao (1973).

Distribución de la configuración

Teorema 4.2 . Con las condiciones del teorema 4.1. y expresando la matriz $Y = TH = rW(\mathbf{U})H$, la densidad de las coordenadas $\mathbf{W}(\mathbf{U})$ esta dada por,

$$f_{W(u)}W(u) = \frac{\prod_{i=1}^n w_{ii}^{K-i} J(u)}{\pi^{\left(\frac{M-nK}{2}\right)} \Gamma_n\left(\frac{K}{2}\right)} \exp\left[-\text{tr}\left(\frac{\mu^t \mu}{2\sigma^2}\right)\right] \sum_{j=0}^{\infty} \frac{2^{n-j-1} \Gamma\left(\frac{M}{2+j}\right)}{j!} \sum_{\lambda} \frac{C_{\lambda}(A)}{\left(\frac{K}{2}\right)_{\lambda}}, \quad (4.7)$$

donde $A = \frac{\mu^t T T^t \mu}{4\sigma^4}$ y $A^* = \frac{\mu^t W W^t \mu}{4\sigma^2} = \frac{\sigma^2 A}{r^2}$.

Demostración:

Expandiendo F_1 en polinomios zonales en la ecuación (4.1) se obtiene

$$f_T(T) = \frac{\prod_{i=1}^n t_{ii}^{K-i}}{\sigma^M 2^{\left(\frac{M-2n}{2}\right)}} \pi^{\left(\frac{M-nK}{2}\right)} \Gamma_n\left(\frac{K}{2}\right) \exp\left[-\frac{1}{\sigma^2} \text{tr}(T T^t + \mu^t \mu)\right] \sum_{l=0}^{\infty} \sum_{\lambda} \frac{C_{\lambda}\left(\frac{K}{2}; A\right)}{\left(\frac{K}{2}\right)_{\lambda} l!}, \quad (4.8)$$

donde $\left(\frac{K}{2}\right)_{\lambda} = \prod_{i=1}^l \left(\frac{K}{2} - \frac{i-1}{2}\right)_{K_i}$.

Sea $W(u) = \frac{T}{r}$, de donde $J(T \rightarrow r, u) = r^m J(u)$ y de las siguientes condiciones,

$$(i) \quad \prod_{i=1}^n t_{ii}^{K-i} = \prod_{i=1}^n (r w_{ii})^{K-i} = \prod_{i=1}^n r^{K-i} w^{K-i},$$

$$= r^{(K-1)+(K-2)+\dots+(K-n)} \prod_{i=1}^n w^{K-i} = r^{nK - \frac{n(n-1)}{2}} \prod_{i=1}^n w^{K-i}.$$

$$(ii) \quad r^{m+nK - \frac{n(n-1)}{2}} = r^{(N-1)K - nK + \frac{n(n+1)}{2} - 1 + nK - \frac{n(n-1)}{2}} = r^{(N-1)K-1} = r^{M-1}.$$

$$(iii) \quad \text{tr}(TT^t) = \text{tr}(rW)(rW^t) = r^2 \text{tr}(WW^t) = r^2 \|W\|^2 = r^2, \text{ ya que } \|W\| = 1.$$

$$(iv) \quad \frac{\mu^t T T^t \mu}{4\sigma^4} = \frac{\mu^t (rW)(rW^t) \mu}{4\sigma^4} = \frac{r^2 \mu^t W W^t \mu}{4\sigma^4}.$$

$$(v) \quad C_\lambda \left(\frac{r^2 \mu^t W W^t \mu}{4\sigma^4} \right) = \frac{r^{2l}}{\sigma^{2l} C_\lambda \left(\frac{\mu^t W W^t \mu}{4\sigma^4} \right)}.$$

Así se cumple lo siguiente,

$$f_{r,u}(r, u) = \frac{\prod_{i=1}^n W^{K-i} J(u) r^{M-1}}{\sigma^M 2^{\binom{M-2n}{2}} \pi^{\binom{M-nK}{2}} \Gamma_n \left(\frac{K}{2} \right)} \exp \left[-\frac{1}{2\sigma^2} \text{tr}(\mu \mu^t) \right] \exp \left[-\frac{r^2}{2\sigma^2} \right] \sum_{l=0}^{\infty} \sum_{\lambda} \frac{C_\lambda \left(\frac{K}{2}; A \right)}{\left(\frac{K}{2} \right)_\lambda l!}.$$

Integrando con respecto a r se tiene; [agrupando las potencias de r , σ y descomponiendo las potencias de (4.8)],

$$f_{W(u)}W(u) = \frac{\prod_{i=1}^n W^{K-i} J(u) 2^n}{\pi^{\left(\frac{M-nK}{2}\right)} \prod \Gamma_n\left(\frac{K}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \text{tr}(\mu\mu^t)\right] \sum_{l=0}^{\infty} \left[\int_0^{\infty} \frac{r^{M+2l-1}}{2^{\left(\frac{M}{2}\right)} \sigma^{M+2l}} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr \right] \sum_{\lambda} \frac{C_{\lambda}\left(\frac{K}{2}; A\right)}{\left(\frac{K}{2}\right)_{\lambda} l!}.$$

Al analizar la integral,

$$I = \int_0^{\infty} \frac{r^{M+2l-1}}{2^{\left(\frac{M}{2}\right)} \sigma^{M+2l}} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr.$$

Dado que $\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx$ y definiendo $x = \frac{r^2}{2\sigma^2}$, de donde $dx = \frac{2r}{2r^2} y$ $dr = \frac{r dr}{\sigma^2}$, de manera tal que,

$$\Gamma(t) = \int_0^{\infty} \left(\frac{r^2}{2\sigma^2}\right)^{t-1} \exp\left(-\frac{r^2}{2\sigma^2}\right) \frac{r}{\sigma^2} dr = \int_0^{\infty} \frac{r^{2(t-1)+1}}{2^{t-1} \sigma^{2(t-1)+2}} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr.$$

Ahora sea $t = \frac{M}{2} + l$, luego

$$\Gamma\left(\frac{M}{2} + l\right) = \int_0^{\infty} \frac{r^{M+2l-1}}{2^{\left(\frac{M}{2}+l-1\right)} \sigma^{M+2l}} \exp\left(-\frac{r^2}{2\sigma^2}\right) dr,$$

de aquí,

$$I = 2^{l-1} \Gamma\left(\frac{M}{2} + l\right).$$

Finalmente,

$$f_{W(u)}W(u) = \frac{\prod_{i=1}^n W^{K-i} J(u)}{\pi^{\left(\frac{M-nK}{2}\right)} \Gamma_n\left(\frac{K}{2}\right)} \exp\left[-\frac{1}{2\sigma^2} \text{tr}(\mu\mu^t)\right] \sum_{l=0}^{\infty} \frac{2^{n+l-1} \Gamma\left(\frac{M}{2} + l\right)}{l!} \sum_{\lambda} \frac{C_{\lambda}\left(\frac{\mu^t \Gamma \Gamma^t \mu}{4\sigma^2}\right)}{\left(\frac{K}{2}\right)_{\lambda}}.$$



Distribuciones excluyendo reflexión

Las densidades presentadas en los teoremas 4.1. y 4.2. incluyen reflexión. Con el objeto de encontrar las correspondientes distribuciones excluyendo reflexión, observe que (vea Goodall y Mardia (1993))

$$\int_{O^+(K)} \exp[\text{tr } \mu^t T H] (H dH^t) = \frac{1}{2} \int_{O(K)} \exp[\text{tr } \mu^t T H] (H dH^t).$$

Así se obtienen los siguientes corolarios a los teoremas 4.1. y 4.2. respectivamente.

Corolario 4.1 . Cuando $N - 1 \geq K$ y el rango de $\mu < K$ la densidad de tamaño-y-configuración para $T = T^{NR}$ excluyendo reflexión está dada por (4.1) dividida por dos, donde $t_{ii} \geq 0$ para todo $i = 1, \dots, K - 1$, y t_{KK} no está restringida. Cuando $N - 1 < K$, la densidad está dada por (4.1) y t_{KK} no está presente.

Corolario 4.2 . Cuando $N - 1 \geq K$ y el rango de $\mu < K$ la densidad de configuración para $W = W^{NR}$ excluyendo reflexión está dada por (4.7) dividida por dos, donde w_{KK} no está restringida. Cuando $N - 1 < K$, la densidad de configuración está dada por (4.7) y W_{KK} no está presente.

Distribución de tamaño-y-configuración mediante la factorización polar

En la presente sección , se encuentra la distribución de R , que por analogía al caso QR , se llama distribución de tamaño-y-configuración bajo la factorización polar. Observe que bajo este nuevo enfoque la distribución de rotación-tamaño-y-configuración coincide con la encontrada bajo la técnica que considera la factorización QR , más aún, tal distribución está dada por (1.2).

Teorema 4.3 . La densidad de R está dada por:

$$f_R(R) = \frac{|R|^{(r-s)} \prod_{i=1}^s (D_{ii} - D_{jj})}{2^{\lfloor \frac{s(r-2)}{2} \rfloor} \sigma^{rs}} \exp \left[-\frac{1}{2} \text{tr}(R^2 + \mu^t \mu) \right] {}_0F_1 \left(\frac{s}{2}; \frac{\mu^t R^2 \mu}{4\sigma^4} \right).$$

Demostración:

$$dF_X(X) = \frac{1}{(2\pi)^{\frac{rs}{2}} \sigma} \exp \left[-\frac{1}{2\sigma^2} \text{tr}(X - \mu)^t (X - \mu) \right] (dX),$$

desarrollando el argumento de la exponencial,

$$\text{tr}(X - \mu)(X - \mu)^t = \text{tr}(XX^t - 2\mu^t X + \mu\mu^t).$$

Factorizando ahora $X = RH_1$ con $(dX) = |R|^{r-s} \prod_{i=1}^q (D_{ii} - D_{jj})(dR)(H_1 dH_1^t)$,

se obtiene que,

$$dF_{H_1, R}(H_1, R) = \frac{|R|^{(r-s)}}{(2\pi)^{\binom{s}{2}} \sigma^{rs}} \exp \left\{ -\frac{1}{2\sigma^2} \operatorname{tr} \left[(R^2 + \mu^t \mu) - 2\mu^t H_1 R \right] \right\} \prod_{i=1}^s (D_{ii} - D_{jj}) (dR) (H_1 dH_1^t),$$

para obtener la marginal de R observe que

$$\int_{V_{s,r}} \exp \left[-\frac{1}{\sigma^2} \operatorname{tr} (R \mu^t H_1) \right] (H_1 dH_1^t) = \frac{2^s \pi^{\frac{rs}{2}}}{\Gamma_s \left(\frac{r}{2} \right)} {}_0F_1 \left(\frac{r}{2}, \frac{\mu^t R^2 \mu}{4\sigma^4} \right),$$

de donde finalmente se obtiene el resultado. ■

LITERATURA CITADA

- Aykroyd R. G., Bowie J. L. and Mardia K. V. 1997. Statistical shape matching for facial photographs. <http://fourier.dur.ac.uk:8000/statsl/pgl/leeds97>. London.
- Bookstein, F. L. 1986. *Size and shape spaces for landmark in two dimensions*. Statist. Sci. 1, p. 181-242, USA.
- Brown, T. W. , Foote K. and Holz R. K. 1997. City Shape and Growth In Austin, Texas. <http://www.utexas.edu/depts/grg/ustudent/geraft/fall96/marsh/City/Shape/Shape.html>. USA.
- Cadet, G. H. 1996. *Polar Coordinates in \mathbb{R}^{np} : Application to the computation of Wishart and Beta laws*, Sankhya A, 58, p. 101-114, India.
- Cootes T. F. and Taylor C. J. 1996. Image analysis and computer vision. <http://www.amsta.leeds.ac.uk/Statistics/leedcove.html>. London.
- Dryden, I. L. and Mardia, K. V. 1991. *General shape distributions in a plane*. Adv. Appl. Prob., 23, p. 259-276, USA.
- Golub G. H. and Van Loan C. F. 1993. Matrix computations (Segunda edición). Baltimore and London: The Johns Hopkins Press. p. 193-259.
- Goodall, C. R. 1991. *Procrustes methods in the statistical analysis of shape*. Journal of the Royal Statistical Society B, 53, p. 285-339. London.
- Goodall, C. R. and Mardia, K. V. 1991. *Multivariate analysis of shape theory*. Technical Report 91-20, Center for Multivariate Analysis. Pennsylvania State Univ. USA.

