

Benemérita Universidad Autónoma de Puebla

Facultad de Ciencias Físico Matemáticas

DE LA DEFINICIÓN DE DISTRIBUCIONES DE PROBABILIDADES PARA TRANSFORMACIONES PROYECTIVAS

TESIS

que para obtener el título de: Licenciado en Matemáticas Aplicadas

presenta:

José Guillermo Herrera Ramírez

dirigido por: Dr. Jean-Bernard Hayet

Puebla, enero 2017

Π

Dedicado a

A mis padres:

María del Pilar Ramírez Juárez y José Alberto Herrera Escutia

A mis hermanos:

Tishbe Pilarh Herrera Ramírez y Angel Alberto Herrera Ramírez IV

Agradecimientos

A Dios.

A mi madre María del Pilar Ramírez Juárez y a mi padre José Alberto Herrera Escutia por su amor y su cariño, por apoyarme incondicionalmente en todas mis decisiones. A mi hermano Angel Alberto Herrera Ramírez y a mi hermana Tishbe Pilarh Herrera Ramírez por su ejemplo de entrega, disciplina y superación, por enseñarme a no rendirme. Los amo.

Al Prof. Juan José Parres Córdova, a la Dra. Lídia Aurora Hernández Rebollar y a la Dra. María Araceli Juárez Ramírez por ayudarme a trazar el camino que he seguido desde mis primeros pasos en la carrera y hasta la culminación de este trabajo.

Al M.C. Edgar Santiago Moyotl Hernández por el tiempo que dedicó para darme elementos para enfrentar problemas algorítmicos y computacionales, al Dr. Wuiyebaldo Fermín Guerrero Sánchez por sus enseñanzas y por motivarme a no sólo resolver problemas teóricos sino también a resolver problemas prácticos. Y junto con el Dr. Carlos Guillén Galván por el tiempo que le dedicaron a revisar este trabajo de tesis. A todos los amigos que hice en la FCFM, en las competencias de Matemáticas y en los congresos, por acompañarme durante estos años. En especial a los Fisiconautas por su hermandad y por todas las risas que pasábamos entre clases.

Al CIMAT por las facilidades que me dio para realizar la estancia de tesis. Al Dr. Ricardo Francisco Vila Freyer por guiarme en temas importantes. A mis amigos del laboratorio de robótica por su compañía y sus consejos.

A mi asesor de tesis el Dr. Jean-Bernard Hayet por su paciencia al explicarme los temas, por revisar mis códigos y demostraciones. Más aún, por brindarme su amistad y por proporcionarme las herramientas necesarias para concluir satisfactoriamente este trabajo. A gradecimientos

Agradecimientos

VIII

Índice general

In	troducción	1
1.	Preliminares 1.1. Visión por computadora 1.2. Estadística	5 5 10
2.	Motivación al análisis estadístico sobre variedades	17
3.	Respaldo matemático: Geometría diferencial3.1. Propiedades de $SL_3(\mathbb{R})$ 3.2. $SL_3(\mathbb{R})$ con estructura métrica3.2.1. Aplicación exponencial3.2.2. Propiedad de invarianza3.2.3. Definición de $expp$ y $logg$ Análisis estadístico sobre $SL_3(\mathbb{R})$ 4.1. Media4.1.1. Otras medidas de tendencia central4.2. Matriz de covarianza	21 21 30 31 32 35 37 37 39 40 46
5 . 6 .	4.2.1. Otras medidas de dispersión $\dots \dots \dots$	 40 48 49 53 61
7.	6.1. Aplicación al problema de clasificación	64 71

Bibliografía

 $\mathbf{75}$

Х

Introducción

Supongamos que tenemos n objetos **planos** distintos y que cada uno representa un modelo M. Supongamos también que, para cada uno de los modelos M, tenemos una muestra aleatoria de fotografías de este objeto tomadas desde distintos puntos de vista. Por ejemplo, en la figura 1 (extraída de [1]), vemos diferentes imágenes de una misma pinza, abierta a distintos ángulos (tres de ellos, más precisamente). Aquí, el modelo M corresponderá al ángulo de apertura de la pinza: pequeño, mediano o grande. Dada una nueva imagen D, de un objeto desde una dirección desconocida, queremos encontrar: a) el modelo M al que corresponde (es decir, en el ejemplo, la apertura de las pinzas) y b) la perspectiva H con la cual lo estamos observando.



Figura 1: Cada fila muestra imágenes con el mismo ángulo de apertura de las pinzas (extraída de [1]).

Este problema pertenece a los llamados "problemas de clasificación", en los cuales se intenta inferir información de la "clase" de un dato, a partir de colecciones de datos-ejemplos. La herramienta utilizada para resolver este tipo de problemas es la inferencia Bayesiana.

En la inferencia Bayesiana utilizamos la información obtenida de una muestra y una distribución a priori de un parámetro desconocido para calcular una densidad posterior que resuma toda la información pertinente acerca del parámetro desconocido [12].

El problema mencionado anteriormente lo podemos expresar, utilizando la inferencia Bayesiana, como:

$$P(M,H|D) = \frac{P(D|M,H)P(M,H)}{P(D)} = \frac{P(D|M,H)P(H|M)P(M)}{P(D)}$$
(1)

es decir, para obtener la probabilidad de tener un modelo M desde un punto de vista H, dada una imagen D, denotada por P(M, H|D), necesitamos calcular la probabilidad de obtener la imagen D dado el modelo M y la perspectiva H, denotada por P(D|M, H); también hay que calcular la distribución a priori de los modelos, denotada por P(M); y la distribución a priori de las perspectivas de cada uno de los modelos, denotado por P(H|M).

El objetivo de este trabajo es dar una definición concreta al término P(H|M), y por ende hablar de estadísticas (nociones de promedio y de varianza) en el espacio de las perspectivas H.

Como estamos utilizando modelos de objetos planos, cada transformación perspectiva H, respecto a una vista fija, la asociamos con una transformación lineal en el plano proyectivo, $H : \mathbb{P}^2 \to \mathbb{P}^2$. A estas transformaciones lineales de \mathbb{P}^2 , las llamamos **homografías** y se pueden definir mediante matrices 3×3 con determinante igual a 1. Recordemos que las matrices que cumplen estas propiedades conforman el grupo lineal especial de orden 3, denotado por $SL_3(\mathbb{R})$. Esto se detallará en el siguiente capítulo.

El problema principal radica en que este grupo no tiene la muy conocida estructura euclidiana de \mathbb{R}^n , lo que complica la definición de distancia, y por

ende las nociones de promedio y varianza.

Metodología

Para este trabajo, partiremos de una muestra aleatoria de imágenes de un mismo objeto desde distintas perspectivas. Luego, encontraremos las homografías subyacentes a esas vistas, que podemos representar con matrices en $SL_3(\mathbb{R})$. Nos concentraremos finalmente en estudiar e implementar las propuestas encontradas en [1] para calcular una media y una matriz de covarianza para un conjunto de datos en $SL_3(\mathbb{R})$. Estos parámetros servirán para definir una función $\phi : SL_3(\mathbb{R}) \to \mathbb{R}$, "similar" a la función de densidad de probabilidad normal, que dará una propuesta consistente para calcular P(H|M). Insistimos en que la muestra aleatoria de las perspectivas H con las que disponemos no vive en un espacio en que naturalmente podemos definir probabilidades. En la figura 2 (extraída de [1]) vemos un ejemplo de un conjunto de imágenes de un reloj de pared tomadas desde distintas perspectivas. En el centro está la imagen correspondiente a la homografía más cercana a la "media" calculada por el algoritmo que describiremos. A la imagen que corresponde a la homografía "media" la llamaremos "vista media".



Figura 2: Diferentes vistas de un reloj de pared con la imagen más cercana a la "vista media" en el centro (extraída de [1]).

Capítulo 1

Preliminares

En esta sección describimos conceptos importantes para la comprensión de este documento. Se hablará de visión por computadora, geometría proyectiva y finalmente de estadística.

1.1. Visión por computadora

Nosotros como humanos podemos percibir estructuras tridimensionales a partir de nuestra vista. Podemos distinguir a dos personas en una fotografía, o saber qué tan lejos o cerca se encuentra un objeto. Sin embargo, una computadora no puede hacer lo mismo "naturalmente".

La visión por computadora es una rama de la computación en la que se trata de describir el mundo que vemos en una o varias imágenes y reconstruir sus propiedades, tales como la forma, iluminación y distribución del color. Hoy en día, la visión por computadora tiene una amplia variedad de aplicaciones en el mundo real, por ejemplo, procesamiento de imágenes médicas, captura de movimiento, detección de rostros, o reconocimiento de objetos para cajas registradoras automáticas [19].

Cuando trabajamos con imágenes, debemos utilizar herramientas geométricas distintas de las usuales, esto porque representamos objetos tridimensionales proyectados en dos dimensiones (por el proceso de proyección inherente a todas las imágenes). Por ejemplo, en la figura 1.1 ilustramos un tablero de ajedrez. Como sabemos tiene rectas paralelas y cuadrados. Si lo observamos desde una perspectiva como en la figura 1.2(a), observaremos que los cuadrados parecen trapecios o romboides. Más aún, en la figura 1.2(b), las prolongaciones de los segmentos de rectas paralelas se intersecan.



Figura 1.1: Tablero de ajedrez.



(a) Desde esta perspectiva, los cuadrados (b) Desde esta perspectiva, las rectas paparecen trapecios o romboides. ralelas se intersecan.

Figura 1.2: Diferentes perspectivas de un tablero de ajedrez.

Para utilizar imágenes, primero debemos hablar de cómo se forman a partir de una cámara. Los ancestros de las cámaras modernas son las llamadas cámaras oscuras, conocidas desde hace siglos. El principio de estos dispositivos es muy simple: En una caja se hace un agujero muy fino (*pinhole*). Para un objeto frente a la caja, por ejemplo, una vela como en la figura 1.3 (extraída de [7]), se observará una imagen invertida de ella en el lado opuesto de la caja. Cada proyección invertida es una fotografía. Llamamos plano de la imagen (*image plane*) al plano donde se obtiene la imagen invertida. En algunas ocasiones es conveniente considerar, en lugar de la imagen invertida, una imagen virtual (*virtual image*) frente al agujero, a la misma distancia que el plano de la imagen. Esta imagen virtual no está invertida y es equivalente a la creada al fondo de la caja [7].



Figura 1.3: Modelo de cámara de "pinhole" (extraída de [7]).

Definiremos sin tanto rigor al plano proyectivo como un espacio que conservará las propiedades geométricas subyacentes a la imagen formada por la cámara oscura, que asocia puntos de \mathbb{R}^3 a proyecciones en \mathbb{R}^2 .

Pensemos en todas las rectas en \mathbb{R}^3 que pasan por el origen. Definimos al **plano proyectivo**, \mathbb{P}^2 , como el conjunto que contiene a todas estas rectas. Notemos que a cada recta la podemos definir sólo con un vector de \mathbb{R}^3 distinto del nulo, y así, para cada recta tomaremos un vector no nulo que será el representante de la recta.

Sea $p = (x, y, z)^t$ un vector en \mathbb{R}^3 tal que $z \neq 0$. Para la recta definida por el vector p, podemos también elegir como "representante" de la recta al vector $p' = (x', y', 1)^t$ con $x' = (\frac{x}{z})$ y $y' = (\frac{y}{z})$. Es decir, proyectaremos al punto en p sobre el plano z = 1 como se ve en la figura 1.4. Es claro que los vectores $(kx, ky, kz)^t$ y $(x, y, z)^t$ tienen la misma proyección en p' cuando $k \neq 0$. Estos vectores son considerados equivalentes, los nombraremos como **vectores homogéneos** y denotaremos la equivalencia como $(kx, ky, kz)^t \equiv (x, y, z)^t$. El plano proyectivo \mathbb{P}^2 se puede ver como el conjunto de las clases de equivalencia construidas con la relación $x' = (\frac{x}{z})$ y $y' = (\frac{y}{z})$. Los puntos para los cuales z = 0 no tienen una proyección sobre el plano z = 1 y son conocidos como puntos ideales (o al infinito). Todos los puntos ideales pertenecen a lo que se conoce como *línea al infinito* [13]. De manera analítica podemos suponer que el origen en \mathbb{R}^3 es el pinhole de la cámara y en el plano z = 1 se encuentran las imágenes con las que trabajaremos.



Figura 1.4: El punto en p' es la proyección del punto en p sobre el plano z = 1. Todos los puntos de la recta que pasa por p y por el origen se proyectan en el mismo punto p'.

Definimos una **transformación proyectiva** como una transformación lineal e invertible entre coordenadas homogéneas, es decir, de \mathbb{P}^2 a \mathbb{P}^2 . La geometría proyectiva se enfoca en caracterizar a estas transformaciones, también llamadas **homografías**. Observemos que cada homografía se puede escribir como una matriz 3×3 no singular, H.

De manera similar a los vectores homogéneos, las homografías $H \ge kH$, con k un número real distinto de 0, también son equivalentes. Sean m un punto en $\mathbb{P}^2 \ge H$ una homografía, tales que

$$m = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad y \quad H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}.$$

Luego

$$kHm = \begin{pmatrix} kh_{11} & kh_{12} & kh_{13} \\ kh_{21} & kh_{22} & kh_{23} \\ kh_{31} & kh_{32} & kh_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} kh_{11}x + kh_{12}y + kh_{13}z \\ kh_{21}x + kh_{22}y + kh_{23}z \\ kh_{31}x + kh_{32}y + kh_{33}z \end{pmatrix}$$
$$= \begin{pmatrix} k(h_{11}x + h_{12}y + h_{13}z) \\ k(h_{21}x + h_{22}y + h_{23}z) \\ k(h_{31}x + h_{32}y + h_{33}z) \end{pmatrix}$$

y por ser vectores homogéneos

$$\equiv \begin{pmatrix} h_{11}x + h_{12}y + h_{13}z \\ h_{21}x + h_{22}y + h_{23}z \\ h_{31}x + h_{32}y + h_{33}z \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

=Hm.

