TRABAJO MONOGRÁFICO

Número de Condición y Matrices Aleatorias

Diego Armentano

Orientador: Mario Wschebor

Licenciatura en Matemática
Facultad de Ciencias
Universidad de la República
Montevideo, Uruguay

11 de Noviembre del 2005

Resumen

El objetivo de esta monografía es estudiar el *número de condición* de sistemas de ecuaciones lineales cuando los coeficientes del sistema sean aleatorios. Exploraremos resultados en el caso que los coeficientes de la matriz sean variables aleatorias independientes con igual distribución. Mostraremos el comportamiento de la cola de la distribución del número de condición en el caso de variables Gaussianas. Con ese fin introduciremos algunos conceptos de campos aleatorios demostrando la *Fórmula de Rice*. Incluiremos también comentarios sobre el número de condición en problemas más generales.

Abstract

The goal of this monograph is to study the *condition number* of systems of linear equations when the coefficients of the system are random. We will explore results in the case that the coefficients of the matrix are independent variables with equal distribution. We will show the behavior of the tail of the distribution of the condition number in the case of Gaussian variables. With that aim we will introduce some concepts of random fields proving *Rice Formula*. We will also include some remmarks on the condition number in more general problems.

Palabras Claves: Número de condición, Fórmula de Rice, Matrices Aleatorias.
Key words: Condition number, Rice Formula, Random matrices.

Índice general

Capítulo 1

Introducción

La resolución de sistemas de ecuaciones lineales por medio de computadoras, es un tema de mucha importancia en las diferenentes áreas de la ciencia. Los analistas numéricos han estudiado el tiempo de ejecución y la memoria necesaria para implementar diferentes algoritmos de álgebra lineal. Este es el caso de algoritmos directos que resuelven un sistema de ecuaciones lineales en una cantidad finita de pasos. Pero las computadoras digitales quiebran la posibilidad de resolver dichos sistemas con precisión exacta. En la mayoría de los casos hay una pérdida en la precisión en cada uno de los pasos del proceso de resolución.
Una de las mayores preocupaciones es estimar el efecto de varios errores. Contaré algunos del los problemas que surgen:

1. Los elementos de la matriz A a considerar pueden traer errores inherentes al problema en si. Por ejemplo que los datos sean extraídos de alguna medida física, como es el caso de muchos problemas surgidos de la ciencia. Para este tipo de errores podemos establecer una cota δ, como por ejemplo la precisión del instrumento, entre otras. Entonces la matriz A es perturbada por una matriz $E = ((e_{ij}))$ donde los $|e_{ij}| < \delta$, y por lo tanto la matriz en la cual se realizarán las operaciones es $A + E$ y no A.

2. Quizás los elementos de la matriz estén exactamente definidos por alguna fórmula algebraica, pero aún así es probable que dichos elementos no estén en el sistema aritmético de la computadora. Las computadoras actuales trabajan con el sistema de aritmética de *punto flotante*, y entonces todas las operaciones aritméticas son hechas en un subconjunto finito $\mathcal{E}$ de $\mathbb{R}$ en vez de todo el conjunto $\mathbb{R}$. Unas de las características de este sistema aritmético es la existencia de una función $r : \mathbb{R} \to \mathcal{E}$ llamada *función redondeo* que aproxima el dato ingresado (en algún sentido le da el valor del elemento más proximo de $\mathcal{E}$). De esta manera todas las operaciones aritméticas en $\mathbb{R}$ pasan a operaciones aritméticas en $\mathcal{E}$ a través de la función r. Es decir se pasan los números a sus redondeos,

luego se realizan las operaciones establecidas y luego se redondea el resultado. Por ejemplo, multiplicar dos elementos x, y en $\mathbb{R}$ es hacer $r(r(x) \cdot r(y))$. De ésta manera aunque estemos seguros que los datos de la matriz son los verdaderos, está el problema de redondeo en cada uno de los datos.

3. Otro problema es que, como se explicó en el ejemplo anterior, si tenemos dos números ya en el subconjunto $\mathcal{E}$ de $\mathbb{R}$, las operaciones pueden caer fuera de $\mathcal{E}$ y por lo tanto al implementar un algoritmo con la matriz es muy probable que en las operaciones realizadas halla una pérdida de exactitud en los datos (aunque evitemos los dos problemas anteriores).

La pregunta que surge para este tipo de problemas (a excepción del primer problema que de alguna manera no nos concierne) es ¿cuánto afecta el error en la entrada de los datos la solución del sistema?. Esta pregunta es la clave para poder hacer algoritmos más eficientes, en el sentido que tengan mayor rapidez y que los errores acumulados no sean excesivos, como para que la solución esté muy lejos de la verdadera.

Muchos científicos han estudiado estos problemas entre los cuales se destacan Von Neumann y Turing entre otros. Von Neumann en particular se preguntaba si era posible resolver sistemas de ecuaciones lineales con grandes números de variables.

Sea A una matriz $n \times n$ con coeficientes reales invertible, y $b \in \mathbb{R}^n$. Estamos interesados en resolver el sistema $A \cdot x = b$, y queremos estudiar como se afecta la solución si se perturban los datos ingresados (*input*) A y b. Sería conveniente conocer de antemano cuán distante es la solución del sistema perturbado $\hat{x}$ de la solución del sistema x. O sea saber de antemano cual es la sensibilidad de la solución a las perturbaciones en el input. Con tales fines existe una cantidad que nos da una medida de esa sensibilidad. Dicha cantidad se denomina **número de condición** y fue identificada por Turing [16], Von Neumann y Goldstine [15]. El número de condición es definido por

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| \tag{1.1}$$

siendo $\|A\|$ la norma usual de operadores, o sea

$$\begin{aligned} \|A\| &= \sup_{x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|} && (1.2) \\ &= \sup_{\|x\|=1} \|Ax\|. && (1.3) \end{aligned}$$

Aquí $\|\cdot\|$ indica la norma Euclideana en $\mathbb{R}^n$

"It is characteristic of ill-conditioned sets of equations that small percentage errors in the coefficients given may lead to large percentage errors in the solution" (Turing [16]).

Es en este artículo que Turing introduce el *número de condición* de una matriz cuadrada.
Probaremos más adelante en el *Teorema 1* que se cumple la siguiente desigualdad:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}} \cdot \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$$

siendo Δb y ΔA el error relativo en el input A y b respectivamente, y Δx el error en la solución ($\Delta x = x - \hat{x}$). Más aún, la constante $\kappa(A)$ es óptima, en el sentido de que cualquier constante menor no verifica siempre la desigualdad. De esa manera, $\kappa(A)$ da una medida de cuánto amplifica el error relativo en el input la solución. Pero en realidad lo que estamos interesados es en saber la magnitud en termino de cifras significativas. O sea lo que queremos es estudiar $log(\kappa(A))$.
Según como sean representados los números en las máquinas digitales se usa el logaritmo en una base conveniente. Por ejemplo si la maquina trabaja con números binarios usaremos el logaritmo en base 2, pero si trabaja con números decimales usaremos logaritmo en base 10. Pero por conveniencia trabajaremos con el logaritmo en base e.
Cuando la matriz A no es invertible (singular) el número de condición no está bien definido, pero lo podemos extender definiendo $\kappa(A) = \infty$.
Llamaremos matrices "*bien-condicionadas*"(*well-conditioned*) aquellas con $\kappa(A)$ pequeño y "*mal-condicionadas*"(*ill-conditioned*) aquellas con $\kappa(A)$ muy grande.

Estamos interesados en estudiar el número de condición para cierta familia de problemas. Es claro que no conocemos a priori el número de condición de cierto input, y que se supone que es más fácil resolver el problema que computar su complejidad. Una manera de atacar estos problemas es asumir que existe alguna ley probabilística donde los coeficientes (input) se rigen bajo la acción de dicha ley. Es decir, estamos interesados en estudiar el número de condición cuando los coeficientes de la matriz en cuestión son aleatorios. A este tipo de matrices las llamaremos *matrices aleatorias.*
La idea de asumir una medida de probabilidad, nos conduce a poder obtener resultados sobre el valor esperado del logaritmo del número de condición y buscar en lo posible resultados sobre la distribución de éste. Esta idea fue impulsada por Stephen Smale en 1997 para analizar la complejidad de este tipo de problemas. Vale la pena comentar, que sobre estos temas se conoce todavía muy poco, y la gran mayoría de resultados solo se restringe a una clase muy particular de problemas en los que los coeficientes son variables aleatorias Gaussianas.
El objetivo de esta monografía es mostrar resultados sobre el valor esperado del logaritmo del número de condición y de la distribución de la cola del número de condición.

De más está decir que la ausencia de referencias en algunas partes de esta monografía, no supone originalidad de mi parte.

Capítulo 2

Número de Condición

En este capítulo daremos algunas propiedades y aspectos geométricos del número de condición. Empezaremos por demostrar la desigualdad que de alguna manera motiva la definición de dicha cantidad. Luego mostraremos algunas propiedades algebraicas sencillas, para terminar mostrando que el número de condición nos da una medida de la "distancia" al conjunto de matrices singulares.
Las referencias principales de este capítulo son: [9], [12], [13] y [14].

2.1. Propiedades

2.1.1. Motivación

Teorema 1. *Consideremos el sistema de ecuaciones lineales*

$$Ax = b. \tag{2.1}$$

Si llamamos ΔA al error de redondeo en A, Δb al error de redondeo en b y Δx la perturbación en la solución de (2.1) entonces si $\|A^{-1} \cdot \Delta A\| < 1$ se tiene:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}} \cdot \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right).$$

Demostración. Haremos la demostración por partes, primero suponiendo que existe un error de redondeo en b y que A es exacta, luego recíprocamente, para luego hacer el caso conjunto. De esta manera la demostración sera más extensa de lo que en realidad es pero seguramente se entienda más si hacemos cada caso por separado.
Supongamos primero que A no tiene error, entonces $A(x + \Delta x) = b + \Delta b$:
Como x es la solución del sistema de ecuaciones lineales $Ax = b$ entonces

$$A \cdot \Delta x = \Delta b \Longrightarrow \Delta x = A^{-1} \cdot \Delta b$$

y entonces tomando norma en ambos miembros y usando la definición de norma de un operador se tiene

$$\begin{aligned} \|\Delta x\| &= \|A^{-1} \cdot \Delta b\| \\ &\leq \|A^{-1}\| \cdot \|\Delta b\|. \end{aligned}$$

Por otro lado

$$Ax = b \Longrightarrow \|A\| \cdot \|x\| \geq \|b\| \Longrightarrow \|x\| \geq \|b\| \cdot \|A\|^{-1}$$

entonces se tiene que el error relativo

$$\begin{aligned} \frac{\|\Delta x\|}{\|x\|} &\leq \|A\| \cdot \|A^{-1}\| \cdot \frac{\|\Delta b\|}{\|b\|} \\ &= \kappa(A) \cdot \frac{\|\Delta b\|}{\|b\|} \end{aligned}$$

entonces

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \cdot \frac{\|\Delta b\|}{\|b\|}.$$

Supongamos ahora b sin error o sea $(A + \Delta A)(x + \Delta x) = b$:
Nuevamente por (2.1) se tiene

$$\Delta A \cdot (x + \Delta x) = -A \cdot \Delta x$$

o sea

$$\Delta A \cdot x + (\Delta A + A)\Delta x = 0 \Longrightarrow (\Delta A + A)\Delta x = -\Delta A \cdot x.$$

Como $A + \Delta A = A \cdot (I + A^{-1} \cdot \Delta A)$ se tiene que $(I + A^{-1} \cdot \Delta A) \cdot \Delta x = -A^{-1} \cdot \Delta A \cdot x$. Si $\|A^{-1} \cdot \Delta A\| < 1$ entonces $I + A^{-1} \cdot \Delta A$ es invertible[1] y por lo tanto $A + \Delta A$ es invertible.