and shape densities. *J. Multivariate Anal.* 40, p. 94-108, USA.

1993. *Multivariate aspects of shape theory*. The Annals of Statistics, Vol 21. No. 2, p. 848-866, USA.

Gower, J. C. 1975. *Generalized procrustes analysis*. *Psychometrika*, 40, p. 33-50, USA.

Hoffman K. and Kunze R. 1984. *Algebra Lineal*. (Primera edición). Prentice Hall. México. 154p.

James, A. T. 1954. *Normal multivariate analysis and the orthogonal group*. *Ann. Math. Statist.*, 25, p. 40-75, USA.

Kendall, D. G. 1984. *Shape manifolds, Procrustean metrics and complex projective spaces*. *Bull. Lond. Math Soc.*, 16, p. 81-121. London.

Kent, J. T. 1994. *The complex Bingham distribution and shape analysis*. *Journal of the Royal Statistical Society B*, 56, p. 285-299. London.

Le, H. y Kendall, D. G. *The Riemannian structure of euclidian shape space: A novel enviroment for statistics*. *Ann. Statist.*, 21, p. 1225-1271, USA.

Lewis, Kendall D. G. and Thomson D. A. 1992. *The statistics of shape*. <http://setosa.uwaterloo.ca/cgsmall/shape.html>. Canada.

Meyer P.L. 1992. *Probabilidad y aplicaciones estadísticas* (Segunda edición). Addison-Wesley Iberoamericana. México. 254p.

Mood A. M., Graybill F. A. and Boes D.C. 1974. *Introduction to the theory of statistics* (third edition). International Student Edition: M.C. Graw-Hill International Book Company. USA. 205p.

Muirhead R. J. 1982. *Aspects of multivariate statistical theory*. John Wiley & Sons. New York. p. 50-77.

Rao, C. R. 1973. *Lineal Statistical Inference And Its Applications*. John Wiley & Sons. New York, p. 597-599.

- Small, C. G. 1981. *Distributions of shape and maximal invariant statistics*. Ph. D. dissertation, Univ. Cambridge. England.
- Takemura A. 1984. Zonal polynomials, Volume 4. Hayward, California: Institute of Mathematical Statistics. Lecture Notes - Monograph Series, Series Editor, Shanti S. Gupta, Purdue University. USA. p. 7-41.
- Wentz E. A. 1994. GIS and shape analysis. <http://gis.psu.edu/Wentz/WentzHTML/Shapes.html>. USA.

APÉNDICES

APÉNDICE A

PRODUCTO EXTERIOR

PRODUCTO EXTERIOR

Al aplicar una transformación a una variable aleatoria, ya sea univariada o multivariada, es de interés llegar a conocer la densidad asociada a la nueva variable. Para determinarla es necesario calcular algunos jacobianos, tarea que en muchas de las ocasiones no es fácil de realizar, motivo por el cual se recurre al uso de técnicas alternativas. Una de estas técnicas consiste en el uso del producto exterior. El siguiente teorema permite determinar la función de densidad asociada a la transformación.

Teorema A.1 . Si X es un vector aleatorio de dimensión \mathbb{R}^r con función de densidad $f_X(x)$ definida en el conjunto $S \subset \mathbb{R}^r$. Suponga que $y(X) = (y_1(X), y_2(X), \dots, y_r(X))^t$ es una transformación 1-1 y sobre de S en T . En donde T denota la imagen de S bajo y , entonces la transformación inversa $x = x(y)$ existe y es continua en T , la función de densidad del vector $Y = y(X)$ está dada por

$$g_Y(y) = f_X(x(y))|J(x \rightarrow y)| \quad (y \in T),$$

donde $J(X \rightarrow Y)$ es el jacobiano de la transformación de X en Y y éste está dado por

$$J(X \rightarrow Y) = \det \begin{bmatrix} \frac{\partial x_1}{\partial y_1} & \cdots & \frac{\partial x_1}{\partial y_r} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_r}{\partial y_1} & \cdots & \frac{\partial x_r}{\partial y_r} \end{bmatrix} = \det \left(\frac{\partial x_i}{\partial y_j} \right).$$

Cuando se trabaja con varias variables es difícil calcular el determinante J . El método alternativo haciendo uso del producto exterior se basa en la forma simétrica y anticonmutativa de la multiplicación de diferenciales. Este desarrollo fue propuesto por James (1954).

$$I = \int_A f(x_1, \dots, x_r) dx_1 \cdots dx_r \text{ donde } A \subset \mathbb{R}^r. \quad (\text{A.1})$$

Si $f_X(x)$ es una función de densidad con $X \in \mathbb{R}^r$, la integral representa la probabilidad de que el vector X tome valores en el conjunto A y al realizar el cambio de variables,

$$\begin{aligned} x_1 &= x_1(y_1, \dots, y_r), \\ &\vdots \\ x_r &= x_r(y_1, \dots, y_r), \end{aligned}$$

y sustituyendo en (A.1) se tiene que,

$$I = \int_{A^t} f_X(x(y)) \det \left(\frac{\partial x_i}{\partial y_j} \right) dy_1 \cdots dy_r,$$

donde A^t denota la imagen de A bajo la función Y .

A continuación se ejemplifica la forma de calcular el jacobiano de la transformación para el caso particular donde $r = 2$, la transformación del cambio de variables esta dada por,

$$x_1 = x_1(y_1, y_2),$$

$$x_2 = x_2(y_1, y_2),$$

y su jacobiano por,

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}.$$

Al desarrollar las operaciones que definen al jacobiano, éste es igual a,

$$J = \left(\frac{\partial x_1}{\partial y_1} \right) \left(\frac{\partial x_2}{\partial y_2} \right) - \left(\frac{\partial x_1}{\partial y_2} \right) \left(\frac{\partial x_2}{\partial y_1} \right). \quad (\text{A.2})$$

Ahora, al calcular el mismo resultado usando las formas diferenciales con el fin de motivar la definición del producto exterior, se tiene:

El diferencial total de una función $x_i = x_i(y_1, \dots, y_r)$ se define por,

$$dx_i = \frac{\partial x_i}{\partial y_1} dy_1 + \dots + \frac{\partial x_i}{\partial y_r} dy_r \quad (i = 1, \dots, r).$$

Al sustituir estas formas diferenciales lineales en la integral I de (A.1) y al considerar el caso particular de $r = 2$, se obtiene el siguiente resultado,

$$I = \int_{A^t} f(x(y)) \left(\frac{\partial x_1}{\partial y_1} dy_1 + \frac{\partial x_1}{\partial y_2} dy_2 \right) \left(\frac{\partial x_2}{\partial y_1} dy_1 + \frac{\partial x_2}{\partial y_2} dy_2 \right). \quad (\text{A.3})$$

Suponiendo que es válido multiplicar los diferenciales con las leyes distributivas, al operar con los diferenciales que aparecen dentro de la integral en (A.3), da como resultado,

$$\frac{\partial x_1}{\partial y_1} \frac{\partial x_2}{\partial y_1} dy_1 dy_1 + \frac{\partial x_1}{\partial y_1} \frac{\partial x_2}{\partial y_2} dy_1 dy_2 + \frac{\partial x_1}{\partial y_2} \frac{\partial x_2}{\partial y_1} dy_2 dy_1 + \frac{\partial x_1}{\partial y_2} \frac{\partial x_2}{\partial y_2} dy_2 dy_2. \quad (\text{A.4})$$

Para que las ecuaciones (A.2) y (A.4) sean iguales se necesita que,

$$dy_1 dy_2 = -dy_2 dy_1.$$

Es decir el producto de dos diferenciales es anti-simétrico, $dy_i dy_j = -dy_j dy_i$, en particular $dy_i dy_i = -dy_i dy_i = 0$, tal producto es llamado el producto exterior y se denota con el símbolo \wedge (en inglés este símbolo es llamado wedge).

Haciendo uso de esta notación y la propiedad anti-simétrica, este producto se denota de la siguiente forma,

$$dy_i \wedge dy_j = -dy_j \wedge dy_i.$$

De manera tal que el jacobiano de esta transformación se expresa como,

$$J = \frac{\partial x_1}{\partial y_1} \frac{\partial x_2}{\partial y_2} - \frac{\partial x_1}{\partial y_2} \frac{\partial x_2}{\partial y_1} = dy_1 \wedge dy_2.$$

Esta forma de trabajo induce a pensar que el multiplicar diferenciales es equivalente a calcular el jacobiano, el siguiente teorema muestra que tal suposición es cierta.

Teorema A.2 . Si dy es un vector en \mathbb{R}^r de formas diferenciales y $dx = Bdy$, donde B es una matriz no singular de dimensión $\mathbb{R}^{r \times r}$, es decir dx es un vector de formas diferenciales lineales, entonces

$$\bigwedge_{i=1}^r dx_i = \det B \bigwedge_{i=1}^r dy_i.$$

Se realizara la demostración de este teorema para el caso particular de $r = 3$, por medio de la definición del determinante.