Como era de esperarse, para cada homografía H, también elegiremos una matriz representante entre las equivalentes a ella. Nosotros elegimos a la matriz kH tal que

$$k = \frac{1}{\sqrt[3]{|H|}}$$

pues de esa manera

$$|kH| = \begin{vmatrix} kh_{11} & kh_{12} & kh_{13} \\ kh_{21} & kh_{22} & kh_{23} \\ kh_{31} & kh_{32} & kh_{33} \end{vmatrix} = k^3 \begin{vmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{vmatrix}$$

$$= \left(\frac{1}{\sqrt[3]{|H|}}\right)^3 |H| = \frac{1}{|H|}|H| = 1.$$

Así, identificaremos al espacio de las homografías con el de las matrices 3×3 y determinante igual a 1. Este conjunto se conoce como el grupo lineal

especial de orden 3 y se denota por $SL_3(\mathbb{R})$.

Finalmente, utilizaremos un resultado importante en visión por computadora. Tomemos dos imágenes, $I_1 \in I_2$, de un objeto plano tomadas desde distintas perspectivas. Sean $p_1 = (x_1, y_1)^t$ y $p_2 = (x_2, y_2)^t$ las coordenadas en $I_1 \in I_2$, respectivamente, de un mismo punto sobre el objeto. Por ejemplo, la esquina inferior derecha de la portada de un libro, como se muestra en la figura 1.5. Se cumple que las representaciones homogéneas, $p'_1 = (x_1, y_1, 1)^t$ y $p'_2 = (x_2, y_2, 1)^t$, de p_1 y p_2 , respectivamente, están relacionadas por una *única* homografía en $SL_3(\mathbb{R})$, es decir, existe una única transformación lineal H en $SL_3(\mathbb{R})$ que va de \mathbb{P}^2 a \mathbb{P}^2 tal que para cualquier par (p'_1, p'_2) se cumple que $Hp'_1 = kp'_2$, donde k es el número real que utilizamos para asegurar que p'_2 tiene la representación homogénea con tercera componente igual a 1.



Figura 1.5: $p_1 \ge p_2$ son las coordenadas de la esquina inferior derecha de la portada del libro en las imágenes $I_1 \in I_2$ respectivamente.

1.2. Estadística

En esta sección daremos las definiciones más importantes para poder hablar de distribución de probabilidad y muestra aleatoria. Con esto describiremos medidas de tendencia central y de dispersión en \mathbb{R} . Recordemos que nosotros no trabajaremos en \mathbb{R} ni en \mathbb{R}^n , por ello estas definiciones sólo servirán de base para explicar nuestras nociones de media y varianza muestral en $SL_3(\mathbb{R})$.

Para mantener la exposición del tema tan simple como sea posible, los conceptos que no son tan utilizados en el desarrollo de esta tesis no tendrán definiciones rigurosas, pero eso no significa que no sean importantes. Al contrario, algunas de estas definiciones abren paso a nuevas investigaciones. En el capítulo 4, en la sección 4.3, se menciona un poco de estos conceptos y dónde encontrar más información al respecto.

Definición 1.1. Dado un experimento aleatorio, si Ω representa el conjunto de todos los resultados posibles de este experimento, una **variable aleatoria** es una función $X : \Omega \to \mathbb{R}$.

Utilizamos los valores que toman las variables aleatorias para definir probabilidades. Por ejemplo, si lanzamos una moneda, una variable aleatoria Xse puede definir como X = 1 si cae águila ó X = 0 si cae sol. Así, la probabilidad de que obtengamos águila la denotaremos por P(X = 1). Para este ejemplo, si la moneda es honesta, $P(X = 0) = P(X = 1) = \frac{1}{2}$.

Para experimentos aleatorios con dimensión mayor a 1, se define a un vector aleatorio.

Definición 1.2. Una función $\underline{X} = (X_1, ..., X_m)^t : \Omega \to \mathbb{R}^m$ es un vector aleatorio si cada coordenada $X_1, ..., X_m$ es una variable aleatoria.

La siguiente definición es útil para referirnos a la distribución de una variable aleatoria.

Definición 1.3. Sea X una variable aleatoria. Si existe una función f tal que $f : \mathbb{R} \to \mathbb{R}$ y satisface las siguientes condiciones:

a) $f(x) \ge 0$ para cada x en \mathbb{R} ,

c) Para cualesquiera $a, b \in \mathbb{R}$ tales que $-\infty < a < b < +\infty$ se cumple que $P(a \le X \le b) = \int_a^b f(x) dx$

decimos que f es una función de densidad de probabilidad de X.

Definición 1.4. Sea X una variable aleatoria, decimos que X tiene una **distribución normal** o Gaussiana, y lo denotaremos por $X \sim N(\mu, \sigma^2)$, si su función de densidad de probabilidad $f : \mathbb{R} \to \mathbb{R}$ es:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

donde $\mu \in \mathbb{R}$ y $\sigma \in \mathbb{R}_+$.

Los parámetros de esta distribución reflejan 2 características importantes, pues, el valor esperado y la varianza son respectivamente μ y σ^2 . La gráfica de la función de densidad de probabilidad de una variable aleatoria normal y el significado geométrico de los parámetros los podemos ver en la figura 1.6 (extraída de [17]).



Figura 1.6: f(x) es la función de densidad de probabilidad de $X \sim N(\mu, \sigma^2)$ (extraída de [17]).

La distribución normal es una distribución con propiedades importantes en estadística, muy utilizada en la modelación estadística para problemas de matemáticas aplicadas o problemas de ingeniería, en particular por el teorema del límite central. Por esta razón nosotros buscaremos definir una función similar a ella para los datos que trabajaremos. Más aún, como nosotros trabajaremos en un espacio de dimensión mayor a 1, la función será más parecida a una distribución normal multivariada.

Definición 1.5. Sea $\underline{X} = (X_1, ..., X_m)^t$ un vector aleatorio, decimos que \underline{X} tiene una **distribución normal multivariada**, la cual denotaremos por $\underline{X} \sim N(\mu, \Sigma)$, si su función de densidad de probabilidad $f : \mathbb{R}^m \to \mathbb{R}$ es:

$$f(\underline{x}) = \frac{1}{(2\pi)^{m/2}\sqrt{|\Sigma|}} \exp\left[-\frac{1}{2}(\underline{x}-\underline{\mu})^t \Sigma^{-1}(\underline{x}-\underline{\mu})\right]$$

donde $\underline{x} = (x_1, ..., x_m)^t \ y \ \underline{\mu} = (\mu_1, ..., \mu_m)^t$ son vectores de números reales, Σ es una matriz positiva definida, es decir, $\underline{x}^t \Sigma \underline{x} > 0$ para cada vector no nulo $\underline{x} \in \mathbb{R}^m$.

En este caso los parámetros $\underline{\mu}$ y Σ representan el vector de medias y la matriz de covarianza, respectivamente. Para $m = 2 \operatorname{con} \underline{\mu}$ y Σ adecuadas, obtenemos la **distribución normal bivariada estándar** cuya gráfica se encuentra en la figura 1.7 (extraída de [17]).



Figura 1.7: f(x, y) es la función de densidad de probabilidad normal bivariada estándar (extraída de [17]).

Para las siguientes definiciones podemos pensar que tenemos una población específica de objetos, y de ellos queremos hacer inferencia sin mirar cada objeto. Es decir, muestreamos [14]. **Definición 1.6.** Una muestra aleatoria es una colección $X_1, ..., X_n$ de variables aleatorias independientes e idénticamente distribuidas. Al número n se le conoce como el tamaño de la muestra. Y denotamos por $x_1, ..., x_n$ a los valores tomados por la muestra.

Esta definición establece que la muestra aleatoria corresponde a n mediciones repetidas de X, hechas básicamente bajo las mismas condiciones [14].

Recordemos que nuestro interés es hacer inferencia sobre algunos parámetros, utilizando la información obtenida de una muestra aleatoria. Por eso definimos lo que es una estadística.

Definición 1.7. Sea $X_1, ..., X_n$ una muestra aleatoria. Una **estadística**, g, es una variable aleatoria que es función de la muestra, $g : \mathbb{R}^n \to \mathbb{R}$ y que toma el valor de $g(x_1, ..., x_n)$.

Si bien podemos definir a una estadística de muchas maneras, hay estadísticas muy importantes, entre ellas, las que dan una medida de tendencia central (media, mediana) y las que dan una medida de dispersión (rango, varianza).

Definición 1.8. Sea $X_1, ..., X_n$ una muestra aleatoria. La **media muestral** es una estadística denotada por \overline{X} y definida como

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i.$$

Antes de definir a la mediana y al rango, definamos las estadísticas de orden.

Definición 1.9. Sea $X_1, ..., X_n$ una muestra aleatoria. A las variables aleatorias ordenadas

$$X_{(1)} = \min\{X_1, ..., X_n\}$$

$$X_{(2)} = \min(\{X_1, ..., X_n\} \setminus \{X_{(1)}\})$$

$$\vdots$$

$$X_{(n)} = \max\{X_1, ..., X_n\}$$

1.2. ESTADÍSTICA

se les conoce con el nombre de **estadísticas de orden**. El subíndice entre paréntesis significa que la variable aleatoria $X_{(i)}$ se encuentra en la posición *i*-ésima de la lista ordenada.

Definición 1.10. Sea $X_1, ..., X_n$ una muestra aleatoria. La **mediana mues**tral es una estadística que denotaremos por \tilde{X} y se define como

$$\tilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{si } n \text{ es impar} \\ \frac{1}{2} \left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)} \right) & \text{si } n \text{ es par} \end{cases}$$

Definición 1.11. Sea $X_1, ..., X_n$ una muestra aleatoria. El **rango muestral** es una estadística denotada por R y definida como

$$R = X_{(n)} - X_{(1)}$$

La varianza muestral de una muestra aleatoria es una medida del grado de dispersión de los diferentes valores tomados por la variable [17]. En el caso de un vector aleatorio, la medida de dispersión es descrita por una matriz llamada "matriz de covarianza", donde cada entrada representa la dispersión direccional de la "diferencia" de las muestras con la media [15].

Definición 1.12. Sea $X_1, ..., X_n$ una muestra aleatoria. La varianza muestral es una estadística denotada por S^2 y definida como

$$S^{2} = \frac{1}{n-1} \sum_{i=1}^{n} (X_{i} - \bar{X})^{2}.$$

Definición 1.13. Sea $(X_1, ..., X_m)$ un vector aleatorio. Tomamos una muestra aleatoria de n elementos respecto a cada una de las m entradas del vector. La **matriz de covarianza muestral** es una matriz $m \times m$ denotada por Σ en la que cada elemento σ_{ij} se define como

$$\sigma_{ij} = \frac{1}{n-1} \sum_{k=1}^{n} (X_{ik} - \bar{X}_i) (X_{jk} - \bar{X}_j)$$

donde X_{ik} y \bar{X}_i representan, respectivamente, la k-ésima medición y el promedio, correspondiente a la muestra de la variable aleatoria X_i . Análogamente X_{jk} y \bar{X}_j representan, respectivamente, la k-ésima medición y el promedio, correspondiente a la muestra de la variable aleatoria X_j .

Podemos deducir que la matriz Σ es simétrica. Al término σ_{ij} lo llamamos la **covarianza muestral** de las variables aleatorias X_i y X_j . Más aún, podemos definir de manera más concisa la **matriz de covarianza muestral** como

$$\Sigma = \frac{1}{n-1} \sum_{k=1}^{n} \left(\left[\begin{pmatrix} X_{1k} \\ \vdots \\ X_{mk} \end{pmatrix} - \begin{pmatrix} \bar{X}_{1} \\ \vdots \\ \bar{X}_{m} \end{pmatrix} \right] \left[\begin{pmatrix} X_{1k} \\ \vdots \\ X_{mk} \end{pmatrix} - \begin{pmatrix} \bar{X}_{1} \\ \vdots \\ \bar{X}_{m} \end{pmatrix} \right]^{t} \right).$$

Capítulo 2

Motivación al análisis estadístico sobre variedades

Sea $\underline{X} = (X_1, ..., X_m)$ un vector aleatorio. Sabemos que si $\underline{X} \sim N(\underline{\mu}, \Sigma)$, donde $\underline{\mu} \in \mathbb{R}^m$ es el vector de medias y $\Sigma \in \mathbb{R}^{m \times m}$ es la matriz de covarianza, cuando \underline{X} es transladada por $\underline{t} \in \mathbb{R}^m$, entonces sólamente el vector de medias es transladado, es decir, $(\underline{X} + \underline{t}) \sim N(\mu + \underline{t}, \Sigma)$.

Ahora, sean n imágenes de un objeto plano tomadas desde distintas posiciones. Sea Ref_1 una imagen de referencia del objeto. Calculemos las homografías $H_1, ..., H_n$ que transforman a cada una de las n imágenes en Ref_1 . Luego, con los algoritmos que describiremos, encontramos la homografía "media", $H_{(\mu_1)}$, que transforma la "vista media" del objeto en Ref_1 , al tomar la "transformación media" entre las diferentes H_i . Sea Ref_2 otra imagen de referencia. Si $H_{(2,1)}$ es la homogafía que transforma a Ref_1 en Ref_2 , entonces las homografías que transforman a cada una de las n imágenes en Ref_2 serán $H_{(2,1)}H_1, ..., H_{(2,1)}H_n$. Nosotros queremos que la homografía "media", $H_{(\mu_2)}$, transforme la "vista media" del objeto en Ref_2 . Para que esto ocurra, la propiedad de invarianza que necesitamos es que la media $H_{(\mu_2)}$ se obtenga como composición de la homografía media $H_{(\mu_1)}$ y de la que transforma a Ref_1 en Ref_2 , es decir, buscamos que sólo se "translade" la media haciendo $H_{(\mu_2)} = H_{(2,1)}H_{(\mu_1)}$.

En la figura 2.1 mostramos un ejemplo con un conjunto de 6 fotografías de una cartera (extraídas en parte de [1]), tomadas desde distintas posiciones. En el centro está la imagen correspondiente a la "vista media".



Figura 2.1: Las homografías $H_1, ..., H_6$ las encontramos a partir de Ref_1 . Queremos que $H_{(\mu_2)} = H_{(2,1)}H_{(\mu_1)}$ (extraídas en parte de [1]).

Recordemos que aunque utilizamos imágenes de distintas vistas de un objeto plano, y queremos encontrar su "vista media", nosotros trabajaremos con las homografías subyacentes, es decir, matrices 3×3 con determinante igual a 1. Podríamos intentar definir una función de densidad de probabilidad en el grupo tratándolo como un subconjunto de \mathbb{R}^9 ; sin embargo, sólo podríamos tomar 8 elementos, pues uno de ellos estaría en términos del determinante de la matriz, esto para asegurar que el determinante es igual a 1.