Teníamos que $(A + \Delta A) \cdot \Delta x = -\Delta A \cdot x$ entonces $\Delta x = -(A + \Delta A)^{-1} \cdot \Delta A \cdot x$ y entonces tomando normas se tiene

$$\frac{\|\Delta x\|}{\|x\|} \leq \|(A + \Delta A)^{-1} \cdot \Delta A\|. \tag{2.2}$$

Llamando $F = A^{-1} \cdot \Delta A$ tenemos que $A + \Delta A = A \cdot (I + F)$ entonces

$$(A + \Delta A)^{-1} = (I + F)^{-1} \cdot A^{-1}$$

[1] En el apéndice *Serie de Neumann* incluimos la demostración.

y como $\|(I+F)^{-1}\| \leq \frac{1}{1-\|F\|}$ se tiene que

$$\|(A+\Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\cdot\|\Delta A\|}. \tag{2.3}$$

Entonces por (2.2) y (2.3)

$$\begin{aligned}
\frac{\|\Delta x\|}{\|x\|} &\leq \frac{\|A^{-1}\cdot\Delta A\|}{1-\|A^{-1}\cdot\Delta A\|} \\
&\leq \frac{\|A^{-1}\|\cdot\|\Delta A\|}{1-\|A^{-1}\cdot\Delta A\|} \\
&\leq \frac{\|A^{-1}\|\cdot\|\Delta A\|}{1-\|A^{-1}\|\cdot\|\Delta A\|},
\end{aligned}$$

y si multiplicamos y dividimos por $\|A\|$ se tiene

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)\cdot\|\Delta A\|}{\|A\|-\kappa(A)\cdot\|\Delta A\|}$$

y escribiéndolo en función del error relativo de la matriz A se tiene:

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)\cdot\frac{\|\Delta A\|}{\|A\|}}{1-\kappa(A)\cdot\frac{\|\Delta A\|}{\|A\|}}.$$

Concluyendo obtuvimos los siguientes resultados:

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A)\cdot\frac{\|\Delta b\|}{\|b\|} \tag{2.4}$$

en el caso A sin perturbar y

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)\cdot\frac{\|\Delta A\|}{\|A\|}}{1-\kappa(A)\cdot\frac{\|\Delta A\|}{\|A\|}}$$

en el caso b sin perturbar.

Ahora con las dos discusiones anteriores estamos preparados para atacar el caso general en que se perturban a la vez A y b:

$$(A+\Delta A)\cdot(x+\Delta x) = b+\Delta b.$$

Desarrollando y usando el hecho de que x es la solución de $Ax=b$ se tiene que

$$(A+\Delta A)\cdot\Delta x = \Delta b - \Delta A\cdot x.$$

Si $\|A^{-1} \cdot \Delta A\| < 1$ entonces $(A + \Delta A)$ es invertible y

$$\|(A + \Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \tag{2.5}$$

(ver apéndice en la sección de Serie de Neumann). Entonces en ese caso $\Delta x = (A + \Delta A)^{-1} \cdot (\Delta b - \Delta A \cdot x)$ entonces

$$\begin{aligned}
\|\Delta x\| &= \|(A + \Delta A)^{-1} \cdot (\Delta b - \Delta A \cdot x)\| \\
&\leq \|(A + \Delta A)^{-1}\| \cdot \|(\Delta b - \Delta A \cdot x)\| \\
&\leq \|(A + \Delta A)^{-1}\| \cdot (\|\Delta b\| + \|\Delta A \cdot x\|) \\
&\leq \|(A + \Delta A)^{-1}\| \cdot (\|\Delta b\| + \|\Delta A\| \cdot \|x\|)
\end{aligned}$$

entonces dividiendo por $\|x\|$

$$\frac{\|\Delta x\|}{\|x\|} \leq \|(A + \Delta A)^{-1}\| \cdot \left(\frac{\|\Delta b\|}{\|x\|} + \|\Delta A\| \right). \tag{2.6}$$

Usando la definición de norma de un operador en (2.1) se tiene que $\|b\| \leq \|A\| \cdot \|x\|$ y por lo tanto $\|x\| \geq \|A\|^{-1}\|b\|$. Entonces sustituyendo en (2.6) y luego usando (2.5)

$$\begin{aligned}
\frac{\|\Delta x\|}{\|x\|} &\leq \|(A + \Delta A)^{-1}\| \cdot \left(\frac{\|\Delta b\|}{\|x\|} + \|\Delta A\| \right) \\
&\leq \|(A + \Delta A)^{-1}\| \cdot \left(\frac{\|\Delta b\|}{\|A\|^{-1} \cdot \|b\|} + \|\Delta A\| \right) \\
&\leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \left(\frac{\|\Delta b\|}{\|A\|^{-1} \cdot \|b\|} + \|\Delta A\| \right) \\
&\leq \frac{\|A^{-1}\| \cdot \|A\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)
\end{aligned}$$

o sea

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A) \cdot \frac{\|\Delta A\|}{\|A\|}} \cdot \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$$

que es la desigualdad que buscábamos. □

Como ya comentamos, cuando una matriz A no es invertible podemos extender la definición de $\kappa(A)$ para que tome el valor ∞.
Una matriz se dice singular cuando su determinante es cero, y por tanto se podría esperar que las matrices que tengan determinante pequeño tienen que ser las mal condicionadas. Pero ese no es el caso, como veremos en el siguiente ejemplo.

Ejemplos 2.2. Sea $A_\varepsilon = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$, entonces $\det(A_\varepsilon) = \varepsilon^2$.
Como $A_\varepsilon^{-1} = \frac{1}{2}\begin{pmatrix} \varepsilon^{-1} & 0 \\ 0 & \varepsilon^{-1} \end{pmatrix}$, se tiene que $\kappa(A_\varepsilon) = 1$. Entonces si tendemos ε a cero, el determinante tiende a cero mientras que el número de condición se mantiene constante igual a 1.

En la siguiente sección demostraremos un resultado que relaciona el $\det(A)$ con $\kappa(A)$.
Es fácil probar las siguientes dos afirmaciones

Afirmación 2. Sea A una matriz $n \times n$ entonces:

1. $\kappa(\lambda A) = \kappa(A)$ si $\lambda \neq 0$.
2. $\kappa(A^{-1}) = \kappa(A)$ si A es invertible.

2.3. Aspectos Geométricos

Sea Σ el conjunto de matrices singulares (o sea las *ill-posed*). Estas matrices caen en una subvariedad del espacio $\mathbb{R}^{n^2}$. Dicha subvariedad está definida por las matrices con determinante nulo, y como la función determinante es un polinomio, la subvariedad de la que hablamos es una *variedad algebraica* en el espacio $\mathbb{R}^{n^2}$ ($\Sigma = det^{-1}(\{0\})$). Probaremos que Σ es un conjunto de medida de Lebesgue nula en el espacio $\mathbb{R}^{n^2}$.

Lema 1. *Si* $\Sigma = \{A \in M_n(\mathbb{R}) : \ det(A) = 0\}$ *entonces*

$$\lambda_{n^2}(\Sigma) = 0$$

siendo λ_{n^2} *la medida de Lebesgue en el espacio* $\mathbb{R}^{n^2}$.

Demostración. Quizás uno se ve tentado a usar el Teorema de Sard, y argumentar que Σ es una variedad de dimensión menor a n^2 en $\mathbb{R}^{n^2}$ y por tanto la medida de Lebesgue sería cero, pero no estamos en las condiciones del teorema ya que 0 no es un valor regular. Por tanto hay que recurrir a otros métodos dejando de lado resultados de la Geometría Diferencial.
Una matriz $A = ((a_{ij})_{i,j=1})$ la podemos escribir en el espacio $\mathbb{R}^{n^2}$ como

$$A = (a_{11}, a_{12}, \ldots, a_{1n}, a_{2n}, \ldots\ldots, a_{nn}) = (a_{11}, b)$$

entonces si escribimos el determinante desarrollando por la primera fila (o primera columna) se tiene que

$$det(A) = a_{11}\sigma_{11} + B$$

donde σ_{11} y B solo dependen de b. Entonces por Fubini

$$\lambda_{n^2}(\Sigma) = \int_{\mathbb{R}^{n^2}} \lambda_1(\Sigma_b)\, db \quad \text{siendo} \quad \Sigma_b = \{a_{11} : \ a_{11}\sigma_{11} + B = 0\}.$$

Si $\sigma_{11} \neq 0$ entonces $\Sigma_b = \{-\frac{B}{\sigma_{11}}\}$ y por tanto $\lambda_1(\Sigma_b) = 0$.
Entonces

$$\begin{aligned}
\lambda_{n^2}(\Sigma) &= \lambda_{n^2}(\Sigma \cap \{\sigma_{11} \neq 0\}) + \lambda_{n^2}(\Sigma \cap \{\sigma_{11} = 0\}) \\
&= \lambda_{n^2}(\Sigma \cap \{\sigma_{11} = 0\}) \\
&\leq \lambda_{n^2}(\{\sigma_{11} = 0\}),
\end{aligned}$$

pero

$$\{\sigma_{11} = 0\} = \mathbb{R}^{2n-1} \times \Sigma_{n-1} \tag{2.7}$$

siendo $\Sigma_{n-1} = \{A \in M_{n-1}(\mathbb{R}) : \ det(A) = 0\}$. De esta manera si hacemos un razonamiento por inducción usando Fubini en (2.7) se tiene el resultado. Detallemos esto. Para matrices 1×1 es claro que el resultado del Lema es cierto. La hipótesis de inducción es que para matrices $(n-1)\times(n-1)$ vale el resultado, es decir $\lambda_{(n-1)^2}(\Sigma_{n-1}) = 0$. Entonces aplicando Fubini en (2.7) se tiene

$$\begin{aligned}
\lambda_{n^2}(\{\sigma_{11} = 0\}) &= \lambda_{n^2}(\mathbb{R}^{2n-1} \times \Sigma_{n-1}) \\
&\overset{Fubini}{=} \int_{\mathbb{R}^{2n-1}} \lambda_{(n-1)^2}\big((\Sigma_{n-1})_x\big)\, dx \\
&= \int_{\mathbb{R}^{2n-1}} \overbrace{\lambda_{(n-1)^2}(\Sigma_{n-1})}^{=0}\, dx \\
&= 0.
\end{aligned}$$

$\square$

Existe un teorema que nos dice que el número de condición está relacionado con la distancia de A a la variedad algebraica Σ.
Consideraremos otra norma en el espacio de las matrices $n \times n$ distinta a la norma de operadores que comentamos en (1.2)

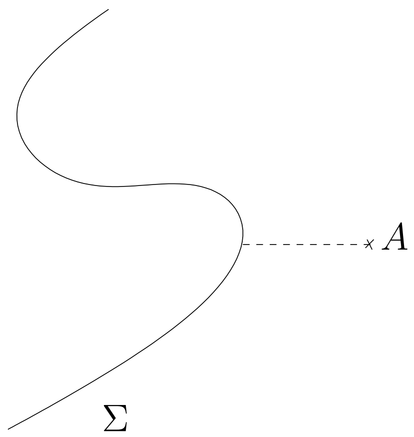
$$\|A\|_F = \sqrt{traza(AA^T)} = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2}$$

conocida como norma de Frobenius o Hilbert-Schmidt, siendo $A = ((a_{ij})_{i,j=1\ldots n})$ y A^T la matriz traspuesta de A, que no es otra cosa que la norma Euclideana en $\mathbb{R}^{n^2}$. En el caso de que las entradas de la matriz sean complejas en vez de la matriz traspuesta se considera la matriz conjugada traspuesta (o sea la matriz adjunta).