Demostración :

El lado derecho de la ecuación $dX = BdY$, se puede escribir como

$$\bigwedge_{i=1}^r dx_i = p(B) \bigwedge_{i=1}^r dy_i. \quad (\text{A.5})$$

donde $p(B)$ es un polinomio con elementos en B . Para el caso de $r = 3$, se tiene que

$$\begin{pmatrix} dx_1 \\ dx_2 \\ dx_3 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \begin{pmatrix} dy_1 \\ dy_2 \\ dy_3 \end{pmatrix}.$$

Ahora desarrollando el producto de la derecha para cada diferencial dx_i donde $i = 1, 2$ y 3 , se llega a,

$$dx_1 = b_{11}dy_1 + b_{12}dy_2 + b_{13}dy_3,$$

$$dx_2 = b_{21}dy_1 + b_{22}dy_2 + b_{23}dy_3,$$

$$dx_3 = b_{31}dy_1 + b_{32}dy_2 + b_{33}dy_3,$$

de aquí,

$$\begin{aligned} dx_1 \wedge dx_2 \wedge dx_3 &= (b_{11}b_{22}b_{33} - b_{12}b_{21}b_{33} - b_{11}b_{23}b_{32} + b_{13}b_{21}b_{32} \\ &\quad + b_{12}b_{23}b_{31} + b_{13}b_{22}b_{31})(dy_1 \wedge dy_2 \wedge dy_3). \end{aligned}$$

En general se cumplen las siguientes tres propiedades,

- (i) $p(B)$ es lineal en cada columna de B .
- (ii) Si el orden de dos factores dx_i, dx_j se invierten, entonces el signo de $\bigwedge_{i=1}^r dx_i$ cambia. Pero esto es equivalente a intercambiar la i -ésima y la j -ésima columnas de B . Es decir al intercambiar dos columnas de B , el signo de $p(B)$ es contrario.
- (iii) $p(I_r) = 1$.

Las propiedades (i), (ii) y (iii) caracterizan a la función determinante de la definición dada por Weierstrass. Por lo tanto $p(B) = \det B$.

■

Así al sustituir este resultado en la ecuación (A.5) y al considerar la matriz B como la matriz de derivadas parciales que aparece en la ecuación $dX = BdY$, la integral I se escribe como

$$I = \int_A f(x_1, \dots, x_r) dx_1 \wedge \dots \wedge dx_r \text{ donde } A \subset \mathbb{R}^r.$$

De manera tal que al determinar la distribución asociada a la transformación, en lugar de calcular el jacobiano se calcula el producto de los diferenciales asociados.

A continuación se dan algunos resultados del producto exterior para vectores y matrices. El siguiente teorema ejemplifica la forma de calcular el jacobiano para la transformación de coordenadas rectangulares a coordenadas polares generalizadas.

Teorema A.3 . (*Coordenadas polares generalizadas.*) Cuando se definen las componentes x_1, \dots, x_s en coordenadas rectangulares, si se requiere hacer un cambio a coordenadas polares $r, \theta_1, \dots, \theta_{s-1}$, cada una de las componentes x_i para $i = 1, \dots, s$ se expresan por medio de:

$$x_1 = r \operatorname{sen} \theta_1 \operatorname{sen} \theta_2 \dots \operatorname{sen} \theta_{s-2} \operatorname{sen} \theta_{s-1}.$$

$$x_2 = r \operatorname{sen} \theta_1 \operatorname{sen} \theta_2 \dots \operatorname{sen} \theta_{s-2} \cos \theta_{s-1}.$$

$$x_3 = r \operatorname{sen} \theta_1 \operatorname{sen} \theta_2 \dots \cos \theta_{s-2}.$$

$$\vdots$$

$$x_{s-1} = r \operatorname{sen} \theta_1 \cos \theta_2.$$

$$x_s = r \cos \theta_1.$$

donde $r > 0$, $0 < \theta_i < \pi$ ($i = 1, 2, \dots, s-2$), $0 < \theta_{s-1} \leq 2\pi$, entonces se tiene que

$$\bigwedge_{i=1}^s dx_i = r^{s-1} \operatorname{sen}^{s-2} \theta_1 \operatorname{sen}^{s-3} \theta_2 \dots \operatorname{sen} \theta_{s-2} \left(\bigwedge_{i=1}^{s-1} d\theta_i \right) \wedge dr,$$

y por lo tanto el jacobiano de la transformación es,

$$J(x \rightarrow r, \theta_1, \dots, \theta_{s-1}) = r^{s-1} \operatorname{sen}^{s-2} \theta_1 \operatorname{sen}^{s-3} \theta_2 \dots \operatorname{sen} \theta_{s-2}.$$

Demostración :

Note que:

$$\begin{aligned}
x_1^2 &= r^2 \sin^2 \theta_1 \sin^2 \theta_2 \dots \sin^2 \theta_{s-2} \sin^2 \theta_{s-1}. \\
x_1^2 + x_2^2 &= r^2 \sin^2 \theta_1 \sin^2 \theta_2 \dots \sin^2 \theta_{s-2}. \\
&\vdots \\
x_1^2 + \dots + x_s^2 &= r^2.
\end{aligned}$$

Diferenciando los primeros términos de estas ecuaciones se obtiene que,

$$\begin{aligned}
2x_1 dx_1 &= 2r^2 \sin^2 \theta_1 \dots \sin^2 \theta_{s-2} \sin \theta_{s-1} \cos \theta_{s-1} d\theta_{s-1}. \\
&+ \text{términos que incluyen } dr, d\theta_1, \dots, d\theta_{s-2}.
\end{aligned}$$

$$\begin{aligned}
2x_1 dx_1 + 2x_2 dx_2 &= 2r^2 \sin^2 \theta_1 \dots \sin \theta_{s-2} \cos \theta_{s-2} d\theta_{s-2}. \quad (\text{A.6}) \\
&+ \text{términos que incluyen } dr, d\theta_1, \dots, d\theta_{s-2}.
\end{aligned}$$

\vdots

$$2x_1 dx_1 + \dots + 2x_s dx_s = 2r dr.$$

Al realizar el producto exterior del lado izquierdo de las ecuaciones (A.6) y recordando que el producto exterior de términos repetidos es igual a cero, se genera la siguiente expresión

$$2^s x_1 \dots x_s \bigwedge_{i=1}^s dx_i. \quad (\text{A.7})$$

Al calcular el producto exterior del lado derecho de (A.6) este es igual a

$$2^s r^{2s-1} \sin^{2s-3} \theta_1 \sin^{2s-5} \theta_2 \dots \sin \theta_{s-1} \cos \theta_1 \cos \theta_2 \dots \cos \theta_{s-1} \bigwedge_{i=1}^{s-1} d\theta_i \wedge dr.$$

y de la definición de las x_i , este da,

$$2^s x_1 \cdots x_s r^{s-1} \sin^{s-2} \theta_1 \sin^{s-3} \theta_2 \cdots \sin \theta_{s-2} \bigwedge_{i=1}^{s-1} d\theta_i \wedge dr, \quad (\text{A.8})$$

de (A.7) y (A.8) se deduce el resultado deseado. ■

Enseguida se enuncian algunas definiciones del producto exterior dada una matriz de diferenciales y se exponen algunas propiedades y resultados de este producto.

Definición A.1 . (*Matriz de diferenciales*) Dada una matriz X , la matriz de sus diferenciales se define como el diferencial de cada una de sus componentes, se denota por dX ó por (dx_{ij}) .

Proposición A.1 . (*Producto*) Si $X \in \mathbb{R}^{r \times s}$ y $Y \in \mathbb{R}^{s \times p}$ son matrices, entonces la matriz de diferenciales del producto de X con Y se denota por $d(XY)$ y se cumple la siguiente propiedad

$$d(XY) = XdY + dXY.$$

La demostración se realiza utilizando la definición del producto de matrices.

Definición A.2 . Para una matriz X en $\mathbb{R}^{r \times s}$ el producto exterior de sus diferenciales se denota por (dX) y este es igual a,

$$(dX) = \bigwedge_{i=1}^r \bigwedge_{j=1}^s dx_{ij}.$$

Observación A.1 . Si $X \in \mathbb{R}^{r \times r}$ es una matriz simétrica, se tiene que $dx_{ij} = dx_{ji}$, es decir aparecen elementos repetidos, de manera tal que el total de elementos diferentes en la matriz X es igual a $\frac{r(r+1)}{2}$ y el producto exterior de sus elementos esta dado por,

$$(dX) = \bigwedge_{1 \leq i \leq j \leq r} dx_{ij}.$$

Si $X \in \mathbb{R}^{r \times r}$ es una matriz anti-simétrica ($X = -X^t$), entonces en el producto existen $\frac{r(r-1)}{2}$ elementos diferentes y se reduce a multiplicar los elementos que aparecen por arriba de la diagonal superior.

$$(dX) = \bigwedge_{i < j} dx_{ij}.$$

Teorema A.4 . (*Transformación lineal*) Si $X = BY$ donde X y Y son matrices en $\mathbb{R}^{r \times s}$ y B es una matriz no singular en $\mathbb{R}^{r \times r}$, entonces

$$(dX) = (\det B)^s (dY).$$

de manera tal que $J(X \rightarrow Y) = (\det B)^s$.

Demostración :

Dado que $X = BY$, entonces de la proposición A.1. se tiene que $dX = dBY + BdY = 0Y + BdY$, de manera tal que $dX = BdY$, donde $dX = [dx_1, \dots, dx_s]$ y $dY = [dy_1, \dots, dy_s]$.

Ahora aplicando el teorema A.2. para cada $dx_j = Bdy_j$ ($j = 1, 2, \dots, s$), el producto exterior de los elementos de X es,

$$\bigwedge_{i=1}^r dx_{ij} = (\det B) \bigwedge_{i=1}^r dy_{ij}.$$

Por lo tanto,

$$\begin{aligned} (dX) &= \bigwedge_{j=1}^s \bigwedge_{i=1}^r dx_{ij}, \\ &= \bigwedge_{j=1}^s \left[(\det B) \bigwedge_{i=1}^r dy_{ij} \right], \\ &= (\det B)^s \bigwedge_{j=1}^s \left[\bigwedge_{i=1}^r dy_{ij} \right], \\ &= (\det B)^s \bigwedge_{j=1}^s \bigwedge_{i=1}^r dy_{ij}, \\ &= (\det B)^s (dY). \end{aligned}$$

■

Lema A.1 . Si $X = BYC$, donde X y Y son matrices en $\mathbb{R}^{r \times s}$, B y C son matrices no singulares en $\mathbb{R}^{r \times r}$ y $\mathbb{R}^{s \times s}$ respectivamente, entonces

$$(dX) = (\det B)^s (\det C)^r (dY),$$

donde el jacobiano de la transformación se expresa como,

$$J(X \rightarrow Y) = (\det B)^s (\det C)^r.$$

Lema A.2 . Sea $X = BYB^t$, donde X y Y son matrices simétricas en $\mathbb{R}^{s \times s}$ y suponga que B es una matriz no singular en $\mathbb{R}^{s \times s}$, entonces

$$(dX) = (\det B)^{s+1} (dY).$$

Lema A.3 . Suponga $X = BYB^t$ donde X y Y son matrices anti-simétricas en $\mathbb{R}^{s \times s}$ y B es una matriz no singular $\mathbb{R}^{s \times s}$, entonces

$$(dX) = (\det B)^{s-1} (dY).$$

Lema A.4 . Si $X = Y^{-1}$, donde Y es una matriz simétrica en $\mathbb{R}^{s \times s}$, entonces

$$(dX) = (\det Y)^{-(s+1)} (dY).$$

Teorema A.5 . Si A es una matriz definida positiva en $\mathbb{R}^{s \times s}$ y $A = T^t T$, donde T es una matriz triangular superior con elementos en la diagonal positivos, entonces

$$(dA) = 2^s \prod_{i=1}^s t_{ii}^{s+1-i} (dT).$$

Demostración :