De ahí que si seguimos este método, no sería posible mantener la propiedad de invarianza que acabamos de mencionar. Por eso, utilizamos el hecho de que $SL_3(\mathbb{R})$ además de ser un grupo, también es una variedad diferenciable.

Definiremos una función $\phi : SL_3(\mathbb{R}) \to \mathbb{R}$ "similar" a una función de densidad de probabilidad normal multivariada utilizando propiedades intrínsecas de este espacio. Dada una media μ y una matriz de covarianza Σ definiremos a ϕ buscando que se mantenga la propiedad de invarianza descrita anteriormente.

Capítulo 3

Respaldo matemático: Geometría diferencial

En este capítulo presentaremos los elementos necesarios para describir a $SL_3(\mathbb{R})$ como un grupo, variedad diferenciable y grupo de Lie. Además, se exhibirán otros espacios y funciones que serán la base para desarrollar los algoritmos de cálculo de medidas de tendencia central y de dispersión.

3.1. Propiedades de $SL_3(\mathbb{R})$

Comencemos definiendo a un grupo y probemos que $SL_3(\mathbb{R})$ lo es.

Definición 3.1. Sea G un conjunto distinto del vacío y * una operación binaria definida en él. Decimos que G es un **grupo** si se cumple que

- Para cada $x, y, z \in G$: (x * y) * z = x * (y * z). [Ley asociativa]
- Existe $e \in G$ tal que para cada $x \in G$: x * e = e * x = x. [Existencia del elemento neutro]
- Para cada $x \in G$ existe $x^{-1} \in G$ tal que: $x * x^{-1} = x^{-1} * x = e$. [Existencia de los inversos]

Proposición 3.2. $SL_3(\mathbb{R})$ es un grupo respecto al producto de matrices.

Demostración. Sabemos que en el conjunto de matrices el producto es asociativo. También sabemos que la matriz

$$I_3 = \left(\begin{array}{rrr} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$$

cumple que det $(I_3) = 1$, es decir $I_3 \in SL_3(\mathbb{R})$ y para cualquier $g' \in SL_3(\mathbb{R})$, $g'I_3 = I_3g' = g'$. Finalmente, sea $g \in SL_3(\mathbb{R})$. Como det $(g) = 1 \neq 0$ entonces existe una matriz g^{-1} tal que $g * g^{-1} = I_3$, así det $(g * g^{-1}) = 1$ de ahí que det(g)det $(g^{-1}) = 1$, donde obtenemos que det $(g^{-1}) = 1$ y por lo tanto el inverso de g pertenece a $SL_3(\mathbb{R})$.

Definamos ahora a una variedad diferenciable.

Definición 3.3. Sea $X \subseteq \mathbb{R}^n$. Decimos que X es una variedad diferenciable de dimensión $k \leq n$ si es localmente difeomorfa a \mathbb{R}^k . Es decir, si para cada $x \in X$, existe un abierto V en X que contiene a x, un abierto $U \subseteq \mathbb{R}^k$ y cumplen que existe una función invertible $\phi : U \to V$ tal que ϕ y ϕ^{-1} son suaves.

Llamamos parametrización local de la variedad diferenciable a la función $\phi: U \to V$ y sistema coordenado de la variedad diferenciable a la función $\phi^{-1}: V \to U$. Cuando mencionemos que la dimensión de X es igual a k lo denotaremos por dim X = k. Y de aquí en adelante utilizaremos indistintamente los términos "variedad" y "variedad diferenciable".

En la figura 3.1 ilustramos una variedad, X, de 2 dimensiones inmersa en \mathbb{R}^3 . V es un abierto de X que contiene a x, U es un abierto de \mathbb{R}^2 que contiene a u y V está parametrizado por la función $\phi : U \to V$ y es tal que $\phi(u) = x$.

Antes de demostrar que $SL_3(\mathbb{R})$ es una variedad diferenciable, demos algunas definiciones más que nos servirán en la prueba.



Figura 3.1: V es un abierto en la variedad X y U es un abierto en \mathbb{R}^2 . $\phi: U \to V \neq \phi^{-1}: V \to U$.

Utilizamos las derivadas para identificar el espacio lineal que mejor se aproxima a una variedad X en el punto x.

Definición 3.4. Sea $X \subseteq \mathbb{R}^n$ una variedad de dimensión $k \leq n$. Dado $x \in X$ con una paramerización local $\phi : U \subseteq \mathbb{R}^k \to V \subseteq \mathbb{R}^n$ tal que $\phi(u) = x$. Si $D\phi_u : \mathbb{R}^k \to \mathbb{R}^n$ es la derivada de ϕ en u, definimos al **espacio tangente** a X en x como $T_x(X) := D\phi_u(\mathbb{R}^k) \subseteq \mathbb{R}^n$.

Recordemos que la derivada $D\phi_u: \mathbb{R}^k \to \mathbb{R}^n$ es una transformación lineal que es representada por la matriz

$$D\phi_u = \begin{pmatrix} \frac{\partial\phi_1}{\partial u_1} & \cdots & \frac{\partial\phi_1}{\partial u_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial\phi_n}{\partial u_1} & \cdots & \frac{\partial\phi_n}{\partial u_k} \end{pmatrix}$$

donde $\frac{\partial \phi_i}{\partial u_j}$ denota la derivada parcial de la *i*-ésima función de ϕ respecto a la *j*-ésima entrada de *u*, evaluada en *u*.

Aunque parece haber una ambigüedad en la definición (porque la definición podría parecer específica a una parametrización local), como se puede leer en [9], distintas parametrizaciones locales de $x \in X$ producen el mismo espacio tangente $T_x(X)$. Además, $T_x(X)$ es un subespacio vectorial de \mathbb{R}^n de dimensión de k, y la translación paralela $x + T_x(X) = \{y \in \mathbb{R}^n | y = x + z \text{ con } z \in T_x(X)\}$ es la aproximación por un subespacio de dimensión kmás cercana a X que pasa por x (es el "plano tangente" a la variedad). En la figura 3.2 ilustramos la translación del espacio tangente a una variedad 2-dimensional X en el punto x.



Figura 3.2: $x + T_x(X)$ es la translación del espacio tangente a X en x.

Supongamos que X y Y son dos variedades de dimensiones k y p respectivamente, $f : X \to Y$ es una función suave. Supongamos también que X y Y están parametrizadas respectivamente por $\phi : U \subseteq \mathbb{R}^k \to X$ y $\psi : V \subseteq \mathbb{R}^p \to Y$ tales que $\phi(0) = x, \psi(0) = y$ y f(x) = y. Notemos que si U es suficientemente pequeño podemos definir una función suave $h = \psi^{-1} \circ f \circ \phi$, que mapea de U a V. En la figura 3.3 se ilustran estas funciones.



Figura 3.3: X y Y son variedades parametrizadas por ϕ y ψ respectivamente. Para un abierto $U \subseteq \mathbb{R}^k$ suficientemente pequeño existe una función h tal que $h: U \to V$.

Luego, utilizando la regla de la cadena definimos

$$Df_x = D\psi_0 \circ Dh_0 \circ D\phi_0^{-1}$$

que mapea de $T_x(X)$ a $T_y(Y)$ como se ilustra en la figura 3.4 [9].



Figura 3.4: Df_x es una función que mapea de $T_x(X)$ a $T_y(Y)$.

Definición 3.5. Sean $X, Y \subseteq \mathbb{R}^n$ variedades diferenciables. Para una transformación suave entre variedades $f : X \to Y$, un punto $y \in Y$ es un **valor regular** de f si, en cada punto $x \in X$ tal que f(x) = y, $Df_x : T_x(X) \to T_y(Y)$ es suprayectiva.

Si dim X = p y dim Y = q. La aplicación $Df_x : T_x(X) \to T_y(Y)$ es una transformación lineal que es representada por la matriz

$$Df_x = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_q}{\partial x_1} & \cdots & \frac{\partial f_q}{\partial x_p} \end{pmatrix}$$

donde $\frac{\partial f_i}{\partial x_j}$ denota la derivada parcial de la *i*-ésima función de f respecto a la j-ésima entrada de x, evaluada en x.

Teorema 3.6. [Teorema de la imagen inversa] Sean $X, Y \subseteq \mathbb{R}^n$ variedades diferenciables. Si y es un valor regular de $f : X \to Y$, entonces la imagen inversa $f^{-1}(y)$ es una variedad diferenciable contenida en X, con dim $f^{-1}(y) = \dim X - \dim Y$.

Teorema 3.7. Sean $X, Y \subseteq \mathbb{R}^n$ variedades diferenciables. Sea Z la imagen inversa de un valor regular $y \in Y$ bajo la transformación suave $f : X \to Y$. Entonces el núcleo de la derivada $Df_x : T_x(X) \to T_y(Y)$ en cualquier punto $x \in Z$ es el espacio tangente a Z, $T_x(Z)$. Las demostraciones de los teoremas 3.6 y 3.7 se encuentran en la sección 4 del capítulo I de [9].

Para demostrar que $SL_3(\mathbb{R})$ es una variedad diferenciable, primero hagamos 2 observaciones.

Observación 3.8. Cada vector en \mathbb{R}^9 se puede asociar de manera única con una matriz 3×3 de la siguiente forma:

$$\left(\begin{array}{c}a_1\\\vdots\\a_9\end{array}\right)\leftrightarrow \left(\begin{array}{c}a_1&a_4&a_7\\a_2&a_5&a_8\\a_3&a_6&a_9\end{array}\right).$$

Observación 3.9. Sea $n \in \mathbb{N}$ $y \ u \in \mathbb{R}^n$. La función $\phi_u : \mathbb{R}^n \to \mathbb{R}^n$ tal que $\phi_u(x) = x$ es un difeomorfismo de \mathbb{R}^n a \mathbb{R}^n , por lo tanto, \mathbb{R}^n es una variedad diferenciable. Además, se puede ver que $D_{\phi_u} = I_n$. Más aún, $T_u(\mathbb{R}^n) = D_{\phi_u}(\mathbb{R}^n) = \mathbb{R}^n$, es decir, para cualquier punto u en \mathbb{R}^n el espacio tangente a \mathbb{R}^n en u es el mismo \mathbb{R}^n .

Ahora, utilizando todo lo mencionado anteriormente ya podemos demostrar la siguiente proposición.

Proposición 3.10. $SL_3(\mathbb{R})$ es una variedad diferenciable de dimensión 8.

Demostración. Sea $f : \mathbb{R}^9 \to \mathbb{R}$ tal que, si $a = \begin{pmatrix} a_1 \\ \vdots \\ a_9 \end{pmatrix}$, $f(a) = \begin{vmatrix} a_1 & a_4 & a_7 \\ a_2 & a_5 & a_8 \\ a_3 & a_6 & a_9 \end{vmatrix}.$

Veamos que el número real 1 es un valor regular de f. Sea x en \mathbb{R}^9 tal que f(x) = 1. Es decir, x se puede asociar a una matriz 3×3 con determinante igual a 1 dada la observación 3.8. Su derivada, Df_x , mapea de \mathbb{R}^9 a \mathbb{R} , viéndolos como espacios tangentes por lo descrito en la observación 3.9. Luego, viendo a la función f como un determinante y desarrollandolo por columnas tenemos:
$$f(x) = +x_1 \begin{vmatrix} x_5 & x_8 \\ x_6 & x_9 \end{vmatrix} - x_2 \begin{vmatrix} x_4 & x_7 \\ x_6 & x_9 \end{vmatrix} + x_3 \begin{vmatrix} x_4 & x_7 \\ x_5 & x_8 \end{vmatrix}$$
$$= -x_4 \begin{vmatrix} x_2 & x_8 \\ x_3 & x_9 \end{vmatrix} + x_5 \begin{vmatrix} x_1 & x_7 \\ x_3 & x_9 \end{vmatrix} - x_6 \begin{vmatrix} x_1 & x_7 \\ x_2 & x_8 \end{vmatrix}$$
$$= +x_7 \begin{vmatrix} x_2 & x_5 \\ x_3 & x_6 \end{vmatrix} - x_8 \begin{vmatrix} x_1 & x_4 \\ x_3 & x_6 \end{vmatrix} + x_9 \begin{vmatrix} x_1 & x_4 \\ x_2 & x_5 \end{vmatrix}$$

De ahí es claro que Df_x es la transformación lineal representada por la matriz

$$Df_x = \left((-1)^{1+1} \left| \begin{array}{cc} x_5 & x_8 \\ x_6 & x_9 \end{array} \right|, \cdots, (-1)^{9+1} \left| \begin{array}{cc} x_1 & x_4 \\ x_2 & x_5 \end{array} \right| \right).$$

Es decir, la entrada *i*-ésima de Df_x es $(-1)^{i+1}$ multiplicada por el determinante de la matriz que se obtiene de x eliminando el renglón y la fila correspondiente al *i*-ésimo término de x (representando a x como matriz).

Verifiquemos que $Df_x : \mathbb{R}^9 \to \mathbb{R}$ es suprayectiva. Sea y en \mathbb{R} , probemos que existe un z en \mathbb{R}^9 tal que $Df_x(z) = y$.