Teorema 3 (Teorema del número de condición). *Para toda matriz real A de $n \times n$ se tiene*

$$d_F(A, \Sigma) = \frac{\|A\|}{\kappa(A)}$$

siendo d_F la distancia en el espacio de las matrices $n \times n$ con respecto a la norma de Frobenius.



Este teorema fue probado por [Eckart y Young 1936].
Es necesario especificar algunos detalles antes de empezar la demostración.
Si $A \in M_n(\mathbb{R})$ (matrices reales cuadradas de dimensión n) entonces podemos considerar A como un operador lineal $A : \mathbb{R}^n \to \mathbb{R}^n$ con matriz asociada A en la base canónica. Como estamos en dimensión finita es conocido de los cursos básicos de Cálculo que A es una función continua, y como $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ es también una función continua tenemos que la composición también lo es. Pero el conjunto $\{x \in \mathbb{R}^n : \|x\| = 1\}$ es compacto, entonces por Weirestrass podemos sustituir la definición de $\|A\|$ dada en (1.2) por

$$\|A\| = \max_{\|x\|=1} \|Ax\|. \tag{2.8}$$

Para la prueba del teorema de Eckart-Young es necesario usar que la norma de Frobenius es invariante por transformaciones ortogonales.
Recordemos que una matriz B es ortogonal cuando $BB^T = I$ siendo I la matriz identidad.
Notar que en las matrices ortogonales los vectores columnas de la matriz B (B_i con $i = 1 \dots n$) son ortogonales entre sí y de norma 1. Esto se ve fácilmente pues $\langle B_i, B_j \rangle = \delta_{i,j}$ siendo

$$\delta_{i,j} = \begin{cases} 1 & i = j \\ 0 & i \neq j. \end{cases}$$

Lema 2. *Sea A una matriz $n \times n$ y B una matriz ortogonal también $n \times n$. Entonces*

$$\|AB\|_F = \|BA\|_F = \|A\|_F.$$

Demostración. Notar que cuando multiplicamos por B a la izquierda, la fila i del producto es

$$(\langle A_i, B_1 \rangle, \ldots, \langle A_i, B_n \rangle)$$

siendo A_i la fila i de la matriz A y B_j la columna j de B. Pero como los B_j son ortogonales y de norma uno, por Pitágoras la longitud de la fila es igual a la longitud de A_i.
Análogamente, si se multiplica por la matriz B a la derecha, la longitud de las columnas se mantiene igual a las columnas de A. Por lo tanto la norma de Frobenius se mantiene constante. □

Prueba de Teorema 3. Probaremos que

$$d_F(A, \Sigma) = \frac{1}{\|A^{-1}\|}.$$

Sea $v \in \mathbb{R}^n$ donde se toma el máximo en (2.8) para A^{-1}. Es decir sea v tal que $\|v\| = 1$ y $\|A^{-1}(v)\| = \|A^{-1}\|$.
Sea $u = \frac{1}{\|A^{-1}\|} A^{-1}(v)$, por lo tanto $\|u\| = 1$, $A(u) = \frac{1}{\|A^{-1}\|} v$ y $\|A(u)\| = \frac{1}{\|A^{-1}\|}$.
Si $u = e_1$ y llamamos $\hat{A}$ a la matriz que resulta de cambiar la primer columna de A por todos ceros, tenemos que $\hat{A} \in \Sigma$. Entonces $d_F(A, \Sigma) \leq \|\hat{A} - A\|_F = \|A(u)\| = \frac{1}{\|A^{-1}\|}$ es decir $d_F(A, \Sigma) \leq \frac{1}{\|A^{-1}\|}$.
Si $u \neq e_1$ consideramos B una matriz ortogonal tal que $B(e_1) = u$. Entonces $(AB)_1 = AB(e_1) = A(u)$ es decir la primera columna de la matriz AB es $A(u)$. De forma análoga tomamos la matriz $\hat{AB}$ que resulta de cambiar la primer columna de AB por ceros. Entonces procediendo igual que antes llegamos a que $d_F(AB, \Sigma) \leq \frac{1}{\|A^{-1}\|}$, pero usando el Lema 2 se tiene

$$d_F(A, \Sigma) \leq \frac{1}{\|A^{-1}\|}.$$

Falta la otra desigualdad.
Sea $M \in \Sigma$ la matriz donde se minimiza $d_F(A, \Sigma)$, o sea $d_F(A, \Sigma) = \|A - M\|_F$. Como M es una matriz singular el núcleo no es el trivial. Entonces existe v de norma uno tal que $M(v) = 0$.
Sea B una matriz ortogonal tal que $B(e_1) = v$. De esta manera la primera columna de MB es cero ($(MB)_1 = MB(e_1) = M(v) = 0$). Entonces $\|AB(e_1)\| \leq \|AB - MB\|_F = d_F(A, \Sigma)$.
De la definición de norma de operadores en (1.2) se desprende que

$$\|Au\| \leq \|A\| \, \|u\| \; \forall u \in \mathbb{R}^n,$$

de donde

$$\|w\| = \|A^{-1} \cdot A(w)\| \leq \|A^{-1}\| \cdot \|A(w)\|$$

entonces

$$\|A^{-1}\| \geq \frac{\|w\|}{\|A(w))\|},$$

y tomando $w = B(e_1)$ se tiene que

$$\|A^{-1}\| \geq \frac{\|B(e_1)\|}{\|AB(e_1)\|} \geq \frac{1}{d_F(A, \Sigma)}.$$

□

Lo que nos dice este teorema es que las matrices mal-condicionadas (las de número de condición grande) están cercas de la variedad algebraica Σ. Es decir están en alguna vecindad de cada punto de Σ, lo que significa que caen en algún entorno tubular de Σ.

2.4. Matrices mal condicionadas

Las matrices mal condicionadas son aquellas en que pequeñas perturbaciones en los coeficientes de un sistema de ecuaciones lineales ($Ax = b$) provocan grandes perturbaciones en la solución del sistema. Daremos ejemplos de este tipo de matrices, pero tratemos el caso de matrices sencillas (como por ejemplo las diagonales) para facilitar los cálculos.
Sea $A = \begin{pmatrix} \alpha & 0 \\ 0 & \eta \end{pmatrix}$ con $\alpha >> \eta > 0$. Supongamos A fija y un error en b (δb). Entonces $A \cdot (X + \delta X) = b + \delta b$ y por lo tanto $A \cdot \delta X = \delta b$.
Es claro que A es invertible, entonces $\delta X = A^{-1} \cdot \delta b$, y si tomamos norma se tiene que $\|\delta X\| = \|A^{-1} \cdot \delta b\|$. Como $X = A^{-1} \cdot b$ obtenemos que

$$\frac{\|\delta X\|}{\|X\|} = \frac{\|A^{-1} \cdot \delta b\|}{\|A^{-1} \cdot b\|}. \tag{2.9}$$

Entonces eligiendo de manera adecuada b y δb tal que el numerador en (2.9) sea grande, y el denominador en (2.9) sea chico, tendremos un gran error relativo. Como el comportamiento de A^{-1} según los valores propios es opuesto al de A, basta tomar b en la dirección $(1, 0)$ y δb en la dirección $(0, 1)$. Tomemos $b = (1, 0)$ y $\delta b = \varepsilon \cdot (0, 1)$.
En ese caso la solución del sistema $A \cdot X = b$ es

$$X = \begin{pmatrix} 1/\alpha \\ 0 \end{pmatrix}.$$

La matriz inversa de A es $\frac{1}{\alpha\cdot\eta}\begin{pmatrix} \eta & 0 \\ 0 & \alpha \end{pmatrix}$ entonces

$$\begin{aligned} \delta X &= \frac{1}{\alpha\cdot\eta}\begin{pmatrix} \eta & 0 \\ 0 & \alpha \end{pmatrix}\cdot\begin{pmatrix} 0 \\ \varepsilon \end{pmatrix} \\ &= \frac{1}{\alpha\cdot\eta}\begin{pmatrix} 0 \\ \alpha\cdot\varepsilon \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ \frac{\varepsilon}{\eta} \end{pmatrix} \end{aligned}$$

entonces $\|\delta X\| = |\frac{\varepsilon}{\eta}|$ y $\|X\| = |\frac{1}{\alpha}|$.
Por (2.9) el error relativo es

$$\frac{\|\delta X\|}{\|X\|} = \frac{\alpha}{\eta}\cdot\varepsilon$$

y como estábamos en el caso $\alpha >> \eta$, podemos controlar el ε de la máquina para hacer que el error relativo sea tan grande como querramos.
Es claro que lo anterior lo podríamos haber hecho para matrices tomando distintos vectores propios y ajustando como antes los valores propios para obtener nuevamente una gran perturbación en la solución.
Observar que $\kappa(A) = \alpha\cdot\frac{1}{\eta} >> 1$.

Ejemplos 2.5. $A = \begin{pmatrix} 10^6 & 0 \\ 0 & 10^{-6} \end{pmatrix}$, $\varepsilon = 10^{-3}$ entonces $X = \begin{pmatrix} 10^{-6} \\ 0 \end{pmatrix}$ y $\delta X = (0, 10^3)$ o sea la solución perturbada es $X + \delta X = (10^{-6}, 10^3)$ y $\frac{\|\delta X\|}{\|X\|} = 10^9$.

Capítulo 3

Fórmula de Rice

En este capítulo introduciremos algunas nociones sobre procesos estocásticos y campos aleatorios, orientados hacia el estudio de los extremos de un campo aleatorio. Enunciaremos y demostraremos la Formula de Rice que sera de utilidad para poder estudiar la cola de la distribución del número de condición en el caso Gaussiano.
Las referencias de este capítulo son: [2], [3] y [7].

3.1. Introducción

Un *procesos estocástico* (p.e.) es una familia de variables aleatorias que están definidas en un mismo espacio de probabilidad, que están indexadas por cierto conjunto T. Es decir, dado un espacio de probabilidad $(\Omega, \mathscr{A}, P)$, un p.e. es una función $X : \Omega \times T \to \mathbb{R}$ medible en la primera variable. Los procesos estocásticos serán notados por $\{X(t) : t \in T\}$.
Los procesos para los cuales T se identifica con los enteros o alguna parte de ellos, serán llamados *procesos estocásticos de tiempo discreto.*
Si T coincide con algún intervalo $[a, b]$ (o la recta real, o alguna semirrecta) diremos que $\{X(t) : t \in T\}$ es de *tiempo continuo.*

Si en vez de tener una familia de variables aleatorias tenemos una familia de vectores aleatorios definidos en el mismo espacio de probabilidad a valores en $\mathbb{R}^n$, diremos que es un *campo estocástico* o *campo aleatorio* en $\mathbb{R}^n$.