Con las condiciones del teorema se tiene que,

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1s} \\ a_{12} & a_{22} & \dots & a_{2s} \\ \vdots & & \ddots & \\ a_{1s} & a_{2s} & \dots & a_{ss} \end{bmatrix} = \begin{bmatrix} t_{11} & 0 & \dots & 0 \\ t_{12} & t_{22} & \dots & 0 \\ \vdots & & \ddots & \\ t_{1s} & t_{2s} & \dots & t_{ss} \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1s} \\ 0 & t_{22} & \dots & t_{2s} \\ \vdots & & \ddots & \\ 0 & 0 & \dots & t_{ss} \end{bmatrix}$$

Desarrollando el producto de las matrices que aparecen del lado derecho de la ecuación anterior y expresándolo en función de cada miembro de la matriz de la izquierda, se concluye que,

$$\begin{aligned} a_{11} &= t_{11}^2, \\ a_{12} &= t_{11}t_{12}, \\ &\vdots \\ a_{1s} &= t_{11}t_{1s}, \\ a_{22} &= t_{12}^2 + t_{22}^2, \\ &\vdots \\ a_{2s} &= t_{12}t_{1s} + t_{22}t_{2s}, \\ &\vdots \\ a_{ss} &= t_{1s}^2 + \dots + t_{ss}^2, \end{aligned}$$

de aquí y calculando el diferencial para cada a_{ij} , siendo este igual a,

$$\begin{aligned}
da_{11} &= 2t_{11}dt_{11}. \\
da_{12} &= t_{11}dt_{12} + \cdots \\
&\vdots \\
da_{1s} &= t_{11}dt_{1s} + \cdots \\
da_{22} &= 2t_{22}dt_{22} + \cdots \\
&\vdots \\
da_{2s} &= t_{22}dt_{2s} + \cdots \\
&\vdots \\
da_{ss} &= 2t_{ss}dt_{ss} + \cdots
\end{aligned}$$

Por lo tanto el producto exterior sobre los a_{ij} es,

$$(dA) = \bigwedge_{i \leq j}^s da_{ij} = 2^s t_{11}^{s-1} t_{22}^{s-2} \cdots t_{ss} \bigwedge_{i \leq j}^s dt_{ij}.$$

■

Corolario A.1 . Si A es una matriz definida positiva en $\mathbb{R}^{s \times s}$ y $A = GG^t$, donde G es una matriz triangular inferior con elementos en la diagonal positivos, entonces

$$(dA) = 2^s \prod_{i=1}^s g_{ii}^{s+1-i} (dG).$$

Demostración: Sólo observe que $G = T^t$.

■

El siguiente lema muestra la forma de calcular el producto exterior en la factorización de una matriz de rango m , resultado que será de mucha utilidad al aplicarlo en la factorización QR .

Lema A.5 . Si Z es una matriz en $\mathbb{R}^{r \times s}$, ($r \geq s$) de rango s y que se puede expresar como $Z = H_1 T$, donde H_1 es una matriz en $\mathbb{R}^{r \times s}$, donde $H_1^t H_1 = I_s$ y T es una matriz triangular superior en $\mathbb{R}^{s \times s}$ con elementos positivos en la diagonal. Si H_2 (como una función de H_1) es una matriz en $\mathbb{R}^{r \times (r-s)}$, tal que $H = [H_1 : H_2]$ es una matriz ortogonal en $\mathbb{R}^{r \times r}$ y se puede escribir como la matriz $H = [h_1 \dots h_s : h_{s+1} \dots h_r]$, donde $[h_1 \dots h_s]$ son las columnas de H_1 y $h_{s+1} \dots h_r$ son las columnas de H_2 . Entonces

$$(dZ) = \prod_{i=1}^s t_{ii}^{r-i} (dT)(H_1^t dH_1),$$

donde,

$$(H_1^t dH_1) \equiv \bigwedge_{i=1}^s \bigwedge_{j=i+1}^r h_j^t dh_i.$$

Demostración:

Como $Z = H_1 T$ y de la proposición A.1. se tiene que $dZ = dH_1 T + H_1 dT$.

Ahora

$$\begin{aligned}
H^t dZ &= \begin{bmatrix} H_1^t \\ H_2^t \end{bmatrix} dZ, \\
&= \begin{bmatrix} H_1^t \\ H_2^t \end{bmatrix} [dH_1 T + H_1 dT], \\
&= \begin{bmatrix} H_1^t dH_1 T + H_1^t H_1 dT \\ H_2^t dH_1 T + H_2^t H_1 dT \end{bmatrix}, \\
&= \begin{bmatrix} H_1^t dH_1 T + dT \\ H_2^t dH_1 T \end{bmatrix},
\end{aligned} \tag{A.9}$$

dado que $H_1^t H_1 = I_s$ y $H_2^t H_1 = 0$.

Ahora al calcular el producto exterior del término $H^t dZ$ del teorema A.4. este es igual a,

$$(H^t dZ) = (\det H^t)^s (dZ) = (dZ),$$

ya que $\det H^t = \det H = 1$, por ser H una matriz ortogonal.

El j -ésimo renglón de la matriz $H_2^t dH_1 T$, está dado por el primer renglón de,

$$(h_j^t dh_1, \dots, h_j^t dh_s) T, \quad (s+1 \leq j \leq r).$$

Aplicando el teorema A.2. se sigue que el producto exterior de los elementos de cada renglón es,

$$(\det T) \bigwedge_{i=1}^s h_j^t dh_i.$$

Entonces el producto exterior de todos los elementos en $H_1^t dH_1 T$ se expresa como,

$$\bigwedge_{j=s+1}^r \left[(\det T) \bigwedge_{i=1}^s h_j^t dh_i \right] = (\det T)^{r-s} \bigwedge_{j=s+1}^r \bigwedge_{i=1}^s h_j^t dh_i. \quad (\text{A.10})$$

Ahora al considerar el término $H_1^t dH_1 T + dT$ y dado que $H_1^t H_1 = I_s$, se cumple que,

$$H_1^t dH_1 + dH_1^t H_1 = 0,$$

se sigue que,

$$H_1^t dH_1 = -dH_1^t H_1 = -(H_1^t dH_1)^t.$$

Por lo tanto $H_1^t dH_1$ es una matriz anti-simétrica que se puede escribir como,

$$H_1^t dH_1 = \begin{bmatrix} 0 & -h_2^t dh_1 & \dots & -h_s^t dh_1 \\ h_2^t dh_1 & 0 & \dots & -h_s^t dh_2 \\ h_3^t dh_1 & h_3^t dh_2 & \dots & -h_s^t dh_3 \\ \vdots & \vdots & \ddots & \vdots \\ h_s^t dh_1 & h_s^t dh_2 & & 0 \end{bmatrix}, \quad \in \mathbb{R}^{s \times s}.$$

Ahora, multiplicando la matriz anterior por la matriz triangular superior T , donde se denota a los términos que se repiten en los diferenciales con el símbolo $*$,

$$H_1^t dH_1 T = \begin{bmatrix} 0 & * & \dots & * & * \\ h_2^t dh_1 t_{11} & * & \dots & * & * \\ h_3^t dh_1 t_{11} & h_3^t dh_2 t_{22} + * & \dots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ h_s^t dh_1 t_{11} & h_s^t dh_2 t_{22} + * & \dots & h_s^t dh_{s-1} t_{s-1,s-1} + * & * \end{bmatrix}. \quad (\text{A.11})$$

El producto exterior columna por columna de la ecuación (A.11) da como resultado

$$\begin{aligned} & t_{11}^{s-1} \bigwedge_{j=2}^s h_j^t dh_1 \bigwedge t_{22}^{s-2} \bigwedge_{j=3}^s h_j^t dh_2 \dots \bigwedge t_{s-1,s-1} h_s^t dh_{s-1} \\ & = t_{11}^{s-1} t_{22}^{s-2} \dots t_{s-1,s-1} \bigwedge_{i=1}^s \bigwedge_{j=i+1}^s h_s^t dh_{s-1}. \end{aligned} \quad (\text{A.12})$$

De (A.10) y (A.12) se tiene que el producto exterior de los elementos de $H_2^t dH_1 T$ y los elementos de la subdiagonal de $H_1^t dH_1 T + dT$ es,

$$\begin{aligned} & \left(\prod_{i=1}^s t_{ii}^{r-s} \right) \left(\bigwedge_{i=1}^s \bigwedge_{j=s+1}^r h_j^t dh_i \right) \left(\prod_{i=1}^s t_{ii}^{s-i} \right) \left(\bigwedge_{i=1}^s \bigwedge_{j=i+1}^s h_j^t dh_i \right) \\ & = \prod_{i=1}^s t_{ii}^{r-i} \bigwedge_{i=1}^s \bigwedge_{j=i+1}^r h_j^t dh_i = \prod_{i=1}^s t_{ii}^{r-i} (H_1^t dH_1). \end{aligned} \quad (\text{A.13})$$

El producto exterior de los elementos de $H_1^t dH_1 T + dT$ en su diagonal superior está dado por,

$$\bigwedge_{i \leq j}^s dt_{ij} + \text{términos que envuelven a } dH_1. \quad (\text{A.14})$$

Finalmente al multiplicar (A.13) y (A.14), y sin tomar en cuenta a los términos que envuelven a dH_1 , el producto del lado derecho de (A.9) es igual a,

$$\prod_{i=1}^s t_{ii}^{r-i} (dT)(H_1^t dH_1).$$

■

El siguiente lema es similar al anterior y con el cual se puede calcular el producto exterior de la factorización QR , cuando el rango de la matriz es igual al número de renglones.

Lema A.6 . Si Z es una matriz en $\mathfrak{R}^{r \times s}$, ($r \leq s$) de rango r y esta se puede expresar como $Z = HT$, donde $H \in \mathfrak{R}^{r \times r}$ es una matriz ortonormal, es decir, $H^t H = I_s$, y $T \in \mathfrak{R}^{r \times s}$ es una matriz quasi-triangular superior con elementos en la diagonal positivos. Se puede escribir a la matriz $T = [T_1 | T_2]$, donde $T_1 \in \mathfrak{R}^{r \times r}$ y $T_2 \in \mathfrak{R}^{r \times (s-r)}$, entonces

$$(dZ) = \prod_{i=1}^r t_{ii}^{r-i} (dT)(H^t dH).$$

Demostración:

Dado que $Z = HT$, entonces $dZ = dHT + HdT$, así que al multiplicar esta última expresión por la matriz H^t se tiene que $H^t dZ = H^t dHT + H^t HdT$ y dado que $H^t H = I_m$, se tiene que,

$$dZ = H^t dHT + dT.$$

Por otra parte $(H^t dZ) = |H^t|^s(dZ) = (dZ)$.