Proponemos a
$$z = \frac{y}{3}x = \begin{pmatrix} \frac{yx_1}{3} \\ \vdots \\ \frac{yx_9}{3} \end{pmatrix}$$
, así

$$Df_{x}(z) = \left((-1)^{1+1} \begin{vmatrix} x_{5} & x_{8} \\ x_{6} & x_{9} \end{vmatrix} , \cdots, (-1)^{9+1} \begin{vmatrix} x_{1} & x_{4} \\ x_{2} & x_{5} \end{vmatrix} \right) \left(\begin{pmatrix} \frac{yx_{1}}{3} \\ \vdots \\ \frac{yx_{9}}{3} \end{pmatrix}$$
$$= \frac{yx_{1}}{3} (-1)^{1+1} \begin{vmatrix} x_{5} & x_{8} \\ x_{6} & x_{9} \end{vmatrix} + \cdots + \frac{yx_{3}}{3} (-1)^{3+1} \begin{vmatrix} x_{4} & x_{7} \\ x_{5} & x_{8} \end{vmatrix}$$
$$+ \frac{yx_{4}}{3} (-1)^{4+1} \begin{vmatrix} x_{2} & x_{8} \\ x_{3} & x_{9} \end{vmatrix} + \cdots + \frac{yx_{6}}{3} (-1)^{6+1} \begin{vmatrix} x_{1} & x_{7} \\ x_{2} & x_{8} \end{vmatrix}$$
$$+ \frac{yx_{7}}{3} (-1)^{7+1} \begin{vmatrix} x_{2} & x_{5} \\ x_{3} & x_{6} \end{vmatrix} + \cdots + \frac{yx_{9}}{3} (-1)^{9+1} \begin{vmatrix} x_{1} & x_{4} \\ x_{2} & x_{5} \end{vmatrix}$$
$$= \frac{y}{3} \begin{vmatrix} x_{1} & x_{4} & x_{7} \\ x_{2} & x_{5} & x_{8} \\ x_{3} & x_{6} & x_{9} \end{vmatrix} + \frac{y}{3} \begin{vmatrix} x_{1} & x_{4} & x_{7} \\ x_{2} & x_{5} & x_{8} \\ x_{3} & x_{6} & x_{9} \end{vmatrix} + \frac{y}{3} \begin{vmatrix} x_{1} & x_{4} & x_{7} \\ x_{2} & x_{5} & x_{8} \\ x_{3} & x_{6} & x_{9} \end{vmatrix} + \frac{y}{3} \begin{vmatrix} x_{1} & x_{4} & x_{7} \\ x_{2} & x_{5} & x_{8} \\ x_{3} & x_{6} & x_{9} \end{vmatrix} + \frac{y}{3} \begin{vmatrix} x_{1} & x_{4} & x_{7} \\ x_{2} & x_{5} & x_{8} \\ x_{3} & x_{6} & x_{9} \end{vmatrix}$$
$$= \frac{y}{3} f(x) + \frac{y}{3} f(x) + \frac{y}{3} f(x) = 3 \left(\frac{y}{3} f(x)\right) = yf(x) = y.$$

Con esto probamos que Df_x es una función suprayectiva, lo cual implica que 1 es un valor regular de f. Entonces, por el teorema 3.6, $f^{-1}(1)$ es una variedad diferenciable contenida en \mathbb{R}^9 y dim $f^{-1}(1) = \dim \mathbb{R}^9 - \dim \mathbb{R} =$ 9-1=8. Es decir, por la observación 3.8, el conjunto de matrices 3×3 con determinante igual a 1 es una variedad de dimensión 8.

Notemos que esta variedad tiene estructura de grupo, de ahí que podemos encontrar el espacio tangente a $SL_3(\mathbb{R})$ en el elemento identidad, I_3 . A este conjunto lo denotaremos por \mathfrak{g} .

Proposición 3.11. \mathfrak{g} es el conjunto de matrices 3×3 cuya traza es cero.

Demostración. Sea $f : \mathbb{R}^9 \to \mathbb{R}$ definida como en la demostración de la proposición 3.10, recordemos que su derivada es

$$Df_x = \left((-1)^{1+1} \left| \begin{array}{cc} x_5 & x_8 \\ x_6 & x_9 \end{array} \right|, \cdots, (-1)^{9+1} \left| \begin{array}{cc} x_1 & x_4 \\ x_2 & x_5 \end{array} \right| \right).$$

Veamos qué es el núcleo de Df_{I_3} . Sea y en \mathbb{R}^9 tal que $Df_{I_3}(y) = 0$, entonces

$$\left((-1)^{1+1} \left| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right|, \cdots, (-1)^{9+1} \left| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right| \right) \left(\begin{array}{c} y_1 \\ \vdots \\ y_9 \end{array} \right) = 0$$

asi

$$1(y_1) + 0(y_2) + 0(y_3) + 0(y_4) + 1(y_5) + 0(y_6) + 0(y_7) + 0(y_8) + 1(y_9) = 0$$

por lo tanto, $y_1 + y_5 + y_9 = 0$, es decir, viendo a y como matriz por la observación 3.8, el núcleo de Df_{I_3} son las matrices con traza igual a cero.

Finalmente, por el teorema 3.7, el conjunto de matrices 3×3 con traza igual a cero es el espacio tangente a $SL_3(\mathbb{R})$ en I_3 , \mathfrak{g} .

Definición 3.12. Sea G una variedad diferenciable con estructura de grupo, decimos que G es un grupo de Lie si cumple que las funciones

$$f_1: G \times G \to G, f_1(x, y) = x * y$$
$$y f_2: G \to G, f_2(x) = x^{-1}$$

son suaves.

Proposición 3.13. $SL_3(\mathbb{R})$ es un grupo de Lie.

Demostración.

Sean $A, B, C, D \in SL_3(\mathbb{R})$ tales que AB = C y $D = A^{-1}$ donde

$$A = \left(\begin{array}{rrrr} a_1 & a_4 & a_7 \\ a_2 & a_5 & a_8 \\ a_3 & a_6 & a_9 \end{array}\right)$$

y B, C y D tienen su respectiva representación análoga.

Sean también $g_1 : \mathbb{R}^{18} \to \mathbb{R}^9$ tal que

$$g_1 \begin{pmatrix} a_1 \\ \vdots \\ a_9 \\ b_1 \\ \vdots \\ b_9 \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_9 \end{pmatrix},$$

 $g_2: \mathbb{R}^9 \to \mathbb{R}^9$ tal que

$$g_2 \begin{pmatrix} a_1 \\ \vdots \\ a_9 \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_9 \end{pmatrix}.$$

Observemos que la función g_1 es suave, pues cada elemento de C se forma como composición de suma y producto de números reales. Luego, empleando la observación 3.8, g_1 es equivalente a f_1 . Por lo tanto f_1 es una función suave.

También g_2 es suave, pues cada elemento de D se forma como un cociente con denominador igual a 1 y numerador dado por una composición de sumas y productos de números reales, relativos a la transpuesta de la matriz de cofactores. Luego, empleando la observación 3.8, g_2 es equivalente a f_2 . Por lo tanto f_2 también es una función suave.

Hasta aquí sólo hemos exhibido características intrínsecas de $SL_3(\mathbb{R})$.

3.2. $SL_3(\mathbb{R})$ con estructura métrica

Ahora que hemos caracterizado a $SL_3(\mathbb{R})$ como grupo de Lie, definiremos una estructura métrica sobre él. Esta métrica será útil para definir a la función ϕ mencionada en la introducción.

Para cada punto g en el grupo, definiremos un producto interior \langle , \rangle_g sobre el espacio tangente $T_g(SL_3(\mathbb{R}))$, luego, inducimos la norma $||v||_g = \sqrt{\langle v, v \rangle_g}$. Esto nos permitirá medir la longitud de un vector tangente, y será útil para medir la distancia entre 2 puntos en $SL_3(\mathbb{R})$. Llamaremos variedad riemanniana a una variedad en la que podamos definir un producto interior en cada espacio tangente, y con ello, definir una distancia sobre la variedad. $SL_3(\mathbb{R})$ es un ejemplo de una variedad riemanniana [1]. De aquí en adelante denotaremos por T_g a $T_g(SL_3(\mathbb{R}))$.

3.2.1. Aplicación exponencial

Definimos la longitud de una trayectoria $\sigma : [a, b] \to SL_3(\mathbb{R})$ como

$$L(\sigma) = \int_{a}^{b} ||\sigma'(t)||_{\sigma(t)} dt.$$

Si $t_0 \in [a, b]$, $\sigma(t_0)$ es un punto en $SL_3(\mathbb{R})$ y $\sigma'(t_0)$ es la derivada de la trayectoria σ evaluada en t_0 , entonces $\sigma'(t_0)$ es un vector en el espacio tangente a $\sigma(t_0)$. Por ello el subíndice de la norma, dentro de la integral, es $\sigma(t)$.

Finalmente, definimos la distancia entre los puntos g_1 y g_2 en $SL_3(\mathbb{R})$ como

 $d(g_1, g_2) = \inf \{ L(\sigma) : \sigma \text{ es una trayectoria que une a } g_1 \text{ con } g_2 \}.$

Una trayectoria σ^* donde se alcanza el ínfimo se llama geodésica. Y para cada $g_1 y g_2$ en $SL_3(\mathbb{R})$ se puede probar que existe una única geodésica que va de g_1 a g_2 [1]. Este hecho permite construir, para cada g en el grupo, la aplicación exponencial $\exp_g : T_g \to SL_3(\mathbb{R})$ tal que para cada v en $T_g, d(g, \exp_g(v)) = ||v||_g$. En la figura 3.5 ilustramos un ejemplo en el que $d(g, g') = ||v||_g$ donde v está en T_g y es tal que $\exp_g(v) = g'$.

La aplicación exp_g no siempre será biyectiva, pero en el conjunto donde sí lo sea (en general esto sucede en una vecindad de g), denotaremos a su inversa por \log_g . Así, también podremos decir que si $d(g, g' = \exp_g(v)) = ||v||_g$



Figura 3.5: La trayectoria marcada de g a g' es la geodésica. $||v||_g$ es la distancia entre $g \neq g'$, donde $g' = \exp_q(v)$.

entonces $v = \log_g(g')$ y $d(g, g') = ||\log_g(g')||_g$. Haremos una distinción con la aplicación exponencial bajo la identidad, \exp_{I_3} , y con su inversa \log_{I_3} , y las denotaremos por *expp* y $\log g$ respectivamente. También denotaremos al producto interior sobre el plano tangente a $SL_3(\mathbb{R})$ en I_3 , \mathfrak{g} , como \langle , \rangle_e y a su norma como $|| \cdot ||_e$. Hablaremos de ellos más adelante.

3.2.2. Propiedad de invarianza

Para que la función ϕ mantenga la propiedad de invarianza descrita en el capítulo 2, la métrica debe cumplir una propiedad importante.

Sea $g \in SL_3(\mathbb{R})$, definimos la función de translación por izquierda *(left translation by g)* $L_g : SL_3(\mathbb{R}) \to SL_3(\mathbb{R})$ tal que $L_g(g') = gg'$. De la definición de grupo de Lie, L_g es un difeomorfismo dentro del grupo, y además, para cada g en el grupo, $(L_g)^{-1} = L_{g^{-1}}$ [20].

Para que se mantenga la propiedad de invarianza, se debe cumplir que la distancia entre 2 puntos debe ser la misma sin importar una translación por izquierda. En particular queremos que si $g, g' \in SL_3(\mathbb{R})$ entonces

$$d(g,g') = d\left(L_{q^{-1}}(g), L_{q^{-1}}(g')\right).$$
(3.1)

En el capítulo 4 se verifica que esta igualdad es suficiente para mantener

la propidedad de invarianza.

Definamos funciones entre dos espacios tangentes a la variedad, específicamente, entre \mathfrak{g} y T_q .

Sea $L_q^* : \mathfrak{g} \to T_g$ tal que

$$L_g^*(v) = \log_g(g \cdot expp(v)).$$

Gráficamente representa que si $v \in \mathfrak{g}$, primero se mapea a la variedad, cerca de la identidad, con la función expp, luego se translada cerca de g con la función L_g y finalmente se mapea a T_g con la función \log_g . Esta función es invertible localmente, su inversa $(L_g^*)^{-1}$ va de T_g a \mathfrak{g} y es tal que

$$(L_g^*)^{-1}(u) = \log g(g^{-1} \cdot \exp_g(u)) = L_{g^{-1}}^*(u).$$

De manera análoga, para un punto $u \in T_g$, primero se mapea a la variedad, cerca de g, con la función \exp_g , luego se translada cerca de la identidad con la función $L_{q^{-1}}$ y finalmente se mapea a \mathfrak{g} con la función $\log g$.

Ya que exhibimos a L_g^* y a $L_{g^{-1}}^*$, veamos que si definimos el producto interior \langle , \rangle_g , en T_g , en función de $L_{g^{-1}}^*$ y \langle , \rangle_e obtendremos la igualdad deseada 3.1.

Proposición 3.14. Sea $u \in T_g$, definiendo $\langle u, u \rangle_g := \langle L_{g^{-1}}^*(u), L_{g^{-1}}^*(u) \rangle_e$ se cumple que $d(g, g') = d(L_{g^{-1}}(g), L_{g^{-1}}(g')).$

Demostración. Sean $g, g' \in SL_3(\mathbb{R})$

$$\begin{split} (d(g,g'))^2 &= (||\log_g(g')||_g)^2 \\ &= < \log_g(g'), \log_g(g') >_g \\ &= < L_{g^{-1}}^*(\log_g(g')), L_{g^{-1}}^*(\log_g(g')) >_e \\ &= < \log g(g^{-1} \cdot \exp_g(\log_g(g'))), \log g(g^{-1} \cdot \exp_g(\log_g(g'))) >_e \\ &= < \log g(g^{-1} \cdot g'), \log g(g^{-1} \cdot g') >_e \\ &= (||\log g(g^{-1}g')||_e)^2 = (d(I_3, g^{-1}g'))^2 \\ &= (d(g^{-1}g, g^{-1}g'))^2 = (d(L_{g^{-1}}(g), L_{g^{-1}}(g')))^2. \end{split}$$

Es decir $(d(g,g'))^2 = (d(L_{g^{-1}}(g), L_{g^{-1}}(g')))^2$, por lo tanto $d(g,g') = d(L_{g^{-1}}(g), L_{g^{-1}}(g'))$.

Esta demostración también nos dice cómo calcular la distancia entre dos elementos cualesquiera en $SL_3(\mathbb{R})$, $g \ge g'$, utilizando solamente las funciones $logg \ge || \cdot ||_e = \sqrt{\langle \cdot \rangle_e}$, definidas en el plano tangente al grupo en el elemento identidad, \mathfrak{g} . Es claro que

$$d(g,g') = ||logg(g^{-1}g')||_e.$$

Más aún, definiendo así a la distancia entre $g \ge g'$ se cumple que d(g, g') = d(Hg, Hg') para cualquier homografía H, pues

$$g^{-1}g' = g^{-1}H^{-1}Hg' = (Hg)^{-1}(Hg').$$

En la figura 3.6 ilustramos todas las funciones mencionadas anteriormente: \exp_g , expp, \log_g , $\log g$, L_g , $L_{g^{-1}}$, L_g^* y $L_{g^{-1}}^*$. Para $v \in \mathfrak{g}$ y $u \in T_g$ tales que $expp(v) = g^{-1}g'$ y $\exp_g(u) = g'$ se cumple que $L_g^*(v) = u$ y $L_{g^{-1}}^*(u) = v$.