Consideremos el proceso $\{X(t) : t \in T\}$. Si fijamos $\omega \in \Omega$ obtenemos una función $X(\cdot, \omega) : T \to \mathbb{R}$ tal que $t \mapsto X(t, w)$ que llamaremos *trayectoria* del proceso.
De esta manera podemos ver al proceso como una variable aleatoria que a cada ω le asigna una función.
Diremos que un p.e. tiene trayectorias *continuas* (o C^r) si el mapa $X(\cdot, \omega) : T \to \mathbb{R}$ es continuo (o C^r) salvo en un conjunto de probabilidad nula.
Llamaremos $p_{X(t)}(x)$ a la densidad (en caso de que exista) de la variable aleatoria

$X(t)$ en el punto x.
Diremos que un campo aleatorio $\{Z(t) : t \in T\}$ es *Gaussiano* si dado cualquier conjunto finito $t_1, \ldots, t_m$ de índices el vector $(X(t_1), \ldots, X(t_m))$ tiene distribución conjunta Gaussiana en $\mathbb{R}^m$.

A lo largo de este capítulo, consideraremos que el conjunto de parámetros es una variedad diferenciable compacta I de dimensión d en $\mathbb{R}^d$ y el campo (aleatorio) $Z : I \to \mathbb{R}^d$ es Gaussiano.

3.1.1. Notaciones

- Sea A una matriz simétrica $n \times n$, notamos $A \succ 0$ ($A \prec 0$) si A es definida positiva (definida negativa).

- Sea $f : U \subset \mathbb{R}^d \to \mathbb{R}$, definimos el *módulo de continuidad* para $\delta > 0$ como

$$\omega_f(\delta) = \sup_{x,y \in U} \{|f(x) - f(y)| : \|x - y\| < \delta\}.$$

  Más generalmente si $A(\cdot) : U \subset \mathbb{R}^d \to \mathbb{R}^{m \times n}$ es un campo matricial ($A(t) = (a_{ij}(t))$ con $i = 1 \ldots m$, $j = 1 \ldots n$) definimos $\omega_A(\delta)$ como el máximo módulo de continuidad de sus entradas, es decir

$$\omega_A(\delta) = \max_{\substack{i=1\ldots m \\ j=1\ldots n}} \omega_{a_{ij}}(\delta).$$

- Sea $Z : I \to \mathbb{R}^d$ un campo y $J \subset I \subset \mathbb{R}^d$, definimos $N_u^Z(J)$ como el número de soluciones en J del sistema de ecuaciones $Z(t) = u$. O sea

$$N_u^Z(J) = \#\{t \in J : Z(t) = u\}.$$

La fórmula de Rice establece de manera explícita una fórmula que nos permite expresar el valor esperado de $N_u^Z(J)$ en el caso que Z sea Gaussiano y bajo ciertas hipótesis sobre su regularidad y sobre la variedad I.
El resultado es el siguiente

$$E(N_u^Z(I)) = \int_I E\big(|det(Z'(t))|/Z(t) = u\big) \cdot p_{Z(t)}(u)\, dt.$$

3.2. Prueba Heurística

Teorema 4. *Sea $Z : I \to \mathbb{R}^d$ con $I \subset \mathbb{R}^d$ compacto, un campo aleatorio, y $u \in \mathbb{R}^d$. Asumimos lo siguiente:*

H0 Z *es Gaussiano,*

H1 *c.s. el mapa* $t \longmapsto Z(t)$ *es* C^1*,*

H2 $\forall\ t \in I$*,* $Z(t)$ *tiene distribución no-degenerada* ($var(Z(t)) \succ 0$)*,*

H3 $P\{\exists\, t \in \dot{I},\ Z(t) = u,\ det(Z'(t)) = 0\} = 0$*,*

H4 $\lambda(\partial I) = 0$*.*

Entonces

$$E(N_u^Z(I)) = \int_I E\big(|det(Z'(t))|/Z(t) = u\big) \cdot p_{Z(t)}(u)\, dt. \tag{3.1}$$

En esta sección mostraremos una prueba no formal de la Fórmula de Rice pero que ayuda a entender y justificar la fórmula (3.1).

-Fijemos ω en el espacio muestral.

Tomemos $u \in \mathbb{R}^d$. Por tener medida de Lebesgue nula la frontera de I, con probabilidad uno las pre-imágenes caen todas en el interior de I. Además podemos suponer que u es un valor regular para el campo Z pues lo es con probabilidad uno por $H3$. Por ser Z de clase C^1 y I compacto, $N_u^Z(I) < \infty$. Si $n > 0$, sean $\tau_1, \ldots, \tau_n$ las pre-imágenes por Z de u ($N_u^Z(I) = n$).

Tomando un δ chico, podemos aislar cada una de las pre-imágenes $\tau_1, \ldots, \tau_n$ con entornos abiertos disjuntos $U_1, \ldots, U_n$ respectivamente, de forma tal que:

· Z sea un difeomorfismo local entre U_i y la bola $B_d(u, \delta)$

· si $x \notin \cup_{i=1}^n U_i$ entonces $Z(x) \notin B_d(u, \delta)$

donde $B_d(u, \delta)$ es la bola Euclideana en $\mathbb{R}^d$ de centro u y radio δ.

Entonces

$$\begin{aligned}\int_I |\det(Z'(t))|\, Ind_{\{\|Z(t)-u\|<\delta\}}\, dt &= \sum_{i=1}^n \int_{U_i} |\det(Z'(t))|\, dt \\ &= \lambda_d\big(B_d(u,\delta)\big) \cdot n\end{aligned}$$

entonces

$$N_u^Z(I) = \lim_{\delta \downarrow 0} \frac{1}{\lambda_d(B_d(u,\delta))} \int_I |\det(Z'(t))|\, Ind_{\{\|Z(t)-u\|<\delta\}}\, dt.$$

Tomando esperanza en ambos miembros se tiene

$$\begin{aligned}E(N_u^Z(I)) &= \lim_{\delta \downarrow 0} \frac{1}{\lambda_d(B_d(u,\delta))} \int_I E(|\det(Z'(t))|\, Ind_{\{\|Z(t)-u\|<\delta\}})\, dt \\ &= \lim_{\delta \downarrow 0} \int_I dt\, \frac{1}{\lambda_d(B_d(u,\delta))} \int_{B_d(u,\delta)} E\big(|\det(Z'(t))| \,\big|\, Z(t) = y\big)\, p_{Z(t)}(y)\, dy \\ &= \int_I E\big(|\det(Z'(t))|/Z(t) = u\big) \cdot p_{Z(t)}(u)\, dt\end{aligned}$$

probándose el resultado.
No nos es posible justificar cada una de las igualdades por tanto para la prueba formal necesitamos un trabajo preliminar previo.

3.3. Preliminares Técnicos

En el siguiente lema probaremos que con probabilidad uno no hay pre-imágenes por Z de u en el borde de la variedad.

Lema 3. *Sea I compacto en $\mathbb{R}^d$, $Z : I \to \mathbb{R}^d$ un campo vectorial aleatorio y $u \in \mathbb{R}^d$ tal que:*

1. *El mapa $t \longmapsto Z(t)$ es C^1 c.s.,*
2. *$\lambda_d(\partial I) = 0$,*
3. *$\forall t \in I$, $p_{Z(t)}(x) \leq C$ con $x \in V(u)$ entorno de u,*

entonces

$$P\big(N_u^Z(\partial I) \neq 0\big) = 0.$$

Demostración. Como I es compacto, 1. nos dice que

$$\sup_{\substack{t\in I\\ i,j=1...d}} |\frac{\partial Z_i}{\partial t_j}(t)| \leq M$$

con $M(\omega) < \infty$ con probabilidad 1. Entonces existe M_0 suficientemente grande tal que $P(M > M_0) < \varepsilon$ y sea $A_\varepsilon = \{M > M_0\}$.
Como $\lambda_d(\partial I) = 0$ y ∂I es también compacto, dado $\varepsilon' > 0$, encontramos bolas (determinísticas) $B_1, ..., B_m$ de radio δ_j con $j = 1...m$ respectivamente tales que

$$\sum_{j=1}^m \lambda_d(B_j) < \varepsilon' \quad y \quad \partial I \subset \bigcup_{j=1}^m B_j.$$

$$P\big(N_u^Z(\partial I) \neq 0\big) \leq P\left(N_u^Z\big(\bigcup_{j=1}^m B_j \cap I\big) \geq 1\right) \tag{3.2}$$

$$\leq P\left(\Big\{N_u^Z\big(\bigcup_{j=1}^m B_j \cap I\big) \geq 1\Big\} \cap A_\varepsilon^c\right) + \underbrace{P(A_\varepsilon)}_{<\varepsilon} \tag{3.3}$$

$$< P\left(\Big\{N_u^Z\big(\bigcup_{j=1}^m B_j \cap I\big) \geq 1\Big\} \cap A_\varepsilon^c\right) + \varepsilon. \tag{3.4}$$

Estudiemos el primer miembro después de la desigualdad en (3.4).

$$P\left(\left\{N_u^Z(\bigcup_{j=1}^m B_j \cap I) \geq 1\right\} \cap A_\varepsilon^c\right) \leq \sum_{j=1}^m P\left(\left\{N_u^Z(B_j \cap I) \geq 1\right\} \cap A_\varepsilon^c\right).$$

Si $\omega \in \left\{N_u^Z(B_j \cap I) \geq 1\right\}$ entonces $\exists\ \tau \in B_j \cap I$ tal que $Z(\tau) = u$.
Sean $\hat{t}_j \in B_j \cap I$ fijos.
Entonces si $\omega \in \left\{N_u^Z(B_j \cap I) \geq 1\right\} \cap A_\varepsilon^c$ se tiene

$$\begin{aligned} \|Z(\hat{t}_j) - Z(\tau)\| &= \|Z(\hat{t}_j) - u\| \\ &\leq M_0 d \|\hat{t}_j - t\| \\ &\leq M_0 d 2\delta_j \end{aligned}$$

donde la primera desigualdad sigue del teorema de valor medio para campos vectoriales, que por completitud agregamos su demostración en el Apéndice (6.1).
Tenemos que

$$\|Z(\hat{t}_j) - u\| \leq 2M_0 d\delta_j. \tag{3.5}$$

Concluimos que: si $\omega \in \left\{N_u^Z(B_j \cap I) \geq 1\right\} \cap A_\varepsilon^c$ entonces por (3.5) $\|Z(\hat{t}_j)(\omega) - u\| \leq 2M_0 d\delta_j$. Juntando lo anterior se tiene

$$\begin{aligned} P(\{N_u^Z(B_j \cap I) \geq 1\} \cap A_\varepsilon^c) &\leq P(\|Z(\hat{t}_j) - u)\| \leq 2M_0 d\delta_j) \\ &= \int_{B(u,2M_0 d\delta_j)} P_{Z(\hat{t}_j)}(x)\, dx \\ &\leq C \cdot Vol(B(u, 2M_0 d\delta_j)) \\ &= C \cdot Vol(B(0, 2M_0 d\delta_j)) \\ &= C \cdot (2M_0 d)^d Vol(B(0, \delta_j)). \end{aligned}$$

Entonces

$$\begin{aligned} P(N_u^Z(\partial I) \neq 0) &< \sum_{j=1}^m 2M_0 d)^d Vol(B(0, \delta_j) + \varepsilon \\ &< \varepsilon' C(2M_0 d)^d + \varepsilon. \end{aligned}$$

Como $M_0 = M_0(\varepsilon)$ depende sólo de ε, dado $\varepsilon > 0$ elegimos M_0 y luego ε' tal que $\varepsilon' C(2M_0 d)^d < \varepsilon$, entonces se tiene

$$P(N_u^Z(\partial I) \neq 0) < 2\varepsilon$$

lo que concluye la prueba. □

En la demostración anterior hay un detalle que pasamos por alto. Para usar el teorema de valor medio es necesario que los extremos (en este caso $\hat{t}_j$ y τ) estén conectados. Pero esto no es un problema pues en la hipótesis cuando decimos que el mapa $t \longmapsto Z(t)$ es C^1 en I c.s, queremos decir que tiene una extensión C^1 y por tanto el teorema de valor medio lo podemos usar en un conjunto mayor a I en el que para cada punto de I existe una bolita incluida en la extensión. De esa manera trabajamos en vez de con I, con la extensión, y ahí hacemos el procedimiento.