Como H es una matriz ortogonal y del lema A.5. se sabe que la matriz $H^t dH$ es una matriz anti-simétrica en $\mathfrak{R}^{r \times r}$ y que se expresa como,

$$H^t dH = \begin{pmatrix} h_1^t \\ h_2^t \\ \vdots \\ h_r^t \end{pmatrix} \begin{pmatrix} dh_1, dh_2, \dots, dh_r^t \end{pmatrix} = \begin{pmatrix} h_1^t dh_1 & h_1^t dh_2 & \dots & h_1^t dh_r \\ h_2^t dh_1 & h_2^t dh_2 & \dots & h_2^t dh_r \\ \vdots & \vdots & \ddots & \vdots \\ h_r^t dh_1 & h_r^t dh_2 & \dots & h_r^t dh_r \end{pmatrix}.$$

Por anti-simetría,

$$\begin{pmatrix} 0 & h_1^t dh_2 & \dots & h_1^t dh_r \\ h_2^t dh_1 & 0 & \dots & h_2^t dh_r \\ \vdots & \vdots & \ddots & \vdots \\ h_r^t dh_1 & h_r^t dh_2 & \dots & 0 \end{pmatrix}.$$

Posmultiplicando por la matriz T , es decir

$$H^t dHT = H^t dH(T_1|T_2) = (H^t dHT_1|H^t dHT_2).$$

Todos los elementos de la forma $h^t dh_j$ son ignorados si aparecen en las columnas previas, así

$$\left(\begin{array}{cccc|cccc} 0 & * & \dots & * & * & * & * & * \\ h_2^t dh_1 t_{11} & * & \dots & * & * & * & * & * \\ h_3^t dh_1 t_{11} & h_3^t dh_2 t_{22} & \dots & * & * & * & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_r^t dh_1 t_{11} & h_r^t dh_2 t_{22} & \dots & h_{r-1}^t dh_{r-1} t_{r-1,r-1} & * & * & * & * \end{array} \right).$$

El producto exterior columna por columna esta dado por,

$$\begin{aligned} t_{11}^{r-1} \bigwedge_{j=2}^r h_j^t dh_1 t_{22}^{r-2} \bigwedge_{j=3}^r h_j^t dh_2 t_{r-1,r-1} h_r^t dh_{r-1} &= t_{11}^{r-1} t_{22}^{r-2} t_{r-1,r-1} \bigwedge_{i=1}^r \bigwedge_{j=i+1}^r h_j^t dh_i \\ &= \prod_{i=1}^r t_{ii}^{r-i} (H^t dH). \end{aligned}$$

Finalmente se puede verificar que,

$$(dT) = \bigwedge_{i \leq j}^r dt_{ij} \bigwedge_{i=1}^r \bigwedge_{j=r+1}^r dt_{ij},$$

y procediendo como en el lema A.5.,

$$(dZ) = \prod_{i=1}^r t_{ii}^{r-i} (dT)(H^t dH).$$

■

Los resultados de los lemas A.5. y A.6. se pueden resumir en el siguiente teorema.

Teorema A.6 . Si $Z = HT \in \mathbb{R}^{r \times s}$ y el rango de la matriz es igual a $K = \min(s, r)$, entonces

$$(dZ) = \prod_{i=1}^K t_{ii}^{r-i} (dT)(H^t dH),$$

donde T y H son matrices quasi-triangular superior y semiortogonal respectivamente.

En el análisis de configuraciones usualmente existen más puntos que dimensiones ($N > K$), y H opera del lado izquierdo en las factorizaciones QR y polar, lo cual es el caso contrario del análisis multivariado, donde la matriz H opera del lado derecho y se tienen más dimensiones que puntos. Por lo tanto se tiene el siguiente resultado.

Corolario A.2 . Si $Z = TH \in \mathbb{R}^{r \times s}$ y el rango de la matriz es igual a $K = \min(s, r)$, entonces

$$(dZ) = \prod_{i=1}^K t_{ii}^{r-i} (dT)(HdH^t),$$

donde T es una matriz quasi-triangular inferior y H es una matriz semiortogonal (filas ortogonales).

El siguiente teorema involucra la forma de calcular el producto exterior de la transpuesta de una matriz por si misma, resultado de gran importancia al definir la distribución Whisart.

Teorema A.7 . Con las suposiciones del teorema A.6. y $A = Z^t Z$, se tiene que

$$(dZ) = 2^{-s}(\det A)^{\frac{r-s-1}{2}}(dA)(H_1^t dH_1).$$

Demostración:

Del teorema A.5.

$$(dZ) = \prod_{i=1}^s t_{ii}^{r-i} (dT)(H_1^t dH_1). \quad (\text{A.15})$$

Se tiene que $A = Z^t Z = T^t T$, de manera tal que,

$$(dA) = 2^s \prod_{i=1}^s t_{ii}^{-s-1+i} (dA),$$

de donde se desprende que,

$$(dT) = 2^{-s} \prod_{i=1}^s t_{ii}^{-s-1+i} (dA). \quad (\text{A.16})$$

Al sustituir (A.16) en (dZ) de (A.15) se tiene que,

$$\begin{aligned} (dZ) &= 2^{-s} \prod_{i=1}^s t_{ii}^{r-s-1} (dA)(H_1^t dH_1), \\ &= 2^{-s}(\det A)^{\frac{r-s-1}{2}}(dA)(H_1^t dH_1). \end{aligned}$$

Dado que

$$\prod_{i=1}^s t_{ii} = \det T = (\det T^t T)^{\frac{1}{2}} = \det A^{\frac{1}{2}}.$$

■

Corolario A.3 . Con las suposiciones del corolario A.2. y definiendo $A = ZZ^t$, se tiene que

$$(dZ) = 2^{-s}(\det A)^{\frac{r-s-1}{2}}(dA)(H_1 dH_1^t).$$

Teorema A.8 . (*Descomposición Polar*) Sea $X \in \mathbb{R}^{r \times s}$, $r \geq s$ y de rango s , entonces existe $H_1 \in V_{s,r}$ y $R \in \mathbb{R}^{s \times s}$, $R^t = R$ tal que $X = H_1 R$.

Demostración: Golub (1993).

Por analogía al caso de la factorización QR , se tiene la siguiente definición alternativa para la factorización polar.

Corolario A.4 . Sea $X \in \mathbb{R}^{r \times s}$, $r \leq s$ y de rango r , entonces existe $H_1 \in V_{r,s}$ y $R \in \mathbb{R}^{r \times r}$, $R^t = R$ tal que $X = RH_1$.

Teorema A.9 . Bajo las suposiciones del teorema A.8. se tiene que

$$(dX) = |R|^{r-s} \prod_{i < j}^s (D_{ii} - D_{jj})(dR)(H_1^t dH_1),$$

donde $D = \text{diag}(D_{11}, \dots, D_{ss})$ y D_{ii} es el i -ésimo eigenvalor de la matriz R .

Demostración:

Este resultado fue establecido por Cadet(1996) bajo las variedades de Riemann a través del cálculo del determinante de Gram, pero con respecto a la medida de Riemann ds , para el cual

$$\text{Vol}(V_{s,r}) = 2^{\lfloor \frac{s(s+3)}{4} \rfloor} \frac{\pi^{\lfloor \frac{rs}{2} \rfloor}}{\Gamma_s\left(\frac{r}{2}\right)} = \int_{V_{s,r}} ds.$$

Ahora si

$$\mu = \int_{\mathcal{D}} (H_1^t dH_1) \quad y \quad \mu_1 = \int_{\mathcal{D}} ds,$$

y de la unicidad de la medida de Haar $\mu_1 = \alpha\mu, \alpha \in \mathbb{R} > 0$. Para determinar α recuerde que,

$$\int_{V_{s,r}} (H_1^t dH_1) = \frac{2^s \pi^{\frac{rs}{2}}}{\Gamma_s\left(\frac{r}{2}\right)},$$

luego

$$\frac{\int_{V_{s,r}} ds}{\int_{V_{s,r}} (H_1^t dH_1)} = \frac{2^{\lfloor \frac{s(s+3)}{4} \rfloor} \frac{\pi^{\lfloor \frac{rs}{2} \rfloor}}{\Gamma_s\left(\frac{r}{2}\right)}}{\frac{2^s \pi^{\frac{rs}{2}}}{\Gamma_s\left(\frac{r}{2}\right)}} = 2^{\lfloor \frac{s(s-1)}{4} \rfloor},$$

esto es $ds = 2^{\lfloor \frac{s(s-1)}{4} \rfloor} (H_1^t dH_1)$.

Sustituyendo este resultado en la expresión para (dX) en Cadet, se obtiene el resultado deseado. ■

Corolario A.5 . Bajo las suposiciones del corolario A.4. se tiene que

$$(dX) = |R|^{r-s} \prod_{i < j}^s (D_{ii} - D_{jj}) (dR) (H_1 dH_1^t),$$

donde $D = \text{diag}(D_{11}, \dots, D_{ss})$ y D_{ii} es el i -ésimo eigenvalor de la matriz R .

APÉNDICE B

MEDIDAS INVARIANTES

MEDIDAS INVARIANTES

En este apéndice se definen las medidas invariantes con las cuales es posible calcular la distribución uniforme asociada a una matriz ortogonal. Se comienza con una propiedad relacionada con los elementos de la matriz H_1 en la variedad de Stiefel.

Proposición B.1 . Sea $H_1 \in V_{s,r}$, entonces existen $\frac{s(s+1)}{2}$ funciones independientes de los elementos sr de $H_1 \in V_{s,r}$, implicados en la ecuación $H_1^t H_1 = I_s$.

Demostración:

Sea $H_1 = [h_1 h_2 \cdots h_s]$ una matriz en $\mathbb{R}^{r \times s}$, donde h_1, h_2, \dots, h_s son vectores en \mathbb{R}^r , entonces,

$$H_1^t H_1 = \begin{bmatrix} h_1^t \\ h_2^t \\ \vdots \\ h_s^t \end{bmatrix} [h_1 h_2 \cdots h_s] = \begin{bmatrix} h_1^t h_1 & h_1^t h_2 & \cdots & h_1^t h_s \\ h_2^t h_1 & h_2^t h_2 & \cdots & h_2^t h_s \\ \vdots & \vdots & \ddots & \vdots \\ h_s^t h_1 & h_s^t h_2 & \cdots & h_s^t h_s \end{bmatrix},$$

donde $h_i^t h_j = h_j^t h_i$ para $i, j = 1, 2, \dots, s$, de manera tal que la matriz resultante es simétrica y los elementos que aparecen debajo de la diagonal principal son dependientes a su correspondiente simétrico que aparece sobre la diagonal superior.