Figura 3.6: En este diagrama se ilustran las funciones involucradas entre la variedad diferenciable $SL_3(\mathbb{R})$ y los planos tangentes \mathfrak{g} y T_g .

Debido a que el producto interior en cada plano tangente T_g está definido en términos del que está definido en \mathfrak{g} , sólo necesitamos definir un producto interior \langle , \rangle_e adecuado. Sean v_1 y v_2 , en \mathfrak{g} , definimos $\langle v_1, v_2 \rangle_e = Tr(v_1v_2^t)$, es decir, multiplicar v_1 por v_2 transpuesta y luego obtener su traza. Así, la norma que se induce es $||v||_e = \sqrt{Tr(vv^t)}$, la cual se conoce como norma de Frobenius.

3.2.3. Definición de *expp* y *logg*

El producto interior definido como $\langle v_1, v_2 \rangle_e = Tr(v_1v_2^t)$, se conoce como la forma de Killing-Cartan. Este producto interior nos da una ventaja para calcular las distancias entre elementos del grupo, pues, las funciones *expp* y *logg* tienen una forma cerrada.

Se define a $expp : \mathfrak{g} \to SL_3(\mathbb{R})$ como

$$expp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

donde x^k representa la multiplicación por izquierda de la matriz x en k ocasiones. Esta definición es válida, pues la serie converge como se puede ver en el capítulo 4 de [5].

En varios grupos de matrices se tiene esta definición para la función que mapea de \mathfrak{g} al grupo [22]. De hecho, el nombre de "aplicación exponencial" proviene de que en estos grupos tiene un comportamiento similar a la función exponencial en los números reales.

La aplicación inversa, logg, va de una vecindad de la identidad en $SL_3(\mathbb{R})$ a \mathfrak{g} y es tal que

$$logg(g) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (g-I)^k.$$

La definición es válida porque la serie converge, la prueba también se encuentra en el capítulo 4 de [5]. Esta función tiene un comportamiento similar a la función logaritmo en los números reales.

Como podemos observar, ya tenemos todos los elementos necesarios para medir la distancia entre dos elementos de $SL_3(\mathbb{R})$, pues ya tenemos las formas explícitas de logg y $|| \cdot ||_e$ para calcular $||logg(g^{-1}g')||_e = d(g,g')$ con g y g'en $SL_3(\mathbb{R})$.

Capítulo 4

Análisis estadístico sobre $SL_3(\mathbb{R})$

En este capítulo describiremos cómo encontrar medidas de tendencia central y de dispersión de datos dentro de un grupo de Lie. Nuestro objetivo es definir una función $\phi : SL_3(\mathbb{R}) \to \mathbb{R}$ que se ajuste a una muestra aleatoria de datos en $SL_3(\mathbb{R})$.

4.1. Media

Para definir la media utilizamos la noción de centro de masa sobre una variedad riemanniana introducida por Cartan [4] [15]. En específico, utilizaremos la media de Fréchet [21].

Definición 4.1. Sean X una variedad riemanniana $y x_1, ..., x_n \in X$. Definimos la **media de Fréchet** como

$$\mu = \arg\min_{h \in X} \sum_{i=1}^{n} d(x_i, h)^2$$

donde $d(\cdot, \cdot)$ denota la distancia definida sobre la variedad.

Si trabajamos en \mathbb{R} , no es difícil probar que la media de Fréchet es justamente la media muestral mencionada en la definición 1.8. Si trabajamos en \mathbb{R}^2 la media de Fréchet se ilustrará como en la figura 4.1.



Figura 4.1: Los puntos negros representan una colección de datos en \mathbb{R}^2 y la x representa la media de Fréchet.

Utilizando la media de Fréchet podemos definir la media de $g_1, ..., g_n \in SL_3(\mathbb{R})$ como

$$\mu = \arg\min_{h \in SL_3(\mathbb{R})} \sum_{i=1}^n ||logg(g_i^{-1} \cdot h)||_e^2$$

Notemos que al ser un proceso de minimización, no aseguramos que haya un mínimo global. Sin embargo, en [1] se menciona que la unicidad de la media y la convergencia del algoritmo 5.2, donde describimos cómo encontrarla, se siguen de los resultados de [2].

Ahora que hemos definido la media de un conjunto de datos en $SL_3(\mathbb{R})$, podemos verificar la propiedad de invarianza descrita en el capítulo 2. Probemos que si a las homografías de la muestra se les aplica una translación, multiplicando por un g a la izquierda, la media es simplemente transladada por la misma cantidad.

Sea μ_g la media de los datos transladados por $g \in SL_3(\mathbb{R})$, entonces

$$\mu_g = \arg\min_{h \in SL_3(\mathbb{R})} \sum_{i=1}^n d(g.g_i, h)^2.$$

Luego, por la igualdad 3.1

$$\mu_g = \arg\min_{h \in SL_3(\mathbb{R})} \sum_{i=1}^n d(L_{g^{-1}}(g.g_i), L_{g^{-1}}(h))^2,$$

es decir,

$$\mu_g = \arg\min_{h \in SL_3(\mathbb{R})} \sum_{i=1}^n d(g_i, g^{-1}h)^2.$$

Además, utilizando la definición de media de $g_1, ..., g_n$, se deduce que

$$g^{-1}\mu_g = \mu,$$

y finalmente

$$\mu_g = g\mu.$$

Esto nos muestra que la media de los datos transladados es el transladado de la media.

Para calcular el valor de μ utilizaremos la aproximación de segundo orden propuesta en [15], de ahí que,

$$\hat{\mu} = expp(\frac{1}{n}\sum_{i=1}^{n} logg(g_i)).$$

4.1.1. Otras medidas de tendencia central

Recordemos que la media μ es sólo una transformación proyectiva que transforma una "vista media" a una imagen de referencia Ref. Sin embargo, casi nunca tendremos una imagen real (dentro de nuestra muestra) que sea exactamente la "vista media". Por esto, proponemos dos métodos originales para hallar una medida de tendencia central, utilizando las imágenes que vienen dadas por la muestra.

Definición 4.2. Sean X una variedad riemanniana $y x_1, ..., x_n \in X$. Definimos la **media**- α como

$$\mu_{\alpha} = \arg\min_{x_i \in \{x_1, \dots, x_n\}} d(x_i, \mu)$$

donde $d(\cdot, \cdot)$ denota la distancia definida sobre la variedad y μ la media de Fréchet de los datos.

Luego, la media- α de $g_1, ..., g_n \in SL_3(\mathbb{R})$ es

$$\mu_{\alpha} = \arg\min_{g_i \in \{g_1, \dots, g_n\}} ||logg(g_i^{-1} \cdot \mu)||_e.$$

Esta definición propone que después de encontrar la media μ , busquemos en el conjunto de datos el dato que se encuentre más cerca a ella. Al aplicar la transformación μ_{α}^{-1} a una imagen de referencia *Ref* obtendremos la imagen más cercana a la "vista media". La figura 2 es un ejemplo de esto.

También podemos definir una medida de tendencia central similar a la media de Fréchet, pero considerando sólo a los puntos de la muestra.

Definición 4.3. Sean X una variedad riemanniana $y x_1, ..., x_n \in X$. Definimos la **media**- β como

$$\mu_{\beta} = \arg\min_{x_i \in \{x_1, \dots, x_n\}} \sum_{j=1}^n d(x_i, x_j)^2$$

donde $d(\cdot, \cdot)$ denota la distancia definida sobre la variedad.

Así, la media- β de $g_1, ..., g_n \in SL_3(\mathbb{R})$ es

$$\mu_{\beta} = \arg\min_{g_i \in \{g_1, \dots, g_n\}} \sum_{j=1}^n ||logg(g_i^{-1} \cdot g_j)||_e^2.$$

En la figura 4.2 ilustramos la media de Fréchet, la media- α y la media- β con dos conjuntos de datos en \mathbb{R}^2 . En 4.2(b) podemos observar que en algunas ocasiones ocurrirá que $\mu_{\alpha} = \mu_{\beta}$.

4.2. Matriz de covarianza

Recordemos del capítulo 1 que en el caso de un vector aleatorio, la medida de dispersión es descrita por una matriz llamada "matriz de covarianza", donde cada entrada representa la dispersión direccional de la "diferencia" de



(a) Para estos datos la media- α y la (b) Para estos datos la media- α y la media- β son diferentes media- β son iguales

Figura 4.2: Los puntos negros representan una colección de datos en \mathbb{R}^2 , x representa la media de Fréchet, el punto con un segmento hacia la media representa la media- α y el punto con centro blanco representa la media- β .

vectores. Así, para mantener similitud con dicha matriz de covarianza, buscamos "diferencias" y "direcciones" que sean equivalentes a las de los términos $(X_{ik} - \bar{X}_i)$ de la definición 1.13.

Para calcular la matriz de covarianza en $SL_3(\mathbb{R})$ seguiremos la siguiente idea. Primero calcularemos μ , la media de los datos g_1, \ldots, g_n . Luego, mapearemos todos los datos a T_{μ} y finalmente definiremos "diferencias" y "direcciones" sobre T_{μ} para calcular ahí a la matriz de covarianza.

Ahora, como la distancia que definimos entre dos puntos en $SL_3(\mathbb{R})$ es la misma sin importar una translación por izquierda, en lugar de trabajar con $g_1, ..., g_n$, trabajaremos con $\mu^{-1}g_1, ..., \mu^{-1}g_n$. Luego, en lugar de mapear los datos a T_{μ} , mapearemos los datos a \mathfrak{g} con la función logg. Así, sólo nos queda definir las "diferencias" entre $logg(\mu^{-1}g_i)$ y $logg(\mu^{-1}\mu)$ que nos ayuden a calcular la dispersión en ciertas "direcciones" sobre el plano tangente a la variedad en I_3 .

Para hacer esto, primero elegimos una base ordenada del espacio vectorial \mathfrak{g} :

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Sabemos que como las matrices anteriores forman una base de \mathfrak{g} , entonces cada matriz en \mathfrak{g} puede ser expresada, de manera única, como una combinación lineal de los elementos de la base. Definimos la función $b : \mathfrak{g} \to \mathbb{R}^8$, tal que si $a \in \mathfrak{g}$, a se escribe como combinación lineal de los elementos de la base ordenada y los valores de b(a), que son los coeficientes de la combinación lineal. Es decir, si

$$a = \left(\begin{array}{rrrr} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array}\right)$$

entonces

$$b(a) = (a_{11}, a_{12}, a_{13}, a_{21}, a_{23}, a_{31}, a_{32}, -a_{33})^t,$$

pues $-a_{11} - a_{33} = a_{22}$.

Además, es claro que existe su función inversa, $b^{-1} : \mathbb{R}^8 \to \mathfrak{g}$, tal que cada elemento de \mathbb{R}^8 determina, en combinación lineal con los elementos de la base ordenada, una matriz única en \mathfrak{g} .

También sabemos que $logg(\mu^{-1}\mu) = logg(I_3) = 0_3$, donde 0_3 representa a la matriz 3×3 en la que cada elemento es 0, y es claro que $b(0_3) = (0, 0, 0, 0, 0, 0, 0, 0)^t$.

4.2. MATRIZ DE COVARIANZA

Luego, si definimos a las "direcciones" como las direcciones de los vectores de la base mencionada anteriormente, las "diferencias" de los vectores $logg(\mu^{-1}g_i)$ y $logg(\mu^{-1}\mu) = 0_3$ serán simplemente las proyecciones de cada $logg(\mu^{-1}g_i)$ con respecto a cada uno de los elementos de la base dada de g.

Finalmente, si denotamos por $b_k(v)$ al k-ésimo elemento del vector b(v). Podemos definir la matriz de covarianza de $g_1, ..., g_n \in SL_3(\mathbb{R})$ como una matriz 8×8 , Σ , en la que cada elemento σ_{ij} es tal que

$$\sigma_{ij} = \frac{1}{n} \sum_{k=1}^{n} b_i \left(logg(\mu^{-1}g_k) \right) \cdot b_j \left(logg(\mu^{-1}g_k) \right).$$

Podemos deducir que la matriz Σ es simétrica, pues $\sigma_{ij} = \sigma_{ji}$, similar a la matriz de covarianza descrita en la definición 1.13.

Ahora que hemos definido la matriz de covarianza de un conjunto de datos en $SL_3(\mathbb{R})$ podemos verificar por completo la propiedad de invarianza. Probemos que si a las homografías de la muestra se les aplica una translación, multiplicando por un g a la izquierda, la matriz de covarianza continúa siendo la misma.

Sea μ_g la media de los datos transladados por $g \in SL_3(\mathbb{R})$, entonces, para este conjunto de datos la matriz de covarianza se define como

$$\sigma_{ij} = \frac{1}{n} \sum_{k=1}^{n} b_i \left(logg(\mu_g^{-1}gg_k) \right) \cdot b_j \left(logg(\mu_g^{-1}gg_k) \right)$$

y como $\mu_g = g\mu$

$$= \frac{1}{n} \sum_{k=1}^{n} b_i \left(\log g((g\mu)^{-1}gg_k) \right) \cdot b_j \left(\log g((g\mu)^{-1}gg_k) \right)$$

$$= \frac{1}{n} \sum_{k=1}^{n} b_i \left(logg(\mu^{-1}g^{-1}gg_k) \right) \cdot b_j \left(logg(\mu^{-1}g^{-1}gg_k) \right)$$

$$= \frac{1}{n} \sum_{k=1}^n b_i \left(logg(\mu^{-1}g_k) \right) \cdot b_j \left(logg(\mu^{-1}g_k) \right).$$

Esto nos muestra que la matriz de covarianza de los datos originales y la de los datos transladados por g a la izquierda es la misma.

Si bien cada entrada de la matriz Σ ya es una medida de dispersión sobre las direcciones de la base de \mathfrak{g} , aún podemos hacer algo más para relacionar esta información con la muestra de imágenes dada al principio.