Lema 4. *Sea $Z : I \to \mathbb{R}^d$ un campo vectorial C^1, I compacto en $\mathbb{R}^d$ y $u \in \mathbb{R}^d$. Asumimos que:*

1. $\inf_{t \in Z^{-1}(\{u\})} \Big(\lambda_{min}\big(Z'(t)\big)\Big) \geq \Delta > 0$,
2. $\omega_{Z'}(\eta) < \frac{\Delta}{d}$,

siendo $\lambda_{min}(M)$ el mínimo valor propio de la matriz $M^T M$.
Entonces, si t_1 y t_2 son soluciones de $Z(t) = u$ y el segmento comprendido entre t_1 y t_2 (que notaremos $[t_1, t_2]$) está incluido en I entonces

$$\|t_1 - t_2\| > \eta.$$

Demostración. Sean $\widetilde{\eta} = \|t_1 - t_2\|$ y $v = \frac{t_1 - t_2}{\|t_1 - t_2\|}$.
Como $Z(t_1) = Z(t_2)$ y $Z(t) = (Z_1(t), \dots, Z_d(t))$ entonces $Z_i(t_1) = Z_i(t_2)$ para $i = 1 \dots d$. Entonces por ser Z diferenciable se tiene por el teorema del valor medio que $\nabla_{\xi_i} Z_i \cdot v = 0$ con $\xi_i \in [t_1, t_2]$. y entonces $(Z'(t) \cdot v)_i = 0$ cuando $t = \xi_i$, es decir $(Z'(\xi_i) \cdot v)_i = 0$ para todo $i = 1 \dots d$, por ejemplo:

$$Z'(\xi_1) \cdot v = \begin{pmatrix} \nabla_{\xi_1} Z_1 \\ \vdots \\ \nabla_{\xi_1} Z_d \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} 0 \\ * \\ * \\ * \end{pmatrix}.$$

Ahora

$$\left|\big(Z'(t_1) \cdot v\big)_i\right| = \left|\big(Z'(t_1) \cdot v\big)_i - \big(Z'(\xi_i) \cdot v\big)_i\right| = \left|\Big(\big[Z'(t_1) - Z'(\xi_1)\big] \cdot v\Big)_i\right|$$

y si desarrollamos en sus coordenadas tenemos

$$\begin{aligned} \left|\big(Z'(t_1) \cdot v\big)_i\right| &\leq \sum_{k=1}^{d} \left|Z'(t_1)_{ik} - Z'(\xi_i)_{ik}\right| \cdot |v_k| \\ &\leq \omega_{Z'}(\widetilde{\eta}) \cdot \sum_{k=1}^{d} |v_k| \\ &\leq \omega_{Z'}(\widetilde{\eta}) \sqrt{d} \end{aligned}$$

donde la segunda desigualdad es consecuencia de que $\|t_1 - \xi_i\| < \|t_1 - t_2\| = \widetilde{\eta}$
Usando 1. tenemos que

$$\begin{aligned} 0 < \Delta &\leq \lambda_{min}(Z'(t_1)) \\ &\leq \|Z'(t_1) \cdot v\| \\ &= \left(\sum_{i=1}^{d} \left((Z'(t_1) \cdot v)_i \right)^2 \right)^{1/2} \\ &\leq \left(\sum_{i=1}^{d} \omega_{Z'}^2(\widetilde{\eta}) d \right)^{1/2} \\ &= \left(\omega_{Z'}^2(\widetilde{\eta}) d^2 \right)^{1/2} \\ &= \omega_{Z'}(\widetilde{\eta}) d. \end{aligned}$$

Entonces hemos probado que $0 < \Delta \leq \omega_{Z'}(\widetilde{\eta})d$ y por tanto se verifica que $\omega_{Z'}(\eta) < \frac{\Delta}{d} \leq \omega_{Z'}(\widetilde{\eta})$ y entonces es necesario que $\eta < \widetilde{\eta} = \|t_1 - t_2\|$. □

3.4. Demostración de la Fórmula de Rice

Una vez demostrados los dos lemas anteriores, podemos proceder a la demostración formal de la Fórmula de Rice.

Demostración. Sea $F : \mathbb{R} \to \mathbb{R}$ una función continua, creciente en sentido amplio tal que $F(x) = 0$ si $x \leq 1/2$ y $F(x) = 1$ si $x \geq 1$.
Sean Δ y η dos números positivos. Definamos la siguiente función aleatoria:

$$\alpha_{\Delta,\eta}(u, Z, I) = F\left(\frac{1}{2\Delta} \inf_{s \in I} \left[\lambda_{min}(Z'(s)) + \|Z(s) - u\| \right] \right) \cdot \left(1 - F\left(\frac{d}{\Delta} \omega_{Z'}(\eta) \right) \right) \tag{3.6}$$

y el conjunto

$$I_{-\eta} = \{ t \in I : \|t - s\| \geq \eta, \forall s \notin I \}.$$

(Notaremos $\alpha_{\Delta,\eta}(u, Z, I)$ por $\alpha_{\Delta,\eta}(u)$).
Si $\alpha_{\Delta,\eta}(u) > 0$ se tiene que:

(i) $\inf_{s \in I} \left[\lambda_{min}(Z'(s)) + \|Z(s) - u\| \right] > \Delta$,

(ii) $\frac{d}{\Delta} \omega_{Z'}(\eta) < 1$ o sea $\omega_{Z'}(\eta) < \frac{\Delta}{d}$.

De (i) se tiene que $\inf_{s \in Z^{-1}(\{u\})} \left[\lambda_{min}(Z'(s)) \right] > \Delta$ siempre que el conjunto de preimágenes de u sea no vacío.

Usando el teoréma de extensión de pre-medidas de Carathedory, es fácil ver que basta probar la Fórmula para paralelepípedos. De esta manera podemos suponer que I es un paralelepípedo y concluimos que si $\alpha_{\Delta,\eta(u)} > 0$ y $N_u^Z(I_{-\eta})$ no es cero, entonces estamos en las hipótesis del lema 4 por la convexidad de los paralelepípedos.
Entonces cubriendo $I_{-\eta}$ con bolas de radio $\eta/2$, a lo sumo puede haber una raíz en cada bola (de lo contrario las raíces distarían menos de η).
Como $I_{-\eta}$ es compacto (es cerrado y esta incluido en I compacto) tenemos que sólo una cantidad finita de esas bolas cubren $I_{-\eta}$ y por ende hay una cantidad finita de raíces, acotada por una constante $C(I,\eta)$ que depende sólo de I y η.

Sea $f:\mathbb{R}^d \to \mathbb{R}$ otra función continua con soporte compacto. Como c.s. Z es Lipchitz y $\alpha_{\Delta,\eta}(u)\cdot f(u)$ es integrable (pues $\alpha_{\Delta,\eta}(u)$ es acotada) usando la fórmula de co-área de Federer se tiene que:

$$\int_{\mathbb{R}^d} f(u)\, N_u^Z(I_{-\eta})\, \alpha_{\Delta,\eta}(u)\, du = \int_{I_{-\eta}} |det(Z'(t)|\, f(Z(t))\, \alpha_{\Delta,\eta}(Z(t))\, dt. \tag{3.7}$$

Tomando esperanzas en ambos miembros se tiene que el primer miembro en (3.7)

$$E\left(\int_{\mathbb{R}^d} f(u)\, N_u^Z(I_{-\eta})\, \alpha_{\Delta,\eta}(u)\, du\right) \overset{Fubini}{=} \overbrace{\int_{\mathbb{R}^d} f(u)\, \underbrace{E\big(N_u^Z(I_{-\eta})\, \alpha_{\Delta,\eta}(u)\big)}_{(i)}\, du}^{(A)}. \tag{3.8}$$

Antes de estudiar el segundo miembro de (3.7) recordar que si $g(X,Y)$ con X, Y son variables aleatorias

$$E\big(g(X,Y)\big) = E\big[E(g(u,Y)|X=u)\big] = \int_{\mathbb{R}^d} E\Big(g(u,Y)\,|\,X=u\Big)\, p_X(u)\, du \tag{3.9}$$

entonces tomando $g\big(Z(t),Z'(t)\big) = \big|det\big(Z'(t)\big)\big|\, f(Z(t))\cdot \alpha_{\Delta,\eta}(Z(t),Z,I)$ se tiene

$$\begin{aligned}
&E\left(\int_{I_{-\eta}} g\big(Z(t),Z'(t)\big)\, dt\right)\\
&\overset{Fubini}{=} \int_{I_{-\eta}} E\Big(g\big(Z(t),Z'(t)\big)\Big)\, dt\\
&= \int_{I_{-\eta}} E\Big[E\Big(g\big(u,Z'(t)\big)\,|\,Z(t)=u\Big)\Big]\, dt\\
&= \int_{I_{-\eta}}\int_{\mathbb{R}^d} E\Big(g\big(u,Z'(t)\big)\,|\,Z(t)=u\Big)\, p_{z(t)}(u)\, du\, dt
\end{aligned}$$

y si deshacemos la notación y usamos el hecho de que $f(u)$ es determinístico se obtiene que la esperanza del segundo miembro de (3.7) es igual a

$$\int_{I_{-\eta}} dt \int_{\mathbb{R}^d} f(u)\, E\left(\big|det\big(Z'(t)\big)\big|\, \alpha_{\Delta,\eta}(u,Z,I)\,\big|\, Z(t)=u\right)\, p_{z(t)}(u)\, du$$

y si aplicamos nuevamente Fubini llegamos al siguiente resultado

$$E\left(\int_{I_{-\eta}} |\det(Z'(t))|\, f(Z'(t))\, \alpha_{\Delta,\eta}(Z(t),Z,I)\, dt\right)$$
$$= \overbrace{\int_{\mathbb{R}^d} f(u) \underbrace{\int_{I_{-\eta}} E\Big(|\det(Z'(t))|\, \alpha_{\Delta,\eta}(u,Z,I) \mid Z(t)=u\Big)\, p_{Z(t)}(u)\, dt}_{(ii)}\, du}^{(B)}.$$