Al igualar cada componente con las de la matriz I_s se tienen que resolver $\frac{s(s+1)}{2}$ ecuaciones que dependen del producto interior de los vectores mencionados anteriormente.

■

Observe que los elementos de H_1 se pueden identificar como las coordenadas de un punto en una superficie de dimensión $sr - \frac{s(s+1)}{2}$ contenida en el espacio sr dimensional.

Sea $H_1 = (h_{ij})$, ($i = 1, \dots, r; j = 1, \dots, s$), la superficie generada es una esfera de radio $s^{\frac{1}{2}}$ en el espacio sr dimensional y está descrita por la ecuación,

$$\sum_{i=1}^r \sum_{j=1}^s h_{ij}^2 = s.$$

Existen dos casos especiales:

Caso 1: Si $s = r$, el conjunto $V_{s,s} \equiv O(s)$ es llamado el conjunto de matrices ortogonales o grupo ortogonal, con la operación usual de multiplicación de matrices.

Los elementos de $H \in O(s)$ pueden ser representados como las coordenadas de un punto en la superficie correspondiente a una esfera de radio $s^{\frac{1}{2}}$ en $\mathbb{R}^{r \times r}$.

Caso 2: Si $s = 1$, el conjunto $V_{1,r} \equiv S_r \equiv \{h \in \mathbb{R}^{r \times 1} : h^t h = 1\}$, representa una esfera unitaria en \mathbb{R}^r .

Por otra parte en la demostración del lema A.5. la forma diferencial de la matriz H_1 se define como,

$$(H_1^t dH_1) = \bigwedge_{i=1}^s \bigwedge_{j=i+1}^r h_j^t dh_i.$$

Para el caso de un grupo ortogonal de matrices $O(s)$ esta es,

$$(H^t dH) = \bigwedge_{i < j}^s h_j^t dh_i,$$

este producto describe el producto exterior de todos los elementos de la subdiagonal de la matriz anti-simétrica $H^t dH$.

Para el grupo ortogonal $O(s)$ se tiene la propiedad de invarianza que se da al operar por la izquierda o por la derecha a la matriz H y sirve para definir la medida invariante de un elemento $D \in O(s)$.

Proposición B.2 . Si $H_1 \in V_{r \times s}$, $Q \in O(s)$ y $(H_1^t dH_1)$ es la forma diferencial de la matriz H_1 , entonces el diferencial de las matrices QH_1 y $H_1 Q^t$ es el mismo que el de la matriz H_1 , es decir $([QH_1]^t d[QH_1]) = ([H_1 Q]^t d[H_1 Q]) = (H_1^t dH_1)$.

Demostración.

$$H_1^t dH_1 \rightarrow H_1^t Q^t Q dH_1 \rightarrow H_1^t dH_1, \quad \text{dado que } Q^t Q = I_s.$$

■

Existe un resultado similar si se multiplica por la derecha a la matriz H_1 , se demuestra utilizando el teorema A.4. en base al siguiente argumento. Si $H_1 \rightarrow H_1 Q$ para $Q \in O(s)$, entonces

$$(H_1^t dH_1) \rightarrow (H_1 Q)^t dH_1 Q = (Q^t H_1^t dH_1 Q) = \det Q^{s-1} (H_1^t dH_1) = (H_1^t dH_1).$$

■

Esta forma diferencial define una medida μ en $O(s)$, dada por

$$\mu(D) = \int_D (H^t dH), \quad D \in O(s).$$

Donde $\mu(D)$ representa el área de la superficie de la región D , usualmente recibe el nombre de volumen para el caso de la variedad ortogonal. Esta medida tiene la propiedad de ser invariante bajo una traslación ya sea por la izquierda o por la derecha, es decir

$$\mu(QD) = \mu(DQ) = \mu(D), \quad \forall Q \in O(s).$$

Definición B.1 . (*Medida de Haar*) La medida descrita anteriormente recibe el nombre de medida de Haar o medida invariante en $O(s)$. Esta medida existe y es única, difiere de cualquier otra medida en $O(s)$ por un múltiplo finito de μ .

Definición B.2 . (*Volumen*) La medida del conjunto $O(s)$ recibe el nombre de volumen y se define como,

$$\text{Vol}[O(s)] = \mu[O(s)] = \int_{O(s)} (H^t dH).$$

Ejemplo B.1 . Para $s = 2$, es decir para $O(2)$ el conjunto de todas las matrices de dimension 2, se puede parametrizar de la siguiente forma,

$$H = \begin{bmatrix} \cos \theta & -\operatorname{sen} \theta \\ -\operatorname{sen} \theta & \cos \theta \end{bmatrix}; \text{ para } 0 < \theta \leq 2\pi.$$

De aquí,

$$(H^t dH) = h_2^t dh_1 = \begin{pmatrix} -\operatorname{sen} \theta & \cos \theta \end{pmatrix} \begin{pmatrix} -\operatorname{sen} \theta d\theta \\ \cos \theta d\theta \end{pmatrix} = (\operatorname{sen}^2 \theta + \cos^2 \theta) d\theta = d\theta.$$

De manera tal que el volumen de $O(2)$ se calcula a partir de la integral del diferencial $(H^t dH)$ e integrando con respecto a la parametrización anterior, es decir,

$$\operatorname{Vol} [O^+(2)] = \int_{O^+(2)} (H^t dH) = \int_0^{2\pi} d\theta = 2\pi.$$

Ahora al considerar la forma diferencial $(H_1^t dH_1)$ en general, se tiene que

$H_1 \in V_{s,r}$, de manera tal que,

$$(H_1^t dH_1) \equiv \bigwedge_{i=1}^s \bigwedge_{j=i+1}^r h_j^t dh_i,$$

donde $H = [H_1 : H_2] = [h_1 \dots h_s | h_{s+1} \dots h_r] \in O(r)$ y H_2 está en función de H_1 , se puede verificar que esta forma diferencial no depende de la elección de la matriz H_2 y es invariante bajo las transformaciones QH_1 y H_1P , donde Q y $P \in O(r)$. Con estas propiedades se define la medida invariante con respecto a la variedad de Stiefel $(V_{s,r})$.

Definición B.3 . (*Volumen bajo la variedad de Stiefel*) El volumen de la variedad de Stiefel se define como,

$$\text{Vol} [V_{s,r}] = \int_{V_{s,r}} (H_1^t dH_1).$$

Para una matriz $H_1 \in V_{s,r}$, se tiene el siguiente resultado.

Teorema B.1 . Sea $H_1 \in V_{s,r}$, el volumen de la variedad de Stiefel está determinado por,

$$\text{Vol} [O(s)] = \frac{2^s \pi^{\binom{s}{2}}}{\Gamma_s \left(\frac{s}{2} \right)}.$$

Demostración: Muirhead(1982)

De aquí se genera el siguiente corolario para el caso de un grupo ortogonal.

Corolario B.1 . Si $H \in O(s)$, entonces

$$\text{Vol} [O(s)] = \frac{2^s \pi^{\binom{s}{2}}}{\Gamma_s \left(\frac{s}{2} \right)}.$$

Observación B.1 . Las medidas de los grupos $O(s)$ y $V_{s,r}$ están dadas sin normalizar, para que estas representen una medida de probabilidad es necesario que se normalicen. Esto se realiza al redefinir el diferencial (dH) respectivamente de la siguiente forma,

$$(dH) \equiv \frac{1}{\text{Vol}[O(s)]} (H^t dH) = \frac{\Gamma_s \left(\frac{s}{2} \right)}{2^s \pi^{\binom{s}{2}}} \bigwedge_{i \leq j}^s h_j^t dh_i,$$

de manera tal que,

$$\int_{O(s)} (dH) = 1.$$

La medida μ^* en $O(s)$ ya normalizada se define a partir de la siguiente integral,

$$\mu^*(D) \equiv \int_D (dH) = \frac{1}{\text{Vol}[O(s)]} \int_D (H^t dH), \quad [D \in O(s)].$$

La medida obtenida y asociada con la forma diferencial (dH) es una medida de probabilidad, que es llamada en algunos casos la distribución invariante de Haar en $O(s)$ y se trabaja de manera similar para el grupo $V_{s,r}$. En el caso especial de $s = 1$, se obtiene la distribución uniforme de una esfera unitaria S_r en \mathbb{R}^n y es una distribución invariante bajo transformaciones ortogonales.

APÉNDICE C

POLINOMIOS ZONALES

POLINOMIOS ZONALES

La generalización de la distribución χ^2 no centrada al caso multidimensional da como resultado la distribución Wishart no centrada. La distribución χ^2 se escribe frecuentemente como una serie infinita, esta corresponde a una expansión en series de potencias del argumento exponencial de la densidad normal y después se integra término a término. En el caso multivariado la integral se evaluó con respecto a la medida invariante de Haar sobre el grupo ortogonal $O(r)$, los polinomios zonales son una herramienta indispensable para determinar dicha integración. James (1955-56) calculó la integración con respecto a la medida de Haar y expresó la densidad de la distribución Whisart no-centrada como una serie infinita de polinomios zonales.

Existen diferentes propuestas para definir a los polinomios zonales, una de ellas planteada por Akimichi Takemura (1984), basada en el análisis multivariado y el Álgebra Lineal. Muirhead (1982), definió a los polinomios zonales como eigenfunciones de operadores de ecuaciones diferenciales. Farrel's (1976), utilizó el H^* -álgebra de Loomis. Saw(1977), derivó algunas propiedades usando la distribución normal y Whisart multivariada. Finalmente Kates (1980), se basó en funciones esféricas y en la representación de grupos. Muchos autores consideran que las tres últimas propuestas no caen dentro del área de la Estadística Matemática, en este trabajo se utiliza la definición dada por Muirhead.

Los polinomios zonales tienen la propiedad de ser polinomios simétricos y

homogéneos y se definen a partir de una matriz simétrica. Cuando se dice que el polinomio es simétrico significa que si contiene al monomio $y_1^{k_1} y_2^{k_2} \cdots y_s^{k_s}$, entonces también contiene al monomio $y_{l_1}^{k_1} y_{l_2}^{k_2} \cdots y_{l_s}^{k_s}$, donde (l_1, l_2, \dots, l_s) es cualesquier permutación de los enteros $\{k_1, k_2, \dots, k_s\}$. Al decir que un polinomio es homogéneo significa que no contiene términos constantes y al tener grado r la suma de las potencias en cada monomio es igual a éste.

Definición C.1 . (*Partición*) La partición de un entero k se define como la siguiente secuencia ordenada de números enteros $\mathbf{k} = (k_1, k_2, k_3, \dots)$, donde $\sum k_i = k$, y $k_1 \geq k_2 \geq k_3 \geq \dots$, y además $k_i \geq 0$.

Es decir la suma de los elementos de la secuencia ordenada debe ser igual al número k y todos estos elementos deben de ser no negativos, además de estar en orden decreciente.