Definición 4.4. Dada una muestra aleatoria normal multivariada en \mathbb{R}^n , con vector de medias μ y matriz de covarianza Σ . Decimos que los puntos $x^{[0]}, x^{[1]}, ..., x^{[2n]}$ en \mathbb{R}^n son **puntos sigma** de la muestra si $x^{[0]} = \mu$ y

$$\begin{split} x^{[i]} &= \mu + (\sqrt{(n+k)\Sigma})_i & para \ i \in \{1,...,n\} \\ x^{[j]} &= \mu - (\sqrt{(n+k)\Sigma})_{j-n} & para \ j \in \{n+1,...,2n\} \end{split}$$

donde k es un parámetro de escalamiento y el término $(\sqrt{(n+k)\Sigma})_i$ representa la columna i-ésima de la raiz cuadrada de la matriz obtenida al multiplicar el escalar (n+k) con la matriz Σ . Análogamente para el término $(\sqrt{(n+k)\Sigma})_{j-n}$ [18].

En este contexto, utilizamos la siguiente definición de raiz cuadrada de una matriz.

Definición 4.5. Sean A y B matrices $n \times n$ con entradas reales. Decimos que la matriz A es la raiz cuadrada de la matriz B si $B = A \cdot A^t$.

4.2. MATRIZ DE COVARIANZA

Por ejemplo, de la misma manera que un conjunto de datos en \mathbb{R}^2 , tomados de una muestra normal bivariada, se ajusta a una elipse como se ilustra en la figura 4.3, un conjunto de datos en \mathbb{R}^n , tomados de una muestra normal multivariada, se ajusta a una elipsoide de n dimensiones. Es ahí donde toman importancia los puntos sigma.



Figura 4.3: Los puntos representan un conjunto de datos en \mathbb{R}^2 tomados de una muestra normal bivariada.

El punto sigma $x^{[0]}$, al ser la media de la muestra, representa el centro de una elipsoide de *n* dimensiones. Cada uno de los puntos $x^{[1]}, ..., x^{[n]}$ representa un vértice de la elipsoide, todos sobre un eje distinto. Y los puntos $x^{[n+1]}, ..., x^{[2n]}$ representan los vértices opuestos a $x^{[1]}, ..., x^{[n]}$ respectivamente.

Ahora, para mantener similitud con la definición 4.4, dada una muestra en $SL_3(\mathbb{R})$ con media μ y matriz de covarianza Σ , proponemos que los puntos $g^{[0]}, g^{[1]}, \dots, g^{[16]}$ en $SL_3(\mathbb{R})$ sean puntos sigma de la muestra si $g^{[0]} = \mu$ y

$$g^{[i]} = \mu \cdot expp\left(b^{-1}\left((\sqrt{(8+k)\Sigma})_i\right)\right) \qquad \text{para } i \in \{1, \dots, 8\}$$
$$g^{[j]} = \mu \cdot expp\left(b^{-1}\left(-(\sqrt{(8+k)\Sigma})_{j-8}\right)\right) \qquad \text{para } j \in \{9, \dots, 16\}$$

donde k es un parámetro de escalamiento.

Esta definición propone lo siguiente. Primero, en cada entrada del vector $(\sqrt{(8+k)\Sigma})_i$, encontramos los escalares que crean la dirección del *i*-ésimo

punto sigma, así, aplicando b^{-1} encontramos dicha dirección en \mathfrak{g} . Luego, al aplicar la función expp, mapeamos de \mathfrak{g} a la variedad, cerca de la identidad. Finalmente, se translada cerca de la muestra al multiplicar por μ . Deduciendo así que $\mu \cdot expp\left(b^{-1}[(\sqrt{(8+k)\Sigma})_i]\right)$ será un punto en $SL_3(\mathbb{R})$ que representará un vértice de una "elipsoide" que se ajuste a la muestra. Análogamente para el término $\mu \cdot expp\left(b^{-1}[-(\sqrt{(8+k)\Sigma})_{j-8}]\right)$. Además, los puntos $g^{[9]}, ..., g^{[16]}$ representan los vértices opuestos a $g^{[1]}, ..., g^{[8]}$ respectivamente.

Observemos que los puntos sigma $g^{[1]}, ..., g^{[16]}$ también son transformaciones proyectivas, por lo tanto, podemos transformar una imagen de referencia, Ref, en distintas vistas. Cada vista estaría sobre un vértice de la "elipsoide" que se ajusta a la muestra.

En el capítulo 5 mostramos un algoritmo para encontrar un par de puntos sigma opuestos y cómo utilizarlos para visualizar la dispersión direccional de la muestra de imágenes.

4.2.1. Otras medidas de dispersión

Aunque la matriz de covarianza nos da medidas de dispersión direccionales, en algunas ocasiones podremos necesitar una medida de dispersión de toda la muestra que sea más rápida de calcular. Por esto, nosotros proponemos dos métodos para hallar una medida de dispersión que sea un único valor real.

Definición 4.6. Sean X una variedad riemanniana $y x_1, ..., x_n \in X$. Definimos la **varianza**-F como

$$V_F = \min_{h \in X} \sum_{i=1}^n d(x_i, h)^2$$

donde $d(\cdot, \cdot)$ denota la distancia definida sobre la variedad.

Luego, la varianza-F de $g_1, ..., g_n \in SL_3(\mathbb{R})$ es

$$V_F = \min_{h \in SL_3(\mathbb{R})} \sum_{i=1}^n ||logg(g_i^{-1} \cdot h)||_e^2.$$

Esta definición propone que si el argumento que minimiza la función de distancias es la media de Fréchet, entonces el valor mínimo nos dirá información de la dispersión de los datos.

Definición 4.7. Sean X una variedad riemanniana $y x_1, ..., x_n \in X$. Definimos el **rango** como

$$R = \max_{x_i, x_j \in X} d(x_i, x_j)$$

donde $d(\cdot, \cdot)$ denota la distancia definida sobre la variedad.

Notemos que en los números reales, la distancia máxima estará dada por $X_{(n)} - X_{(1)}$, es decir, con esta definición calculamos una medida de dispersión similar al rango en los números reales.

Luego, el rango de $g_1, ..., g_n \in SL_3(\mathbb{R})$ es

$$R = \max_{g_i, g_j \in SL_3(\mathbb{R})} ||logg(g_i^{-1} \cdot g_j)||_e.$$

Finalmente, proponemos una medida de dispersión similar al rango pero aplicado a las direcciones de la elipsoide que se ajusta a la muestra.

Definición 4.8. Sean $g_1, ..., g_n \in SL_3(\mathbb{R})$. Si ordenamos los puntos sigma por su distancia a $g^{[0]}$ de forma decreciente, es decir, $g^{[1]}$ es el más lejano a $g^{[0]}$ y $g^{[8]}$ es el más cercano. Definimos el **rango direccional** k-ésimo como

$$R_{k} = d\left(g^{[k]}, g^{[k+8]}\right) = ||logg\left((g^{[k]})^{-1} \cdot g^{[k+8]}\right)||_{e}$$

 $con \ k \in \{1, ..., 8\}.$

Esta definición propone que R_1 es el mayor rango direccional y R_8 es el menor. Por otra parte, esta propuesta de medida de dispersión podría no ser tan rápida de calcular porque es necesaria la matriz de covarianza Σ , sin embargo, puede ser útil para interpretar a la misma matriz de covarianza.

4.3. Función "Lie-normal"

En esta sección definiremos una función a partir de un conjunto de datos en $SL_3(\mathbb{R})$ "similar" a la distribución Gaussiana en \mathbb{R}^n . Utilizaremos como parámetros de la función a la media μ y a la matriz de covarianza Σ que presentamos en las secciones anteriores.

Primero recordemos que la distribución normal multivariada en \mathbb{R}^n tiene densidad

$$f(\underline{x}) = \frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}(\underline{x} - \underline{\mu})^t \Sigma^{-1}(\underline{x} - \underline{\mu})\right]$$

con vector de medias μ y matriz de covarianza Σ .

También recordemos que la matriz Σ que calculamos para datos en $SL_3(\mathbb{R})$ está definida sobre \mathfrak{g} . Las "direcciones" utilizadas para medir la dispersión son las direcciones de los vectores de la base de \mathfrak{g} , y las "diferencias" están dadas por $b_k (logg(\mu^{-1}g))$.

Luego, para crear una función con dominio en $SL_3(\mathbb{R})$ e imagen en \mathbb{R} , simplemente sustituimos los términos similares a las diferencias de vectores en \mathbb{R}^n . Así, definimos una función **Lie-normal** en $SL_3(\mathbb{R})$ como

$$\phi(g) = \frac{1}{(2\pi)^{8/2}\sqrt{|\Sigma|}} \exp\left[-\frac{1}{2} \left\{ b \left(logg(\mu^{-1}g) \right)^t \Sigma^{-1} b \left(logg(\mu^{-1}g) \right) \right\} \right].$$

La función ϕ , con los parámetros μ y Σ obtenidos de la muestra, será nuestra propuesta para la distribución a priori de las perspectivas de cada uno de los modelos con los que estemos trabajando, es decir, para un modelo fijo, sustituiremos P(H|M) por $\phi(H)$ en la fórmula 1.

Comentarios finales

Como se comentó en el capítulo 1, las definiciones de variable aleatoria y vector aleatorio no fueron rigurosas, sin embargo, en el libro "Curso intermedio de probabilidad" de Luis Rincón podemos encontrar dichas definiciones.

Todas las definiciones del capítulo 1 fueron expuestas para exhibir algunas nociones estadísticas que tenemos en \mathbb{R}^n , además las utilizamos como base para dar nociones estadísticas en espacios distintos de él, específicamente en $SL_3(\mathbb{R})$.

Aunque para cumplir con los objetivos de la tesis no necesitamos conceptos rigurosos de probabilidad, podemos comentar que algunas de las definiciones probabilísticas que se conocen para el espacio euclídeo tienen sus equivalentes sobre grupos, variedades riemannianas y grupos de Lie. Si se desea más información desde este punto de vista, en el artículo de Xavier Pennec, titulado "Probabilities and statistics on riemannian manifolds: basic tools for geometric measurements", podemos leer algunas definiciones útiles para hablar formalmente de probabilidades sobre una variedad riemanniana. En este artículo se da la definición de **primitiva aleatoria**, que es el equivalente de una variable aleatoria pero para una variedad riemanniana. También se define la **función de densidad de probabilidad de una primitiva aleatoria**, más aún, se definen **la media y la varianza de una primitiva aleatoria** utilizando la media de Fréchet.

De hecho, en [1] se menciona que estrictamente hablando, la función ϕ que proponemos no es una distribución normal, sin embargo, en [8] muestran que existe un equivalente al teorema central del límite sobre un grupo. Por lo tanto, en $SL_3(\mathbb{R})$ existe dicha distribución.

4.4. Inferencia Bayesiana en $SL_3(\mathbb{R})$

En esta sección describiremos cada uno de los elementos que conforman la fórmula para hacer inferencia Bayesiana, y también describiremos cómo la utilizaremos para resolver el problema de clasificación de objetos. Primero recordemos que

$$P(M, H|D) = \frac{P(D|M, H)P(H|M)P(M)}{P(D)}$$
(4.1)

donde D representa una imagen de la cual queremos conocer el modelo M

al que pertenece y la homografía H que transforma a D en la imagen de referencia de dicho modelo.

Notemos que debemos calcular el valor de P(M, H|D) para cada modelo M. Por lo tanto, también para cada modelo debemos calcular los valores de P(D|M, H), P(H|M) y P(M).

El valor de la distribución a priori del modelo M denotado por P(M), lo encontramos utilizando una distribución uniforme, es decir, suponiendo que cada modelo tiene la misma probabilidad de ser elegido. Así, en todos los modelos se cumplirá que

$$P(M) = \frac{1}{\text{Cantidad de modelos}}.$$

Dado un modelo M, hacemos el análisis estadístico sobre las imágenes de la muestra de ese modelo y obtenemos una media μ y una matriz de covarianza Σ . Proponemos a la función Lie-normal ϕ , con los parámetros μ y Σ encontrados, como la distribución a priori de las perspectivas. Luego, encontramos una homografía que trate de transformar la imagen D a la imagen de referencia del modelo M. Finalmente, el valor de P(H|M) está dado por la evaluación de la homografía encontrada en la propuesta de distribución a priori, $\phi(H)$.

En general, el término P(D|H, M) es llamado verosimilitud. Mide, a través de una probabilidad, qué tanto los datos extraídos de la imagen D corresponden a una homografía H y un modelo M.

La definición subyacente a las explicaciones del artículo que se revisó, es la siguiente:

$$P(D|H,M) \propto \exp(-\frac{1}{2}\frac{e(D,H,M)^2}{\sigma_M^2})$$

donde \propto significa "proporcional a". σ_M es una desviación elegida (típicamente 1 ó 2 píxeles). El término e(D, H, M) es un error promedio, observado a partir de los datos de la imagen D y de la homografía H que transforma a D en la imagen de referencia del modelo M, concretamente, las correspondencias encontradas entre D y la imagen de referencia. El valor de e(D, H, M) se puede hallar a partir del conjunto de N correspondencias utilizadas para determinar H, $\{(p_i, p'_i) | \text{ con } i \text{ entre 1 y N}\}$. Cada p_i es un punto de la imagen D que está relacionado con un punto, p'_i , de la imagen de referencia, sobre la cual se calcula H.

Este error promedio se define como:

$$e(D, H, M) = \frac{1}{2} \left\{ \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|p_i' - \pi(p_i, H)\|^2} + \sqrt{\frac{1}{N} \sum_{i=1}^{N} \|p_i - \pi(p_i', H^{-1})\|^2} \right\}$$

donde $\pi(p_i, H)$ representa al operador que transforma a p_i a coordenadas homogéneas, luego multiplica la matriz H por dichas coordenadas homogéneas, y finalmente transforma el resultado a coordenadas euclidianas, y de manera análoga $\pi(p'_i, H^{-1})$. Notemos que dentro de las llaves tenemos una suma de dos términos, el de la izquierda mide el ajuste de los puntos p_i correspondidos por H con los puntos p'_i (con una distancia en la imagen de referencia), mientras que el de la derecha mide el ajuste de los puntos p'_i correspondidos por H^{-1} con los puntos p_i (con una distancia en la imagen de prueba). La función e mide globalmente que tan bien H permite ajustar los puntos entre sí.

Si el valor de e(D, H, M) es muy pequeño, indica que la imagen D y la imagen de referencia del modelo M coinciden muy bien con la homografía H calculada. Además, entre más pequeño sea el valor de e(D, H, M), la probabilidad P(D|H, M) será más alta.