De la igualdad entre (A) y (B) se tiene que salvo un conjunto de medida de Lebesgue nula, (i) y (ii) coinciden, es decir

$$E(N_u^Z(I_{-\eta})\alpha_{\Delta,\eta}(u,Z,I))$$

y

$$\int_{I_{-\eta}} E\Big(|det(Z'(t))|\, \alpha_{\Delta,\eta}(u,Z,I) \mid Z(t)=u\Big)\, p_{z(t)}(u)\, dt$$

coinciden para casi todo $u \in \mathbb{R}^d$. Pero probaremos que ambos términos son continuos en u, entonces tendremos la igualdad $\forall\, u \in \mathbb{R}^d$. La prueba de la continuidad dejémosla para más adelante.
Entonces afirmamos que

$$\underbrace{E(N_u^Z(I_{-\eta})\alpha_{\Delta,\eta}(u))}_{(i)} = \underbrace{\int_{I_{-\eta}} E\Big(|det(Z'(t))|\, \alpha_{\Delta,\eta}(u,Z,I) \mid Z(t)=u\Big)\, p_{z(t)}(u)\, dt}_{(ii)}.$$

Fijemos $u \in \mathbb{R}^d$ y hacemos tender $\eta \downarrow 0$ y luego $\Delta \downarrow 0$ en ese orden.

en (i): Es claro que si $\eta \downarrow 0$ entonces $N_u^Z(I_{-\eta}) \uparrow N_u^Z(\dot{I})$ para cada ω, pues si existe $t \in \dot{I}$ solución de $Z(t) = u$ entonces tomando η suficientemente chico se tiene que $t \in I_{-\eta}$, pero c.s. por el Lema 3 $N_u^Z(\dot{I}) = N_u^Z(I)$. Entonces si $\eta \downarrow 0$, $N_u^Z(I_{-\eta}) \uparrow N_u^Z(I)$ c.s..
Ahora por el Lema 3 y por $H3$

$$\inf_{s\in I} \left[\lambda_{min}(Z'(s)) + \|Z(s)-u\|\right] > 0 \quad \text{c.s.}$$

pues si $Z(s) \neq u$ se sigue la desigualdad, y si $Z(s) = u$ entonces con probabilidad uno $u \in \dot{I}$ por Lema 3, y por $(H3)$ c.s. $\lambda_{min}(Z'(s)) > 0$. Además, c.s.

$\lambda_{min}(Z'(s)) + \|Z(s) - u\|$ es una función continua en s.
Si observamos la definición de $\alpha_{\Delta,\eta}$ en (3.6) se tiene

$$\lim_{\Delta\downarrow 0}\lim_{\eta\downarrow 0}\alpha_{\Delta,\eta}(u) = 1 \quad \text{de forma creciente c.s.}$$

(observar que $\omega_{Z'}(\eta)$ decrece a cero cuando $\eta \downarrow 0$). Por lo tanto usando el Teorema de Beppo Levi (convergencia monótona) se tiene

$$\lim_{\Delta\downarrow 0}\lim_{\eta\downarrow 0} E\big(N_u^Z(I_{-\eta})\,\alpha_{\Delta,\eta}(u)\big) = E\big(N_u^Z(I)\big).$$

en (ii): Observemos que

$$(ii) = \int_{\mathbb{R}^d} E\Big(|det(Z'(t))|\,\alpha_{\Delta,\eta}(u,Z,I)\,\Big|\,Z(t) = u\Big)\,Ind_{I_{-\eta}}(t)\,p_{z(t)}(u)\,dt.$$

Si mostramos que

$$E\Big(|det(Z'(t))|\,\alpha_{\Delta,\eta}(u,Z,I)\,|\,Z(t) = u\Big) \uparrow E\Big(|det(Z'(t))\,|\,Z(t) = u\Big) \qquad (3.10)$$

cuando $\eta \downarrow 0$ y luego $\Delta \downarrow 0$, como $Ind_{I_{-\eta}}(t) \uparrow Ind_{\dot{I}}(t)$ cuando $\eta \downarrow 0$ entonces por Beppo Levi se tiene

$$\begin{aligned}
&\lim_{\Delta\downarrow 0}\lim_{\eta\downarrow 0}\int_{\mathbb{R}^d} E\Big(|\det(Z'(t))|\,\alpha_{\Delta,\eta}(Z(t))\,|\,Z(t) = u\Big)\,Ind_{I_{-\eta}}(t)\,p_{z(t)}(u)\,dt \\
&= \int_{\mathbb{R}^d} E\Big(|\det(Z'(t))|\,|\,Z(t) = u\Big)\,Ind_{\dot{I}}(t)\,p_{z(t)}(u)\,dt \\
&= \int_{\dot{I}} E\Big(|\det(Z'(t))|\,|\,Z(t) = u\Big)\,p_{Z(t)}(u)\,dt \\
&\overset{H4}{=} \int_{I} E\Big(|\det(Z'(t))|\,|\,Z(t) = u\Big)\,p_{Z(t)}(u)\,dt.
\end{aligned}$$

Ahora queda por probar la continuidad de (i) y (ii), y (3.10).
Continuidad de (i)*:*
Sea u_0 fijo, probemos que (i) es continua en u_0 usando el teorema de convergencia dominada. Para eso primero veamos que c.s. $N_u^Z(I_{-\eta})$ es constante en un entorno de u_0 como función de u.
Por $H3$ y por el Lema 3 es claro que c.s. no existe $t \in Z^{-1}(u_0)$ tal que $det(Z'(t)) = 0$. Pero entonces por el Teorema de la Función Inversa podemos encontrar entornos disjuntos de cada una de las pre-imágenes de u_0 por Z tal que $Z(.)$ es un difeomorfismo local. Es fácil ver que como I es compacto, la cantidad de pre-imágenes es finita

(consecuencia del Teorema de la Función Inversa). Sean $\{\tau_1, \ldots, \tau_n\}$ las pre-imágenes de u_0 por Z.

Si nos restringimos al conjunto $I_{-\eta}$, la distancia de $Z(t)$ a u_0 para los t que están fuera de la unión de esos entornos, esta acotada por abajo por una constante positiva que solo depende de $I_{-\eta}$. Esto es así ya que de lo contrario existiría una sucesión $\{t_k\}$ en $\hat{I}_{-\eta} = I_{-\eta} \setminus \bigcup_i^n V_i$ (siendo V_i entorno de τ_i donde Z es un difeomorfismo local con su imagen) tal que $\|Z(t_k) - u_0\| < 1/k$. Pero por la compacidad de $\hat{I}_{-\eta}$, existiría t_0 en $\hat{I}_{-\eta}$ límite de alguna subsucesión de $\{t_k\}$, y como Z es continuo se tendría $Z(t_0) = u_0$ lo que es absurdo.
Entonces para cada ω en un conjunto de probabilidad 1, se tiene que $N_u^Z(I_{-\eta}) = N_{u_0}^Z(I_{-\eta})$ para u variando en un entorno (aleatorio) de u_0.
Además, es claro que $\alpha_{\Delta,\eta}(u)$ es continua y acotada por 1 como función de u (basta ver la definición en (3.6)). Recordando que $N_u^Z(I_{-\eta})$ está acotado por una constante $C(I, \eta)$ independiente de u, la continuidad de (i) es consecuencia directa del Teorema de Lebesgue de convergencia dominada

$$\lim_{u \to u_0} E\big(N_u^Z(I_{-\eta})\, \alpha_{\Delta,\eta}(u)\big) = E\big(N_{u_0}^Z(I_{-\eta})\, \alpha_{\Delta,\eta}(u_0)\big).$$

Continuidad de (ii):
La idea es primero mostrar que la esperanza condicional en el integrando de (ii) es continua como función de u para cada $t \in I$ y luego tratar como antes de aplicar el teorema de convergencia dominada.

Es claro que la función aleatoria $\alpha_{\Delta,\eta}(u) = \alpha_{\Delta,\eta}(u, Z, I)$ depende de $Z(s)$, de $Z'(s)$ y del conjunto I. Por lo tanto la función aleatoria $|det(Z'(t))|\, \alpha_{\Delta,\eta}(u, Z, I)$ es un funcional definido en el conjunto $\{(Z(s), Z'(s)) :\ s \in I \subset \mathbb{R}^d\}$.
Haciendo una regresión Gaussiana de $(Z(s), Z'(s))$ con $s \in I$, respecto de $Z(t)$ obtenemos:

$$Z(s) = Y^t(s) + \alpha^t(s)\, Z(t), \qquad Z'_j(s) = Y_j^t(s) + \beta_j^t(s)\, Z(t)$$

donde $Z'_j(s)$ con $j = 1 \cdots d$ son las columnas de la matriz $Z'(s)$, $Y^t(s)$ e $Y_j^t(s)$ para $j = 1 \cdots d$ son vectores Gaussianos independientes de $Z(t)$, y $\alpha^t(s)$ y $\beta_j^t(s)$ son matrices continuas en las variables s y t. La anterior construcción es posible usando el hecho de que $Z(t)$ no degenera en I $(var(Z(t)) \succ 0)$.

Remplazando en la esperanza condicional obtenemos

$$(ii) = \int_{I_{-\eta}} E\Big(|det(M^t(t,u))|\, \alpha_{\Delta,\eta}(u, Y^t(\cdot) + \alpha^t(\cdot)\, u, M^t(\cdot, u))\Big)\, p_{Z(t)}(u)\, dt$$

siendo $M^t(s, u)$ la matriz con columnas $Y_j^t(s) + \beta_j^t(s)\, u$.
Entonces si u tiende a u_0 se tiene que

$$|det(M^t(t,u))|\, \alpha_{\Delta,\eta}(u, Y^t(\cdot) + \alpha^t(\cdot)\, u, M^t(\cdot, u))$$

tiende a
$$|det(M^t(t,u_0))|\,\alpha_{\Delta,\eta}(u_0, Y^t(\cdot)+\alpha^t(\cdot)\,u_0, M^t(\cdot,u_0)).$$
Entonces usando que $E(|det(M^t(t,u))|)$ es acotado variando t en I y u en algún entorno de u_0 (por tener vectores columna Gaussianos) la continuidad sigue de la aplicación del teorema de convergencia de Lebesgue.

Entonces con la continuidad en cada miembro se tiene que $\forall u \in \mathbb{R}^d$

$$E(N_u^Z(I_{-\eta})\alpha_{\Delta,\eta}(u)) = \int_{I_{-\eta}} E\Big(|det(Z'(t))|\,\alpha_{\Delta,\eta}(Z(t))\,|\,Z(t)=u\Big)\,p_{z(t)}(u)\,dt$$

como habíamos afirmado.

Nos resta probar el $\lim_{\Delta\downarrow 0}\lim_{\eta\downarrow 0}$ de (ii), más precisamente probar la afirmación (3.10)

De ahora en adelante $u \in \mathbb{R}^d$ está fijo.
Haciendo la regresión en (ii) y tomando $\eta \downarrow 0$ se obtiene

$$E\big(N_u^Z(I)\,\beta_\Delta(u,Z,I)\big) =$$
$$\int_I E\left\{|det(Z'(t))|\,F\left(\frac{1}{2\Delta}\min_{s\in I}\left[\lambda_{min}(M^t(s))+\|Y^t(s)+\alpha^t(s)u-u\|\right]\right)\right\}\,p_{z(t)}(u)\,dt$$

donde
$$\beta_\Delta(u,Z,I)\big) = F\left(\frac{1}{2\Delta}\min_{s\in I}\left[\lambda_{min}(Z'(s))+\|Z(s)-u\|\right)\right).$$

El problema surge cuando hacemos tender $\Delta \downarrow 0$ pues no sabemos que ocurre con

$$\lambda_{min}(M^t(s))+\|Y^t(s)+\alpha^t(s)u-u\|.$$

Tomemos $\varepsilon > 0$.