Ejemplo C:1 . Para el entero tres se tienen las siguientes particiones: (3), (2, 1), (1, 1, 1) siendo éstas las únicas a menos de que se necesite agregar ceros a la izquierda. La partición (2, 1) también se puede escribir como (2, 1, 0, 0, 0), esto indica que en una partición pueden aparecer términos nulos en las últimas posiciones tantas veces como sea necesario.

Definición C.2 . (*Relación de orden*) Dadas dos particiones $\mathbf{u} = (u_1, u_2, \dots, u_r)$ y $\mathbf{v} = (v_1, v_2, \dots, v_s)$ del entero k , se dice que $\mathbf{u} \geq \mathbf{v}$ si, y sólo si $u_i \geq v_i$, para el primer índice donde los elementos de las secuencias sean diferentes.

Ejemplo C.2 . En el ejemplo (C.1) se tiene que $(3) > (2, 1) > (1, 1, 1)$.

Definición C.3 . (*Monomio de mayor peso*) Dadas dos particiones $\mathbf{u} = (u_1, u_2, \dots, u_r)$ y $\mathbf{v} = (v_1, v_2, \dots, v_s)$ del entero k , donde $\mathbf{u} > \mathbf{v}$, se dice que el monomio $y_1^{u_1} y_2^{u_2} \cdots y_s^{u_s}$ es de mayor peso al del monomio $y_1^{v_1} y_2^{v_2} \cdots y_s^{v_s}$.

Ejemplo C.3 . Continuando con el ejemplo (C.2) se tiene que el monomio y_1^3 es de mayor peso al monomio $y_1^2 y_2$ y este a la vez es de mayor peso que el monomio $y_1 y_2 y_3$.

Definición C.4 . (*Polinomios Zonales*) Si $Y \in \mathbb{R}^{s \times s}$ es una matriz simétrica con eigenvalores y_1, \dots, y_s y dada una partición $\mathbf{k} = (k_1, \dots, k_s)$ del entero positivo k en no más de s partes. El polinomio zonal de Y correspondiente a la partición \mathbf{k} y denotado por $C_{\mathbf{k}}(Y)$ es un polinomio simétrico, homogéneo, de grado k y definido en sus eigenvalores y_1, \dots, y_s , de manera tal que satisface las siguientes tres condiciones:

(i) El término de mayor peso en $C_{\mathbf{k}}$ es el monomio $y_1^{k_1} y_2^{k_2} \cdots y_s^{k_s}$, es decir es de la forma

$$C_{\mathbf{k}}(Y) = d_{\mathbf{k}} y_1^{k_1} y_2^{k_2} \cdots y_s^{k_s} + \text{términos de menor peso, donde } d_{\mathbf{k}} \text{ es constante.}$$

(ii) $C_{\mathbf{k}}$ es una función del operador diferencial Δ_Y y éste está dado por,

$$\Delta_Y = \sum_{i=1}^s y_i^2 \frac{\partial^2}{\partial y_i^2} + \sum_{i=1}^s \sum_{\substack{j=1 \\ j \neq i}}^s \frac{y_i^2}{y_i - y_j} \frac{\partial}{\partial y_i}.$$

(iii) De entre todas las posibles particiones del entero k , los polinomios zonales poseen coeficiente unitario en la expresión $(\text{tr } Y)^k$; Es decir,

$$(\text{tr } Y)^k = (y_1 + y_2 + \cdots + y_s)^k = \sum_{\mathbf{k}} C_{\mathbf{k}}(Y).$$

Teorema C.1 . El polinomio zonal $C_{\mathbf{k}}(Y)$ correspondiente a la partición $\mathbf{k} = (k_1, \dots, k_s)$ de entero k , satisface la ecuación diferencial,

$$\Delta_Y C_{\mathbf{k}}(Y) = [\rho_{\mathbf{k}} + k(s-1)]C_{\mathbf{k}}(Y),$$

donde Δ_Y esta definida por la condición (ii) y

$$\rho_{\mathbf{k}} = \sum_{i=1}^s k_i(k_i - i).$$

Demostración:

Por las condiciones (i) y (ii), es suficiente con demostrar que,

$$\Delta_Y y_1^{k_1} \cdots y_s^{k_s} = [\rho_{\mathbf{k}} + k(s-1)]y_1^{k_1} \cdots y_s^{k_s} + \text{términos de menor peso.}$$

Utilizando la diferenciación se ve que,

$$\begin{aligned} \Delta_Y y_1^{k_1} \cdots y_s^{k_s} &= y_1^{k_1} \cdots y_s^{k_s} \left[\sum_{i=1}^s k_i(k_i - i) + \sum_{i=1}^s \sum_{\substack{j=i \\ j \neq i}}^s \frac{y_i k_i}{y_i - y_j} \right] \\ &= y_1^{k_1} \cdots y_s^{k_s} \left[\sum_{i=1}^s k_i^2 - k + \sum_{i=1}^{s-1} \sum_{i=j+1}^s \left(\frac{y_i k_i}{y_i - y_j} + \frac{y_j k_j}{y_j - y_i} \right) \right]. \end{aligned}$$

Dado que,

$$\frac{y_i k_i}{y_i - y_j} = k_i + \frac{y_j k_i}{y_i - y_j}.$$

Se sigue que

$$\begin{aligned} \Delta_Y y_1^{k_1} \cdots y_s^{k_s} &= y_1^{k_1} \cdots y_s^{k_s} \left\{ \sum_{i=1}^s k_i^2 - k + \sum_{i=1}^{s-1} \sum_{j=i+1}^s \left[k_i + \frac{y_j}{y_j - y_i} (k_j - k_i) \right] \right\} \\ &= y_1^{k_1} \cdots y_s^{k_s} \left[\sum_{i=1}^s k_i^2 - k + \sum_{i=1}^{s-1} k_i (s - i) \right] + \text{m\u00e1s t\u00e9rminos} \\ &\hspace{15em} \text{de menor orden.} \end{aligned}$$

De aqu\u00ed,

$$\sum_{i=1}^{s-1} k_i (s - i) = \sum_{i=1}^s k_i (s - i) = ks - \sum_{i=1}^s i k_i,$$

y se obtiene el resultado deseado. ■

Definici\u00f3n C.5 . (*Funci\u00f3n monomial sim\u00e9trica*) Dada una partici\u00f3n $\mathbf{k} = (k_1, k_2, k_3, \dots)$ del entero k , la funci\u00f3n monomial correspondiente a los eigenvalores y_1, \dots, y_s , se define como,

$$M_{(\mathbf{k})} = \sum y_{i_1}^{k_1} y_{i_2}^{k_2} \cdots y_{i_p}^{k_p}.$$

La sumatoria se realiza sobre todas las permutaciones de los p n\u00fameros enteros k_i , diferentes de cero, donde $i = 1, 2, \dots, s$. La funci\u00f3n monomial se caracteriza por la siguiente ecuaci\u00f3n,

$$M_{(\mathbf{k})} = y_1^{k_1} y_2^{k_2} \cdots y_p^{k_p} + \text{términos simétricos.}$$

Ejemplo C.4 . Dada una matriz simétrica $Y \in \mathbb{R}^{s \times s}$ para el entero uno el término monomial se escribe de la siguiente forma,

$$M_{(1)} = y_1 + y_2 + \cdots y_s.$$

Para el número entero dos existen dos particiones, estas son (2) y $(1, 1)$, las funciones monomiales asociadas son las siguientes,

$$M_{(2)} = y_1^2 + y_2^2 + \cdots y_s^2.$$

$$M_{(1,1)} = \sum_{i < j}^s y_i y_j.$$

El polinomio zonal asociado a la partición (1) del entero uno con respecto a la matriz $Y \in \mathbb{R}^{s \times s}$ cumple la condición (iii) de la definición de los polinomios zonales, esto implica que $C_{(1)} = \text{tr } Y = y_1 + y_2 + \cdots + y_s = M_{(1)}(Y)$.

Para el entero dos existen dos polinomios zonales $C_{(2)}$ y $C_{(1,1)}$. Estos satisfacen la condición (i) de manera tal que el polinomio es simétrico, homogéneo y de grado 2, es decir,

$$\begin{aligned}
C_{(2)} &= d_{(2)}y_1^2 + \text{términos de menor peso}, \\
&= d_{(2)}y_1^2 + \beta y_1 y_2 + \text{términos simétricos}, \\
&= d_{(2)}(y_1^2 + \cdots + y_s^2) + \beta(y_1 y_2 + \cdots + y_{s-1} y_s), \\
&= d_{(2)}M_{(2)}(Y) + \beta M_{(1,1)}(Y),
\end{aligned}$$

donde d y β son constantes. Por otra parte se tiene que,

$$\begin{aligned}
C_{(1,1)} &= d_{(1,1)}y_1 y_2 + \text{términos de menor peso}, \\
&= d_{(1,1)}y_1 y_2 + \cdots + y_{s-1} y_s, \\
&= d_{(1,1)}M_{(1,1)}(Y),
\end{aligned}$$

y de la condición (iii),

$$\begin{aligned}
(\text{tr } Y)^2 &\equiv M_{(2)}(Y) + 2M_{(1,1)}(Y), \\
&= C_{(2)}(Y) + C_{(1,1)}(Y), \\
&= d_{(2)}M_{(2)}(Y) + (\beta + d_{(1,1)})M_{(1,1)}(Y),
\end{aligned}$$

comparando coeficientes se concluye que $d_{(2)} = 1$ y $d_{(1,1)} = 2 - \beta$. De donde

$$C_{(2)}(Y) = M_{(2)}(Y) + M_{(2)}(Y) + \beta M_{(1,1)}(Y), \quad (\text{C.1})$$

y

$$C_{(1,1)}(Y) = (2 - \beta)M_{(1,1)}(Y). \quad (\text{C.2})$$

La constante β se cálcula utilizando la ecuación diferencial en $C_{(2)}(Y)$. Del teorema C.1., $\rho_2 = 2(2 - 1)$ y $C_{(2)}(Y)$ satisface la ecuación diferencial

$$\Delta_Y C_{(2)}(Y) = 2sC_{(2)}(Y).$$

Al operar sobre los polinomois zonales anteriores se tiene que,

$$\Delta_Y M_{(2)}(Y) = 2sM_{(2)}(Y) + 2M_{(1,1)}(Y),$$

y

$$\Delta_Y M_{(1,1)}(Y) = (2s - 3)M_{(1,1)}(Y).$$

Sustituyendo (C.1) en (C.2), esto es igual a,

$$2sM_{(2)}(Y) + 2M_{(1,1)}(Y) + \beta(2s - 3)M_{(1,1)}(Y) = 2sM_{(2)}(Y) + 2s\beta M_{(1,1)}(Y).$$

Igualando los coeficientes de $M_{(1,1)}(Y)$ en ambos lados se obtiene que $\beta = \frac{2}{3}$, de manera tal que los dos polinomios zonales correspondientes al entero $k = 2$, son

$$C_{(2)}(Y) = M_{(2)}(Y) + \frac{2}{3}M_{(1,1)}(Y),$$

y

$$C_{(1,1)}(Y) = \frac{4}{3}M_{(1,1)}(Y).$$

El siguiente teorema es un resultado muy útil al evaluar ciertas integrales.