El valor de P(D) no necesita ser calculado, pues para cada modelo M y una imagen fija D se cumplirá que

$P(M, H|D) \propto P(D|M, H)P(H|M)P(M).$

Finalmente, diremos que la imagen D pertenece a un modelo M_0 y está tomada desde una perspectiva H_0 si $P(M_0, H_0|D)$ es mayor comparado con cualquier P(M, H|D) calculado para otro modelo. El valor de H_0 será calculado a partir de otro método que describiremos en el capítulo 6.

Capítulo 5

Algoritmos e implementación

En esta sección describimos los algoritmos más importantes realizados durante esta investigación. Entre los algoritmos que no presentaremos están: el factorial de un número natural, la potencia de una matriz, las funciones expp y logg, la función norma $||\cdot||_e$, así como las funciones b y b^{-1} utilizadas para la matriz de covarianza y los puntos sigma respectivamente. Todos los algoritmos fueron programados en C++ y fueron corridos bajo Linux. Se utilizaron librerías de OpenCV [23] [11] para trabajar con imágenes. Para dichas librerías una imagen es representada por una matriz en la que cada entrada contiene un pixel de la misma.

El algoritmo 5.1 muestra cómo transformar una imagen de referencia Refutilizando una transformación perspectiva. Recordemos que si $H \in SL_3(\mathbb{R})$ es la homografía que transforma la imagen D a la imagen de referencia del modelo M, entonces la transformación que utilizaremos será H^{-1} . La salida es una imagen nueva representada por una matriz que llamaremos New. El algoritmo está hecho de tal forma que la imagen nueva tenga un tamaño fijo desde el principio sin importar Ref y H^{-1} . Las matrices A y B son utilizadas para pasar a coordenadas homogéneas.

Algoritmo 5.1. Cambio de perspectiva

entrada: Ref, H^{-1} salida: New

	dar el tamaño de la imagen nueva e inicialiar la matriz New con ceros
	para todo i entre 0 y New.rows
	para todo j entre 0 y New.cols
•	A(0,0)=j
	A(1,0) = i
	A(2,0) = 1
•	$B = H \cdot A$
	si $B(2,0) \neq 0$ entonces
•	y = B(0,0)/B(2,0)
	x = B(1,0)/B(2,0)
	si $x > 0$ y $x < Ref.$ rows entonces
	si $y > 0$ y $y < Ref.cols$ entonces
	New(i,j) = Ref(x,y)
	fin si
	fin si
•	fin si
•	fin para
	fin para

El algoritmo 5.2 muestra cómo encontar la media, μ , y la varianza-F, V_F , de $g_1, \ldots, g_n \in SL_3(\mathbb{R})$. La idea del algoritmo es encontrar una propuesta de media en cada iteración. Dicha media será utilizada para acercar la muestra a la identidad y así poder utilizar sin ambigüedad la función *logg*. Finalmente se calcula la varianza-F.

Algoritmo 5.2. Media y varianza-F

 $\begin{aligned} \Delta \mu &= expp(\frac{1}{n}\sum_{i=1}^{n}logg(\Delta g_{i})) \\ \hat{\mu} &= \mu \cdot \Delta \mu \\ \hat{\mu} &= \hat{\mu} \\ norm &= ||logg(\Delta \mu)||_{e} \\ \hat{\mathbf{fin mientras}} \\ V_{F} &= \sum_{i=1}^{n}||logg(g_{i}^{-1} \cdot \mu)||_{e}^{2} \end{aligned}$

El algoritmo 5.3 muestra cómo encontrar la media- α de $g_1, ..., g_n \in SL_3(\mathbb{R})$. Recordemos que $||logg(g_i^{-1} \cdot \mu)||_e$ representa la distancia de g_i a μ .

Algoritmo 5.3. Media- α

entrada: $g_1, ..., g_n$ salida: μ_{α}

El algoritmo 5.4 muestra cómo encontrar la media- β de $g_1, ..., g_n \in SL_3(\mathbb{R})$. La idea es calcular, para cada g_i , la suma de las distancias hacia los demás datos y después buscar el mínimo.

Algoritmo 5.4. Media- β

entrada: $g_1, ..., g_n$

salida: μ_{β}

	calcular μ
	sumAux = 0
	para todo i entre 1 y n
•	para todo j entre 1 y n
	$\mathbf{si} \ i \neq j \ \mathbf{entonces}$
	$sumAux = sumAux + logg(g_i^{-1} \cdot g_j) _e$
	fin si
	fin para
	$sum_i = sumAux$
	sumAux = 0
•	fin para
•	$aux = sum_1$
•	min = 1
	para todo i entre 1 y n
•	$si \ sum_i < aux \ entonces$
•	$aux = sum_i$
•	min = i
	fin si
•	fin para
•	$\mu_eta=g_{min}$

El algoritmo 5.5 muestra cómo calcular el rango de $g_1, ..., g_n \in SL_3(\mathbb{R})$. La idea es calcular la distancia de cada pareja de datos y luego encontrar la máxima entre ellas. Notemos que hay $\frac{n(n-1)}{2}$ parejas distinas.

Algoritmo 5.5. Rango

entrada: $g_1, ..., g_n$ salida: R

 $\begin{array}{ll} k = 1 \\ \text{para todo } i \text{ entre 1 y } n \\ \text{para todo } j \text{ entre 1 y } n \\ \text{si } j > i \text{ entonces} \end{array}$

```
dist_{k} = ||logg(g_{i}^{-1} \cdot g_{j})||_{e}
k = k + 1
fin si
fin para
fin para
aux = dist_{1}
aux = dist_{1}
para todo l entre 1 y \frac{n(n-1)}{2}
si dist_{l} > aux entonces
aux = dist_{l}
max = l
fin si
fin para
R = dist_{max}
```

El algoritmo 5.6 muestra cómo calcular la matriz de covarianza, Σ , de $g_1, ..., g_n \in SL_3(\mathbb{R})$. Recordemos que Σ es una matriz de 8×8 con entradas σ_{ij} .

Algoritmo 5.6. Matriz de covarianza

entrada: $g_1, ..., g_n$ salida: Σ

El algoritmo 5.7 muestra cómo calcular, con $g_1, ..., g_n \in SL_3(\mathbb{R})$, el par de puntos sigma opuestos que presentan mayor dispersión, también muestra el rango direccional sobre ese eje, R_1 .

Más aún, este algoritmo muestra cómo visualizar la dispersión direccional sobre el eje que forma ese par de puntos. Crearemos una cantidad de numImg imágenes, $I_1, ..., I_{numImg}$, sobre este eje, la primera y la última corresponderán a las de los puntos sigma. El número numImg debe ser impar. Utilizamos la función CambioPerspectiva, descrita en el algoritmo 5.1, para transformar una imagen de referencia Ref utilizando la transformación perspectiva eje_l^{-1} . El número k es un parámetro de escalamiento.

Notemos que para una matriz simétrica A, tenemos la siguiente factorización

$$A = VDV^{-1} = VDV^{t} = VPPV^{t} = VP(VP^{t})^{t} = VP(VP)^{t}$$

donde D es una matriz diagonal formada por los valores propios de A, V es la matriz ortogonal formada por sus vectores propios correspondientes, y P es la matriz que se forma al sacar la raiz cuadrada de cada unos de los términos de D. Así, para obtener un vector columna de la raiz cuadrada de una matriz simétrica, sólo utilizamos sus valores y vectores propios. Los vectores propios producirán las direcciones y la raiz cuadrada de los valores propios su distancia al centro. El valor propio más grande produce el eje más grande.

Algoritmo 5.7. Puntos sigma

entrada: $g_1,, g_n$ salida: $R_1 \ge I_1,, I_{numImg}$	
	calcular μ
	calcular Σ
	$\Sigma = (8+k)\Sigma$
	calcular λ , el mayor valor propio de Σ
	calcular v_{λ} , el vector propio asociado a λ
•	$vertice = \sqrt{\lambda} v_{\lambda}, vertice \in \mathbb{R}^8$

```
para todo i entre 1 y numImg
            escalar_i = 2i/(numImg - 1) - 1, escalar_i \in \mathbb{R}
        fin para
        para todo j entre 1 y numImg
           eje_j = b^{-1}(vertice), \ eje_j \in \mathfrak{g}
           eje_j = escalar_j \cdot (eje_i), eje_j \in \mathfrak{g}
           eje_i = expp(eje_i), eje_i \in SL_3(\mathbb{R})
           eje_j = \mu \cdot (eje_j), eje_j \in SL_3(\mathbb{R})
        fin para
.
        R_1 = ||logg\left((eje_1)^{-1} \cdot eje_{numImg}\right)||_e
        para todo l entre 1 y numImg
.
            I_l = CambioPerspectiva(Ref, eje_l^{-1})
        fin para
.
```

Capítulo 6

Experimentos y resultados

En este capítulo mostramos los resultados obtenidos al aplicar nuestros algoritmos en fotografías de portadas de libros tomadas desde distintas perspectivas.

Primero mostramos los resultados obtenidos con 24 fotografías del libro Berkeley Problems in Mathematics. En la figura 6.1 exhibimos la imagen que fue utilizada como referencia. Elegimos esta imagen para recordar que el resultado obtenido no se altera al cambiar la imagen de referencia y que no necesariamente debe ser similar a las imágenes de la muestra.



Figura 6.1: Imagen de referencia del libro *Berkeley Problems in Mathematics* utilizada para probar los algoritmos.

En la figura 6.2 mostramos el proceso con el cual encontramos una homografía a partir de una imagen de referencia. Se utilizó como base el tutorial *Features2D + Homography to find a known object* que se encuentra en [23]. El programa que incluye el tutorial primero detecta puntos clave en la imagen de referencia (lado izquierdo) y puntos clave en la imagen de la muestra (lado derecho). Luego, estos se emparejan con base en la similaridad de su apariencia local. Cada punto está asociado a un vector descriptor local que resume la información de apariencia local contenida dentro de un parche centrado en el punto. El diseño de este descriptor está hecho de tal manera que el vector obtenido sea lo más invariante posible a cambios de perspectiva e iluminación. Finalmente se encuentra la homografía que transforma la imagen de referencia utilizando el algoritmo de Ransac [6].

Notemos que con este proceso calculamos la homografía que transforma una imagen de referencia a una imagen de la muestra, sin embargo, para poder utilizar nuestros algoritmos debemos utilizar la matriz inversa de ésta.



Figura 6.2: Los puntos marcados de ambas imágenes son los puntos clave. Los emparejamientos entre los puntos están marcados con un segmento. El marco del lado derecho nos indica exactamente dónde se encuentra la imagen de referencia dentro de la imagen de muestra.

En la figura 6.3 mostramos 14 de las imágenes utilizadas para este libro. En el centro está la imagen más cercana a la "vista media", es decir, la imagen que se obtiene al aplicar la inversa de la media- α a la imagen de
referencia. En este conjunto de imágenes, la media- α y la media- β son la misma. El rango y la varianza-F de la muestra son R = 192.801 y $V_F = 56048$ respectivamente.



Figura 6.3: Diferentes vistas de la portada del libro *Berkeley Problems in Mathematics* con la imagen más cercana a la "vista media" en el centro.

En la figura 6.4 podemos visualizar la dispersión direccional sobre el eje mayor de la "elipsoide" que se ajusta a las homografías obtenidas de la muestra. En la primera fila están las imágenes correspondientes a las inversas de las homografías entre el punto sigma $g^{[1]}$ y la media μ . La imagen que se encuentra abajo a la izquierda es la "vista media", ésta es precisamente la imagen que se obtiene al aplicar la inversa de la media μ a la imagen de referencia. Las últimas cuatro imágenes corresponden a las obtenidas con las inversas de las homografías entre la media y el punto sigma $g^{[9]}$, opuesto a $q^{[1]}$. Podemos notar que la mayor variación está dada por un movimiento similar al de una puerta que se cierra, incluye el movimiento de rotación (a grosso modo a lo largo de un eje vertical) y de cambio de escala.

En este ejemplo el rango direccional R_1 es igual a 0.928 el cual es mucho menor al rango obtenido (R = 192.801), esto se debe a que en este conjunto de imágenes las homografías subyacentes a ellas están muy lejanas. Nuestra noción de distancia entre los vértices de una elipsoide se ve afectada porque la geodésica que une a los puntos sigma $g^{[1]}$ y $g^{[9]}$ no necesariamente pasa por $g^{[0]}$.



Figura 6.4: Imágenes creadas con homografías sobre el eje mayor de la "elipsoide" que se ajusta a las transformaciones de la muestra.

6.1. Aplicación al problema de clasificación

Veamos una aplicación de los resultados obtenidos con fotografías de dos impresiones del libro *Putnam and Beyond*. Las portadas de dichas impresiones son muy parecidas. En la figura 6.5 mostramos las imágenes que fueron utilizadas como referencia.



Figura 6.5: Las portadas son iguales salvo el número de tomo y la palabra "solutions" en 6.5(b).

Cada portada representa un modelo, M_1 representa al tomo I y M_2 representa al tomo II. Para cada modelo tenemos una muestra aleatoria de imágenes tomadas desde distintos puntos de vista. En las figura 6.6 mostramos algunas de las fotografías utilizadas para la muestra de cada tomo. Las imágenes fueron tomadas tratando de seguir una "distribución" específica para cada caso y motivar el uso de un "prior" sobre las perspectivas para el problema de reconocer el número de tomo.

Dada una nueva imagen D, de una portada desde una dirección desconocida, queremos encontrar: a) el modelo al que corresponde (es decir, el número de tomo) y b) la perspectiva H con la cual lo estamos observando. Para verificar nuestros resultados utilizamos 2 imágenes nuevas, denotadas por D_1 y D_2 . En la figura 6.7 mostramos ambas imágenes.

Para empezar hacemos el análisis estadístico sobre cada muestra. En la figura 6.8 se observan las "vistas medias" generadas con las muestras de cada uno de los tomos. El rango obtenido para la muestra del tomo I es de 207.03 y el rango obtenido para la muestra del tomo II es de 192.776, es decir, la muestra del tomo II tiene menor dispersión. Esto también se verifica utilizando la varianza-F. Para el tomo I $V_F = 78025.3$ y para el tomo II $V_F = 62047.3$. Definimos a $\phi_1 : SL_3(\mathbb{R}) \to \mathbb{R}$ como la función Lie-normal que se obtiene utilizando los parámetros μ y Σ de la muestra del tomo I, y de igual forma definimos a $\phi_2 : SL_3(\mathbb{R}) \to \mathbb{R}$ con los parámetros de la muestra del tomo II.