Consideremos las funciones

$$\overline{\gamma^t}(s) = \lambda_{\text{máx}}(M^t(s)) = \max_{\|x\|=1} x^T(M^t(s))^T M^t(s)x$$

$$\underline{\gamma}^t(s) = \lambda_{\text{mín}}(M^t(s)) = \min_{\|x\|=1} x^T(M^t(s))^T M^t(s)x$$

Como $M^t(s)$ es una función continua como función de las variables s y t en I, se tiene que los valores singulares de $M^t(s)$ (es decir los valores propios de $(M^t(s))^T M^t(s)$) se mueven continuamente con probabilidad uno[1].

[1]Más adelante en la sección 4.5.1 se estudiaran más en detalle los valores singulares de una matriz.

Además se tiene que

$$\sup_{t,s\in I} |\det(M^t(s))|$$

pertenece a L^1 como ya vimos.

Sea λ suficientemente grande tal que

$$E\left(\sup_{t\in I} |\det(M^t(t))|\, Ind_{\{\sup_{s,t\in I}\overline{\gamma^t}(s)>\lambda\}}\right) < \varepsilon$$

y sea $\delta > 0$ suficientemente pequeño tal que

$$E\left(\sup_{t\in I} |\det(M^t(t))|\, Ind_{\{\omega_{\underline{\gamma}}(\delta)\geq \frac{1}{2}\frac{\varepsilon^2}{\lambda^{d-1}}\}}\right) < \varepsilon$$

siendo $\omega_{\underline{\gamma}}$ el módulo de continuidad de $\underline{\gamma}$ en $I \times I$.

Sea $J \subset I$ un cubito tal que su diámetro sea menor a δ,[2] entonces

$$E\big(N_u^Z(I)\,\beta_\Delta(u,Z,J)\big) \geq$$
$$\int_J E\left\{|det(Z'(t))|\, F\left(\frac{1}{2\Delta}\min_{s\in I}\left[\lambda_{min}(M^t(s)) + \|Y^t(s)+\alpha^t(s)u-u\|\right]\right)\, Ind_{F_\varepsilon}\right\} p_{z(t)}(u)\,dt$$

donde

$$F_\varepsilon = \left\{\sup_{s,t\in I}\overline{\gamma^t}(s) \leq \lambda, \omega_{\underline{\gamma}}(\delta) < \frac{1}{2}\frac{\varepsilon^2}{\lambda^{d-1}}, |\det(M^t(s))| > \varepsilon\right\}.$$

De esta manera, si $\omega \in F_\varepsilon$ se tiene:

$$|\det(M^t(t))| = \big(\det\{(M^t(t))^T\, M^t(t)\}\big)^{\frac{1}{2}} \leq \big(\lambda_{\text{mín}}(M^t(t))\,\lambda^{d-1}\big)^{\frac{1}{2}}$$
$$\Longrightarrow \underline{\gamma}^t(t) = \lambda_{\text{mín}}(M^t(t)) \geq \frac{\varepsilon^2}{\lambda^{d-1}},$$

entonces en J $(s,t \in J)$ se tiene que

$$\underline{\gamma}^t(s) \geq \frac{1}{2}\,\frac{\varepsilon^2}{\lambda^{d-1}}$$

(pues $|\underline{\gamma}^t(s) - \underline{\gamma}^t(t)| < \omega_{\underline{\gamma}}(\delta)$ para todo $s,t \in J$), y por lo tanto

$$\forall t \in J, \quad \min_{s\in J}\big(\lambda_{min}(M^t(s)) + \|Y^t(s)+\alpha^t(s)u-u\|\big) > 0$$

[2] Recordar que estamos suponiendo que I es un paralelepípedo.

en F_ε.

Entonces si $\omega \in F_\varepsilon$ se tiene

$$F\left(\frac{1}{2\Delta}\min_{s\in J}\left(\lambda_{min}(M^t(s)) + \|Y^t(s) + \alpha^t(s)u - u\|\right)\right) \uparrow 1 \qquad \text{cuando} \quad \Delta \downarrow 0.$$

Pasando al límite ($\Delta \downarrow 0$) se tiene

$$\begin{aligned} E\big(N_u^Z(J)\big) &\geq \int_J E\big(|\det(M^t(t))|\, Ind_{F_\varepsilon}\big) p_{Z(t)}(u)\,dt \\ &= \int_J E\big(|\det(M^t(t))|\big) p_{Z(t)}(u)\,dt - R(\varepsilon) \end{aligned}$$

siendo

$$R(\varepsilon) = \int_J E\big(|\det(M^t(t))|\, Ind_{F_\varepsilon{}^c}\big) p_{Z(t)}(u)\,dt.$$

Pero

$$\begin{aligned} 0 &\leq E\big(|\det(M^t(t))|\, Ind_{F_\varepsilon{}^c}\big) \\ &\leq E\big(|\det(M^t(t))|\, Ind_{\{\sup_{s,t\in I}\overline{\gamma^t}(s)>\lambda\}}\big) + \\ &\quad + E\big(|\det(M^t(t))|\, Ind_{\{\omega_{\underline{\gamma}}(\delta)\geq\frac{1}{2}\frac{\varepsilon^2}{\lambda^{d-1}}\}}\big) + E\big(|\det(M^t(t))|\, Ind_{\{|\det(M^t(s))|\leq\varepsilon\}}\big) \\ &< 3\varepsilon, \end{aligned}$$

entonces en J se tiene:

$$E\big(N_u^Z(J)\big) \geq \int_J E\big(|\det(M^t(t))|\big)\, p_{Z(t)}(u)\,dt - 3\varepsilon \int_J p_{Z(t)}(u)\,dt$$

ahora tomando una partición de I con cubitos disjuntos y sumando se tiene

$$E\big(N_u^Z(I)\big) \geq \int_I E\big(|\det(M^t(t))|\big)\, p_{Z(t)}(u)\,dt - 3\varepsilon \int_I p_{Z(t)}(u)\,dt,$$

y como ε es arbitrario, se tiene

$$E\big(N_u^Z(I)\big) \geq \int_I E\big(|\det(M^t(t))|\big)\, p_{Z(t)}(u)\,dt.$$

La otra desigualdad resulta de acotar $\alpha_{\Delta,\eta}(u, Z, I)$ por la constante 1.

Observación 1.

$$p_{Z(t)}(u) \leq \frac{1}{(2\pi)^{d/2}\big[\det(Var(Z(t)))\big]^{1/2}}$$

donde $Var(Z(t)) = ((E(Z(t)_i Z(t)_j)))_{i,j=1\ldots d}$, entonces $\det(Var(Z(t)))$ es continuo en un compacto, y siempre positivo para todo t en I como consecuencia de no degenerar en I ($H2$). Por lo tanto esa función tiene un máximo en I y en consecuencia $p_{Z(t)}(u) \leq K$ $\forall t \in I$ y $u \in \mathbb{R}^d$.

Hemos concluido

$$E\big(N_u^Z(I)\big) = \int_I E\big(|\det(M^t(t))|\big)\, p_{Z(t)}(u)\, dt,$$

que es lo que queríamos demostrar. $\square$

Capítulo 4

Matrices Aleatorias

En este capítulo mostraremos resultados sobre el valor esperado del logaritmo del número de condición, y mostraremos como se comporta la cola de la distribución del número de condición en el caso de variables Gaussianas.
Las referencias principales de este capítulo son: [1], [4] y [5].

4.1. Algunos resultados sobre $E(log(\kappa(A)))$

Teorema 5. *Sean a_{ij} con $i, j = 1 \dots n$ variables aleatorias independientes e idénticamente distribuidas (en el espacio $(\Omega, \mathscr{A}, \mu)$), siendo la distribución de probabilidad común P. Asumimos que P satisface las siguientes condiciones:*

1. *Para todo par de números reales α y β tales que $\alpha < \beta$ se tiene que:*

$$P([\alpha, \beta]) \leq P\left([-\frac{\beta-\alpha}{2}, \frac{\beta-\alpha}{2}]\right). \tag{4.1}$$

2. $E(|a_{11}|^r) = \int_{-\infty}^{+\infty} |x|^r P(dx) = 1$ *para algún r positivo.*

3. *Existen constantes positivas C y γ tales que*

$$P([-\alpha, \alpha]) \leq C\alpha^{\gamma}, \quad \text{para todo } \alpha > 0.$$

Bajo esas hipótesis se prueba

$$E[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + \frac{1}{\gamma}\left([(2+\gamma)\log n + \log C]^+ + 1\right)$$

donde $a^+ = \max(a, 0)$.

Demostración. Primero veamos que $\|A\|^2 \leq \sum_{i,j=1}^n a_{ij}^2$ $(= \|A\|_F^2)$.

$$\begin{aligned}
\|A\|^2 &= \sup_{\|x\|=1} \|Ax\|^2 = \\
&= \sup_{x_1^2+\cdots+x_n^2=1} (a_{11}x_1 + \cdots + a_{1n}x_n)^2 + \cdots + (a_{n1}x_1 + \cdots + a_{nn}x_n)^2 \\
&\overset{C-S}{\leq} (a_{11}^2 + \cdots + a_{1n}^2) + \cdots + (a_{n1}^2 + \cdots + a_{nn}^2) \\
&= \sum_{i,j=1}^n a_{ij}^2
\end{aligned}$$

donde la desigualdad sigue de Cauchy-Schwartz.
Calculemos por separado las esperanzas de $\log \|A\|$ y la de $\log \|A^{-1}\|$. Para esto usaremos el siguiente lema conocido de los cursos básicos de probabilidad:

Lema 5. *Sea X una variable aleatoria positiva en sentido amplio, definida en el espacio de probabilidad $(\Omega, \mathscr{A}, \mu)$ entonces*

$$E(X) = \int_0^{+\infty} \mu(X > t)\, dt.$$

Usando el anterior cálculo se tiene

$$\mu(\|A\| > t) = \mu(\|A\|^2 > t^2) \leq \mu(\sum_{i,j=1}^n a_{ij}^2 > t^2)$$

entonces usando <u>sólo</u> el hecho de que las variables aleatorias $|a_{ij}|$ tienen igual distribución, se tiene

$$\begin{aligned}
\mu(\|A\| > t) &\leq \mu\left(n^2 \cdot \sup_{i,j=1\ldots n} a_{ij}^2 > t^2\right) \\
&= \mu\left(\sup_{i,j=1\ldots n} |a_{ij}| > \frac{t}{n}\right) \\
&\leq \mu\left(\bigcup_{i,j=1}^n \left\{|a_{ij}| > \frac{t}{n}\right\}\right) \\
&\leq n^2 \cdot \mu\left(\left\{|a_{11}| > \frac{t}{n}\right\}\right),
\end{aligned}$$

entonces se tiene

$$\mu(\|A\| > t) \leq n^2 \cdot \mu\left(\left\{|a_{11}| > \frac{t}{n}\right\}\right). \tag{4.2}$$