Teorema C.2 . Si X_1 es una matriz definida positiva en $\mathbb{R}^{s \times s}$ y X_2 es una matriz simétrica en $\mathbb{R}^{s \times s}$, entonces

$$\int_{O(s)} C_{(\mathbf{k})}(X_1 H X_2 H^t)(dH) = \frac{C_{(\mathbf{k})}(X_1)C_{(\mathbf{k})}(X_2)}{C_{(\mathbf{k})}(I_s)},$$

donde (dH) es la medida invariante normalizada sobre $O(s)$.

A continuación se da la definición de una función hipergeométrica cuyo argumento es una matriz X definida positiva.

Definición C.6 . (*Función hipergeométrica*) La función hipergeométrica de una matriz $X > 0$, se define como,

$${}_pF_q(a_1, \dots, a_p; b_1, \dots, b_q : X) = \sum_{k=0}^{\infty} \sum_{\mathbf{k}} \frac{(a_1)_{\mathbf{k}} \cdots (a_p)_{\mathbf{k}} C_{\mathbf{k}}(X)}{(b_1)_{\mathbf{k}} \cdots (b_q)_{\mathbf{k}} k!},$$

donde $\sum_{\mathbf{k}}$ denota la sumatoria sobre todas las particiones $\mathbf{k} = \{\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m\}$, donde $\mathbf{k}_1 \leq \mathbf{k}_2 \leq \dots \leq \mathbf{k}_m \leq \mathbf{0}$, con respecto a k y $C_{\mathbf{k}}(X)$ es el polinomio zonal de X

correspondiente a la partición \mathbf{k} y donde el coeficiente generalizado hipergeométrico $(a)_{\mathbf{k}}$ está dado por,

$$(a)_{\mathbf{k}} = \prod_{i=1}^s \left[a - \frac{1}{2}(i-1) \right]_{k_i},$$

donde $(a)_k = a(a+1) \cdots (a+k-1)$, $a_0 = 1$.

Esta expansión en polinomios zonales fue propuesta por Constantino (1963).

El siguiente ejemplo muestra dicha expansión para el caso en que $q = p = 0$,

Ejemplo C.5 .

$$\begin{aligned} {}_0F_0(X) &= \sum_{k=0}^{\infty} \sum_{\mathbf{k}} \frac{C_{\mathbf{k}(X)}}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(\text{tr } X)^k}{k!} \\ &= \exp[\text{tr}(X)]. \end{aligned}$$

La siguiente propiedad de una función hipergeométrica esta relacionada con la distribución Whisart no centrada y permite establecer la relación entre la función ${}_0F_1$ y una integral con argumento matricial.

Teorema C.3 . Si $X \in \mathbb{R}^{s \times r}$ es una matriz real con $s \leq r$ y $H = [H_1 : H_2] \in O(r)$, donde $H_1 \in \mathbb{R}^{r \times s}$, entonces

$$\int_{O(r)} \exp[\text{tr}(XH_1)](dH) = {}_0F_1 \left(\frac{r}{2}; \frac{XX^t}{4} \right),$$

donde (dH) denota la medida invariante normalizada en $O(r)$.

Demostración: Muirhead (1982)

Con lema que sigue a continuación se puede calcular la integral sobre un grupo ortogonal. En el caso de que el rango de la matriz no es completo los vectores se pueden agrupar para obtener una matriz ortogonal cuadrada.

Lema C.1 .

$$\int_{O(s)} f(H_1, H_2) (H^t dH) = \int_{H_1 \in V_{k,s}} \int_{K \in O(s-k)} f(H_1, GK) (K^t dK) (H_1^t) (H_1^t dH_1),$$

donde $H = [H_1, H_2]$, $H_1 \in \mathbb{R}^{s \times k}$ y $G = G(H_1)$ es cualquier matriz en $\mathbb{R}^{s \times (s-k)}$, con columnas ortogonales a las de H_1 (es decir $GG^t = I_s - H_1 H_1^t$).

Demostración: Muirhead (1982)

APÉNDICE D

ESTADÍSTICA

ESTADÍSTICA

A continuación se enuncian algunos conceptos básicos de la estadística.

Definición D.1 . (*Matriz aleatoria*) Se dice que una matriz $A \in \mathbb{R}^{r \times s}$ es una matriz aleatoria si cada una de sus componentes es una variable aleatoria.

Definición D.2 . (*Función gama*) La función gama se define como,

$$\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx, t > 0,$$

y se denota por $\Gamma(\cdot)$. Tiene la propiedad $\Gamma(t+1) = t\Gamma(t)$ y si r es un número entero positivo, entonces se verifica que $\Gamma(r+1) = r!$ (el factorial de r). Otra propiedad es que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Definición D.3 . (*Función gama multivariada*) La función gama multivariada se define como,

$$\Gamma_s(a) = \int_{A>0} \exp[\text{tr}(-A)] \det \left[A^{(a-\frac{s+1}{2})} \right] (dA).$$

y se denota por $\Gamma_s(a)$.

Donde $\text{Re}(a) > \frac{s-1}{2}$ y la integral se evalúa sobre el espacio de todas las matrices definidas positivas (simétricas) en $\mathbb{R}^{s \times s}$.

Observación D.1 . Si $s = 1$, la función $\Gamma_1(\cdot)$ define la función gama $\Gamma(\cdot)$ univariada.

Teorema D.1 . Si $\text{Re}(a) > \frac{s-1}{2}$ y Σ es una matriz simétrica en $\mathbb{R}^{s \times s}$ con $\text{Re}(\Sigma) > 0$, entonces

$$\int_{A>0} \exp\left(\frac{-\Sigma^{-1}A}{2}\right) \det(A)^{\left(a-\frac{s+1}{2}\right)} (dA) = \Gamma_s(a)(\det \Sigma)^a 2^{sa}.$$

Demostración: Muirhead(1982)

De este teorema se desprende la siguiente función de densidad una vez que ésta es no negativa e integra 1.

Definición D.4 . (*Función de densidad Wishart*) Si A es una matriz definida positiva, $\text{Re}(a) > \frac{s-1}{2}$, Σ es una matriz simétrica en $\mathbb{R}^{s \times s}$ con $\text{Re}(\Sigma) > 0$ y $r \geq s$, la función de densidad Wishart se define,

$$f(A) = \frac{1}{2^{\left(\frac{sr}{2}\right)} \Gamma_s\left(\frac{r}{2}\right) (\det \Sigma)^{\frac{r}{2}}} \exp\left[\text{tr}\left(\frac{-\Sigma^{-1}A}{2}\right)\right] (\det A)^{\frac{r-s-1}{2}}, \quad (A > 0).$$

El siguiente teorema muestra que la función gama multidimensional se puede expresar como el producto de funciones gamas ordinarias.

Teorema D.2 .

$$\Gamma_s(a) = \pi^{\frac{s(s-1)}{4}} \prod_{i=0}^{s-1} \Gamma\left[a - \frac{i-1}{2}\right], \quad [\text{Re}(a) > \frac{s-1}{2}].$$

Demostración:

Por definición,

$$\Gamma_s(a) = \int_{A>0} \exp[\operatorname{tr}(-A)] \det \left[A^{(a-\frac{s+1}{2})} \right] (dA).$$

Sea $A = T^t T$ donde T es una matriz triangular superior con elementos en la diagonal positivos, entonces

$$\operatorname{tr} A = \operatorname{tr}(T^t T) = \sum_{i \leq j}^s t_{ij}^2,$$

$$\det A = \det T^t T = (\det T)^2 = \prod_{i=1}^s t_{ii}^2,$$

y del teorema A.5.,

$$(dA) = 2^s \prod_{i=1}^s t_{ii}^{(s+1-i)} \bigwedge_{i \leq j}^s dt_{ij}.$$

Entonces

$$\begin{aligned} \Gamma_s(a) &= 2^s \int \cdots \int \exp \left(- \sum_{i \leq j}^s t_{ij}^2 \right) \prod_{i=1}^s t_{ii}^{(2a-i)} \bigwedge_{i \leq j}^s dt_{ij}, \\ &= \prod_{i < j}^s \left[\int_{-\infty}^{\infty} \exp(-t_{ij}^2) \right] \prod_{i=1}^s \left[\int_0^{\infty} \exp(-t_{ii}^2) (a - \frac{i+1}{2}) dt_{ii}^2 \right]. \end{aligned}$$

El resultado se sigue de,

$$\int_{-\infty}^{\infty} \exp(-t_{ij}^2) dt_{ij} = \pi^{\frac{1}{2}},$$

y

$$\int_0^\infty \exp(-t_{ii}^2)(t_{ii}^2)^{(a-\frac{i+1}{2})} dt_{ii}^2 = \Gamma\left[a - \frac{1}{2}(i-1)\right].$$

■

La siguiente definición establece la densidad de una matriz con componentes normales.

Definición D.5 . (*Función normal multivariada*) Si $Y \in \mathbb{R}^{r \times s}$ se distribuye $N(M, C \otimes D)$, donde $C \in \mathbb{R}^{r \times r}$, $D \in \mathbb{R}^{s \times s}$ son matrices definidas positivas, la función de densidad de Y está dada por,

$$\frac{1}{(2\pi)^{\binom{rs}{2}} |C|^{\binom{s}{2}} |D|^{\binom{r}{2}}} \exp\left[\operatorname{tr}\left(\frac{-C^{-1}(Y-M)D^{-1}(Y-M)^t}{2}\right)\right].$$

Definición D.6 . (*Distribución Wishart no centrada*) Si $A = Z^t Z$, $A \in \mathbb{R}^{r \times s}$, donde Z se distribuye normal multivariada $N(M, I_r \otimes \Sigma)$, entonces A posee una distribución Wishart no centrada con r grados de libertad, matriz de covarianza Σ y con matriz de parametros de no centralidad $\Omega = \Sigma^{-1} M^t M$. Se denota por medio de $A \sim W_s(r, \Sigma, \Omega)$.

Si $M = 0$ y $\Omega = 0$, entonces $A \sim W_r(r, \Sigma)$, es decir A se distribuye Wishart centrada. La función de densidad de $A = Z^t Z$ con las condiciones anteriores y $r \geq s$, está dada por

$$\frac{1}{2^{\binom{sr}{2}} \Gamma_s\left(\frac{r}{2}\right) |\Sigma|^{\binom{r}{2}}} \exp\left[\operatorname{tr}\left(\frac{-1\Sigma^{-1}A}{2}\right)\right] |A|^{\binom{r-s-1}{2}} \exp\left[\operatorname{tr}\left(\frac{-\Omega}{2}\right)\right] {}_0F_1\left(\frac{r}{2}; \frac{\Omega\Sigma^{-1}A}{4}\right),$$