Para clasificar las imágenes D_1 y D_2 seguiremos las ideas presentadas



(b) Tomo II (Solutions)

Figura 6.6: Diferentes vistas de las portadas de las impresiones Tomo I y Tomo II del libro *Putnam and Beyond*.

en la sección 4.4. Recordemos que debemos calcular P(D|M, H), P(H|M) y P(M) para cada modelo.

Empezemos con la distribución a priori de cada modelo. Ya que sólo son 2 modelos, se cumple que $P(M_1) = P(M_2) = \frac{1}{2}$.

Luego, calculemos las homografías que transforman las imágenes D_1 y D_2 en las imágenes de referencia. Aunque estrictamente hablando no es posible encontrar una homografía que transforme, por ejemplo, la imagen D_1 en la imagen de referencia del tomo II, pues ésta no es una transformación de la misma imagen, sí podemos calcular una homografía que "trate de ajustar" la imagen D_1 a la imagen de referencia del tomo II. A esta homografía la denotaremos por $H_{(M_2,D_1)}$ y al resto de las homografías las denotaremos de manera análoga como $H_{(M_2,D_2)}$, $H_{(M_1,D_1)}$ y $H_{(M_1,D_2)}$.

Recordemos que el programa que utilizamos encuentra una homografía a partir de los puntos clave en las imágenes, por lo tanto, este programa es



Figura 6.7: Verificaremos con el método propuesto que D_1 es una fotografía del tomo I y que D_2 es una fotografía del tomo II.



Figura 6.8: "Vistas medias" generadas a partir de las imágenes de referencia y las inversas de su respectivas homografías medias.

útil para calcular una homografía que trate de ajustar las imágenes. En la figura 6.9 mostramos los puntos clave utilizados para encontrar la inversa de la homogafía $H_{(M_2,D_1)}$.

Antes de calcular los valores de la distribución a priori en cada modelo, es sensato mostrar algunos de los valores de las funciones ϕ_1 y ϕ_2 evaluadas en las homografías de las muestras mismas. Los valores mínimo y máximo de ϕ_1 evaluado en las muestras del tomo I son 1.117×10^{-4} y 2.516×10^{-4} respectivamente. Los valores mínimo y máximo de ϕ_2 evaluado en las muestras del tomo II son 6.426×10^{-5} y 1.482×10^{-4} respectivamente.

Ahora que tenemos un rango entre los posibles valores de ϕ_1 y ϕ_2 , calculamos los valores de la distribución a priori.

Para la imagen D_1 :



Figura 6.9: Los puntos marcados de ambas imágenes son los puntos clave. El marco del lado derecho nos indica en dónde se encontraría la imagen D_1 después de aplicar la inversa de la homografía $H_{(M_2,D_1)}$.

$$P(H_{(M_1,D_1)}|M_1) = \phi_1(H_{(M_1,D_1)}) = 1.387 \times 10^{-4}$$
$$P(H_{(M_2,D_1)}|M_2) = \phi_2(H_{(M_2,D_1)}) = 6.585 \times 10^{-5}$$

y para la imagen D_2 :

$$P(H_{(M_1,D_2)}|M_1) = \phi_1(H_{(M_1,D_2)}) = 1.022 \times 10^{-8}$$
$$P(H_{(M_2,D_2)}|M_2) = \phi_2(H_{(M_2,D_2)}) = 2.031 \times 10^{-5}$$

Notemos que podemos utilizar estos resultados como un posible método de clasificación. En este método se elegiría al modelo para el cual se obtenga el mayor valor de P(H|M). Utilizando este método diremos que la imagen D_1 es una fotografía del tomo I pues $P(H_{(M_1,D_1)}|M_1) > P(H_{(M_2,D_1)}|M_2)$. También diremos que la imagen D_2 es una fotografía del tomo II pues $P(H_{(M_1,D_2)}|M_1) < P(H_{(M_2,D_2)}|M_2)$.

Para calcular el valor de la verosimilitud P(D|H, M) primero calculamos el valor del error promedio e(D, H, M). Mostramos los resultados obtenidos en dichas funciones. Para la función de error utilizamos los emparejamientos de puntos clave que fueron necesarios para encontrar la homografía H y la norma euclidiana entre las coordenadas de los píxeles de la imagen. Para la verosimilitud utilizamos una desviación de $\sigma_M = 2$ píxeles.

Para la imagen D_1 :

$$e(D_1, H_{(M_1, D_1)}, M_1) = 1.77 \text{ y } P(D_1 | H_{(M_1, D_1)}, M_1) \propto 0.675$$
$$e(D_1, H_{(M_2, D_1)}, M_2) = 1.17 \text{ y } P(D_1 | H_{(M_2, D_1)}, M_2) \propto 0.842$$

y para la imagen D_2 :

$$e(D_2, H_{(M_1,D_2)}, M_1) = 1.94 \text{ y } P(D_2|H_{(M_1,D_2)}, M_1) \propto 0.624$$

 $e(D_2, H_{(M_2,D_2)}, M_2) = 1.94 \text{ y } P(D_2|H_{(M_2,D_2)}, M_2) \propto 0.624$

Estos resultados obtenidos también los podemos utilizar como un posible método de clasificación. En este método se elegiría al modelo para el cual se obtenga el mayor valor de P(D|H, M). Utilizando este método diremos que la imagen D_1 es una fotografía del tomo II pues $P(D_1|H_{(M_1,D_1)}, M_1) < P(D_1|H_{(M_2,D_1)}, M_2)$, lo cual es un error. Además, no podemos afirmar nada acerca de la imagen D_2 pues en ambos modelos se obtuvo el mismo error promedio y por lo tanto el mismo valor de verosimilitud.

Finalmente calculamos los valores de P(M, H|D) multiplicando los respectivos valores obtenidos para cada imagen y modelo.

Para la imagen D_1 :

$$P(M_1, H_{(M_1, D_1)}|D_1) \propto (1.387 \times 10^{-4})(0.675)(1/2) = 4.681 \times 10^{-5}$$
$$P(M_2, H_{(M_2, D_1)}|D_1) \propto (6.585 \times 10^{-5})(0.842)(1/2) = 2.772 \times 10^{-5}$$

y para la imagen D_2 :

$$P(M_1, H_{(M_1, D_2)}|D_2) \propto (1.022 \times 10^{-8})(0.624)(1/2) = 3.188 \times 10^{-9}$$
$$P(M_2, H_{(M_2, D_2)}|D_2) \propto (2.031 \times 10^{-5})(0.624)(1/2) = 6.336 \times 10^{-6}$$

Así, utilizando la fórmula de Bayes completa, diremos que la imagen D_1 es una fotografía del tomo I pues $P(M_1, H_{(M_1,D_1)}|D_1) > P(M_2, H_{(M_2,D_1)}|D_1)$, y la imagen D_2 es una fotografía del tomo II pues $P(M_1, H_{(M_1,D_2)}|D_2) < P(M_2, H_{(M_2,D_2)}|D_2)$, como era de esperarse.

Capítulo 7

Conclusiones y trabajo a futuro

En este proyecto de tesis se abordó un método de solución al siguiente problema de clasificación: Suponiendo que tenemos varios objetos planos y que cada uno representa un modelo M. Dada una imagen D de un objeto plano desde una dirección desconocida, queremos encontrar el modelo M al que pertenece y la perspectiva H con la cual lo estamos observando.

El problema se dividió en subproblemas utilizando inferencia Bayesiana, y el principal objetivo era proponer una función consistente para calcular P(H|M).

Se estudiaron e implementaron las propuestas encontradas en [1] para calcular una media y una matriz de covarianza para un conjunto de datos en $SL_3(\mathbb{R})$.

Con el fin de utilizar las imágenes dadas por la muestra, se propusieron dos nuevas medidas de tendencia central, **media-** α y **media-** β . Y aunque éstas tienen definiciones distintas, pudimos comprobar que en muchos casos eran iguales.

Para obtener información relacionada con la dispersión de los datos de la muestra, con un único valor en \mathbb{R} , se propusieron dos medidas de dispersión, la **varianza-**F, que proviene de la media de Fréchet, y el **rango**, que coincide con la definición de rango para un conjunto de datos en los números reales.

Se dio una propuesta original para definir **puntos sigma** sobre una va-

riedad y así visualizar la dispersión direccional de un conjunto de datos. Además, se definió un **rango direccional** para obtener un valor numérico de la dispersión sobre una dirección específica y se observó que éste nos dará información más relevante si los datos son muy cercanos entre sí.

En [1] dan una propuesta para la distribución a priori de las perspectivas en cada uno de los modelos. Definen una función ϕ utilizando la media y la matriz de covarianza para un conjunto de datos en $SL_3(\mathbb{R})$, pero suponen que el determinante de la matriz de covarianza para cada modelo es el mismo o que su valor no es realmente importante como para considerarlo. Nosotros eliminamos esa condición y propusimos una función ϕ más general.

Todas estas nuevas propuestas para hacer el análisis estadístico sobre la variedad también se implementaron en C++ utilizando librerías de OpenCV.

Se implementaron también las funciones de error promedio, e(D, H, M), y de verosimilitud, P(D|H, M). Se estudiaron 3 métodos propuestos en [1] para el problema de clasificación. El primero está basado en utilizar sólo el término P(H|M) para cada modelo. El segunto está basado, similarmente, en sólo utilizar el término P(D|M, H) para cada modelo. El tercero está basado en utilizar el término P(M, H|D) aplicando el teorema de Bayes.

Se desarrollaron dos ejemplos de clasificación. Se observó que el método alternativo basado sólo en la distribución a priori de las perspectivas obtiene buenos resultados. Sin embargo, el método alternativo que sólo está basado en la verosimilitud es poco confiable, pues depende en gran medida del error que pueda existir al encontrar una homografía a partir de las imágenes de referencia de cada modelo M hacia una imagen nueva D. Finalmente, se verificó que con el método de la fórmula de Bayes completa se consiguen buenos resultados.

Aunque los resultados de las 2 pruebas que hicimos resultaron satisfactorios, como trabajo a futuro se planea realizar otras pruebas con más modelos distintos (portadas de libros) y con otras imágenes nuevas.

También se tiene contemplado utilizar otro tipo de objetos planos para diversificar las aplicaciones. Por ejemplo:

- Puertas y ventanas a lo largo de un pasillo, con el fin de que un robot pueda entregar algún paquete en alguna puerta específica.
- Fachadas de las casas de cierta zona de una ciudad obtenidas con Street View de Google Maps, con el fin de localizar un domicilio del cual no se tiene la información exacta de su ubicación salvo una fotografía.

Como trabajo a futuro también se pretende desarrollar algo que llamamos "interpolación gráfica". Para cada dos imágenes, visualizaremos la geodésica formada por las homografías subyacentes a estas dos imágenes. En otras palabras, haremos algo similar a la visulización de la dispersión direccional entre 2 puntos sigma de una muestra, pero para 2 puntos cualesquiera en $SL_3(\mathbb{R})$.

Bibliografía

- BEGELFOR, E. y WERMAN, M; How to put probabilities on homographies; IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 27, n. 10, Pág:1-8, 2005.
- [2] BERGER, M.; A Panoramic View of Riemannian Geometry; Springer, USA, 2003.
- [3] BERLANGA, R, HERNÁNDEZ, L., y SÁNCHEZ, A.; Introducción a la geometría de los grupos de Lie; Notas de los cursos "Introducción a los grupos de Lie" y "Geometría de grupos de Lie", impartidos en CIMAT, México, 1998.
- [4] BUSER, P., y KARCHER, H.; *Gromov's almost flat manifolds*; Asterisque, Francia, 1981.
- [5] CURTIS, M.; *Matrix Groups*; Springer, USA, 1984.
- [6] FISCHLER, M.A. y BOLLES, R.C.; Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography.; Communications of the ACM, 24(6):381-395, 1981.
- [7] FORSYTH, D. y PONCE, J.; Computer Vision: A Modern Approach; Prentice Hall, USA, 2011.
- [8] GRENANDER, U.; Probabilities on Algebraic Structures; Wiley, USA, 1963.
- [9] GUILLEMIN, V. y POLLACK, A.; Differential topology; Prentice-Hall, USA, 1974.
- [10] HARTLEY, R., y ZISSERMAN, A.; Multiple View Geometry in computer vision; Cambridge University Press, U.K., 2004.

- [11] KAEHLER, A. y BRADSKI, G.; Learning OpenCV: Computer Vision in C++ with the OpenCV Library; O'REILLY, USA, 2015.
- [12] MENDENHALL, W., y WACKERLY, D.; *Estadística Matemática con Aplicaciones*; Cengage Learning, México, 2010.
- [13] MERY, D.; Visión por Computador; Notas del curso "Visión por Computador", impartido en el Departamento de Ciencia de la Computación, Universidad Católica de Chile, Chile, 2004.
- [14] MEYER, P.; Probabilidad y aplicaciones estadísticas; Sistemas Técnicos de Edición, México, 1986.
- [15] PENNEC, X.; Probabilities and Statistics on Riemannian Manifolds: Basic Tools for Geometric measurements; Proc. of Nonlinear Signal and Image Processing 99, pp. 194-198.
- [16] RAMÍREZ-GALARZA, A., y SEADE, J.; Introducción a la geometría avanzada; Las prensas de ciencias, México, 2013.
- [17] RINCÓN, L.; Curso intermedio de probabilidad; Las prensas de ciencias, México, 2013.
- [18] STACHNISS, C.; Unscented Kalman Filter; Recuperado el 20 de enero de 2016 de: http://ais.informatik.uni-freiburg.de/teaching/ws12/ mapping/pdf/slam05-ukf.pdf; Alemania.
- [19] SZELISKI, R.; Computer Vision: Algorithms and Applications; Springer, U.K., 2011.
- [20] TORRES DEL CASTILLO, G.; Differentiable Manifolds, A Theoretical Physics Approach; Springer, USA, 2010.
- [21] TURAGA, P., y SRIVASTAVA, A.; *Riemannian Computing in Computer Vision*; Springer, USA, 2016.
- [22] WANG, Z.; Lecture 20: The exponential map; Recuperado el 5 de mayo de 20016 de: http://www-personal.umich.edu/~wangzuoq/635W12/ Notes/Lec%2020.pdf; USA.
- [23] OpenCV. 2015. http://opencv.org/.