Usando el *Lema* 5 se tiene

$$\begin{aligned} E[\log \|A\|] &\leq x_n + \int_{x_n}^{\infty} \mu(\log \|A\| > x)\, dx \\ &\leq x_n + \int_{x_n}^{\infty} \mu(\|A\| > e^x)\, dx \\ &\leq x_n + \int_{x_n}^{\infty} n^2 \mu \left(\left\{ |a_{11}| > \frac{e^x}{n} \right\} \right) dx \end{aligned}$$

donde la primera desigualdad resulta de aplicar el *Lema* 5 a la variable aleatoria que coincide con $(\log \|A\|)^+$ cuando ésta toma valores positivos, y toma el valor constante x_n cuando $(\log \|A\|)^+$ toma el valor 0. La última desigualdad sigue de (4.2). Ahora usando la desigualdad de Markov se tiene que

$$\mu \left(|a_{11}| > \frac{e^x}{n} \right) \leq \left(\frac{e^x}{n} \right)^{-r} E(|a_{11}|^r)$$

y por la hipótesis 2. se tiene

$$\mu \left(|a_{11}| > \frac{e^x}{n} \right) \leq \left(\frac{e^x}{n} \right)^{-r},$$

sustituyendo en la integral tenemos

$$\begin{aligned} E[\log \|A\|] &\leq x_n + n^2 \cdot \int_{x_n}^{\infty} \left(\frac{e^x}{n} \right)^{-r} dx \\ &= x_n + n^{2+r} \cdot \int_{x_n}^{\infty} \frac{1}{e^{rx}}\, dx \\ &= x_n + n^{2+r} \cdot \frac{1}{r \cdot e^{x_n r}}, \end{aligned}$$

es decir

$$E[\log \|A\|] \leq x_n + n^{2+r} \cdot \frac{1}{r} \cdot e^{-x_n r}. \tag{4.3}$$

Nos resta por minimizar esta cota sobre $x_n \geq 0$. Como el segundo miembro de (4.3) es una función convexa en x_n, basta ver donde se anula la derivada (y que éste en el eje real positivo). Derivando el segundo miembro según x_n obtenemos

$$1 - n^{2+r} \cdot e^{-x_n r} = 0$$

despejando y tomando logaritmo se obtiene que el x_n buscado es

$$x_n = \log n \left(\frac{2}{r} + 1 \right)$$

que es un número positivo. De esta manera sustituyendo en (4.3) el x_n minimizante se tiene

$$\begin{aligned} E[\log \|A\|] &\leq \left(\frac{2}{r}+1\right)\log n + n^{2+r}\cdot\frac{1}{r}\cdot e^{-\left(\frac{2}{r}+1\right)r\log n} \\ &= \left(\frac{2}{r}+1\right)\log n + n^{2+r}\cdot\frac{1}{r}\cdot n^{-(1+2/r)r} \\ &= \left(1+\frac{2}{r}\right)\log n + \frac{1}{r}, \end{aligned}$$

es decir

$$E[\log \|A\|] \leq \left(1+\frac{2}{r}\right)\log n + \frac{1}{r} \tag{4.4}$$

obteniendo así la primera parte del resultado.

Ahora consideremos $\|A^{-1}\|$. Para eso notemos $A^{-1} = (b_{ij})_{i,j=1\ldots n}$ donde se sabe que

$$b_{ij} = \frac{a^{(ij)}}{\det(A)}$$

siendo $a^{(ij)}$ el adjunto en la posición (i,j). Como las variables a_{ij} son igualmente distribuidas e independientes, las variables aleatorias $|b_{ij}|$ también son igualmente distribuidas, y por tanto podemos realizar el mismo procedimiento que hicimos para A.[1] De esa manera se tiene

$$\begin{aligned} \mu(\|A^{-1}\| > t) &\leq \mu\left(n^2\cdot\sup_{i,j=1\ldots n}|b_{ij}|^2 > t^2\right) \\ &= \mu\left(\sup_{i,j=1\ldots n}|b_{ij}| > \frac{t}{n}\right) \\ &\leq \mu\left(\bigcup_{i,j=1}^{n}\left\{|b_{ij}| > \frac{t}{n}\right\}\right) \\ &\leq n^2\cdot\mu\left(\left\{|b_{11}| > \frac{t}{n}\right\}\right) \\ &= n^2\cdot\mu\left(\left\{\left|\frac{a^{(11)}}{\sum_{j=1}^{n} a_{1j}\cdot a^{(1j)}}\right| > \frac{t}{n}\right\}\right) \\ &= n^2\cdot\mu\left(\left\{\left|a_{11} + \sum_{j=2}^{n} a_{1j}\cdot\frac{a^{(1j)}}{a^{(11)}}\right| < \frac{n}{t}\right\}\right). \end{aligned}$$

[1]Recordar que lo único que usamos era que los módulos de las variables aleatorias eran identicamente distribuidas.

Llamando $\eta = \sum_{j=2}^{n} a_{1j} \cdot \frac{a^{(1j)}}{a^{(11)}}$ se tiene que a_{11} y η son independientes y por lo tanto

$$\begin{aligned}
\mu\left(|a_{11}| + \eta < \frac{n}{t}\right) &= \int_{-\infty}^{+\infty} \mu\left(|a_{11}| + y < \frac{n}{t}\right) P_\eta(dy) \\
&= \int_{-\infty}^{+\infty} P\left(\left(-\frac{n}{t} - y, \frac{n}{t} - y\right)\right) P_\eta(dy) \\
&\leq \int_{-\infty}^{+\infty} P\left(\left[-\frac{n}{t}, \frac{n}{t}\right]\right) P_\eta(dy) \\
&= P\left(\left[-\frac{n}{t}, \frac{n}{t}\right]\right) \\
&\leq C \cdot \left(\frac{n}{t}\right)^{\gamma}.
\end{aligned}$$

Hemos concluido que

$$\mu(\|A^{-1}\| > t) \leq n^2 \cdot C \cdot \left(\frac{n}{t}\right)^{\gamma}. \tag{4.5}$$

Ahora usando nuevamente el *Lema* 5 y (4.5) se tiene

$$\begin{aligned}
E[\log \|A^{-1}\|] &\leq y_n + \int_{y_n}^{\infty} \mu(\|A^{-1}\| > e^x)\, dx \\
&\leq y_n + \int_{y_n}^{\infty} n^2 \cdot C \cdot \left(\frac{n}{e^x}\right)^{\gamma} dx \\
&= y_n + n^{2+\gamma} \cdot C \cdot \int_{y_n}^{\infty} e^{-\gamma x}\, dx \\
&= y_n + n^{2+\gamma} \cdot C \cdot \frac{e^{-\gamma x}}{\gamma}
\end{aligned}$$

entonces

$$E[\log \|A^{-1}\|] \leq y_n + n^{2+\gamma} \cdot C \cdot \frac{e^{-\gamma y_n}}{\gamma}.$$

Resta por minimizar para $y_n \geq 0$.
Derivando e igualando a cero obtenemos

$$1 - n^{2+\gamma} \cdot C \cdot e^{-\gamma y_n} = 0,$$

entonces tomando logaritmo se tiene

$$y_n = \frac{1}{\gamma} \cdot \Big((2+\gamma) \cdot \log n + \log C\Big) \ \text{ e } \ y_n \geq 0,$$

es decir

$$y_n = \left[\frac{1}{\gamma} \cdot \log\left(n^{2+\gamma} C\right)\right]^{+} = \left[\frac{1}{\gamma} \cdot \Big((2+\gamma) \cdot \log n + \log C\Big)\right]^{+}.$$

Observar que

$$n^{2+\gamma} \cdot C \cdot \frac{e^{-\gamma y_n}}{\gamma} = n^{2+\gamma} \cdot C \cdot \frac{e^{-\gamma \cdot \frac{1}{\gamma} \cdot \log(n^{2+\gamma} C)}}{\gamma} \leq \frac{1}{\gamma}$$

donde la desigualdad se debe al siguiente argumento:
-si $n^{2+\gamma}C < 1$ el exponente de e es 0 y se sigue la desigualdad estricta;
-si $n^{2+\gamma}C \geq 1$ desarrollando sigue la desigualdad, por lo tanto hemos concluido que

$$E[\log \|A^{-1}\|] \leq \left[\frac{1}{\gamma} \cdot ((2+\gamma) \cdot \log n + \log C)\right]^{+} + \frac{1}{\gamma}.$$

es decir

$$E[\log \|A^{-1}\|] \leq \frac{1}{\gamma} \cdot \left([\cdot ((2+\gamma) \cdot \log n + \log C)]^{+} + 1\right). \tag{4.6}$$

Entonces de (4.4) y (4.6) se sigue que

$$E[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right) \log n + \frac{1}{r} + \frac{1}{\gamma} \left([(2+\gamma) \log n + \log C]^{+} + 1\right).$$

$\square$

Observación 2. Se puede probar que la hipótesis 1. implica que la distribución P es simétrica alrededor de 0. De esta manera si las variables aleatorias a_{ij} son integrables entonces su valor esperado es 0.

Observación 3. Es fácil ver que la hipótesis 2. a efectos de estimar $E(\log \kappa(A))$ no es más restrictiva que pedir que la distribución P tenga momento de orden r finito para algún $r > 0$. Esto es consecuencia directa de la invariancia del número de condición por homotecias [2]. Si $m_r = \int_{-\infty}^{+\infty} |x|^r \, P(dx) < \infty$ para algún $r > 0$ entonces basta tomar la nueva matriz $m_r^{-1/r} A$ en vez de la matriz A. De esta manera se puede modificar la distribución P sin modificar el número de condición. Resta ver solamente como se modifica la constante C.

Observación 4. Se pueden modificar sutilmente las hipótesis del teorema cambiando las hipótesis 1. y 3. por la siguiente:
Existen constantes positivas C y γ tales que para todo par de números reales α y β tales que $\alpha < \beta$ se tiene que:

$$P([\alpha, \beta]) \leq C \cdot |\beta - \alpha|^{\gamma}.$$

Esto es consecuencia que en la demostración sólo se utilizo el argumento anterior, que claramente es consecuencia de relacionar 1. y 3..

[2] Recordar que en la sección 2.1 se observo que si $\lambda > 0$ entonces $\kappa(\lambda A) = \kappa(A)$.

4.1.1. Ejemplos

Ejemplo 4.2. Supongamos que P tiene densidad f par y decreciente en sentido amplio en $[0,\infty)$. Supongamos además que

$$m_r = \int_{-\infty}^{+\infty} |x|^r f(x)\, dx < \infty \text{ para algún } r > 0.$$

Cambiando f por $m_r^{1/r} f(m_r^{1/r} x)$ y llamando $\hat{P}$ a la nueva distribución, se tiene

$$\int_{-\infty}^{+\infty} |x|^r m_r^{1/r} f(m_r^{1/r} x)\, dx = \int_{-\infty}^{+\infty} \frac{|u|^r}{m_r} f(u)\, du = \frac{m_r}{m_r} = 1.$$

Además,

$$\begin{aligned} \hat{P}([-\alpha,\alpha]) &= \int_{-\alpha}^{\alpha} m_r^{1/r} f(m_r^{1/r} x)\, dx \\ &= 2\int_0^{\alpha} m_r^{1/r} f(m_r^{1/r} x)\, dx \\ &\leq 2m_r^{1/r} f(0)\alpha, \end{aligned}$$

entonces estamos bajo las hipótesis del *Teorema* 5 con $C = 2m_r^{1/r} f(0)$ y $\gamma = 1$, por lo tanto

$$E[\log \kappa(A)] \leq \left(1 + \frac{2}{r}\right)\log n + \frac{1}{r} + \left[3\log n + \frac{1}{r}\log m_r + \log(2f(0))\right]^+ + 1.$$

Ejemplo 4.3. Supongamos ahora que P tiene distribución uniforme en el intervalo $[-H,H]$. Entonces P tiene densidad

$$f(x) = \frac{1}{2H}, \;\; x \in [-H,H]$$

$$\begin{aligned} m_r &= \int_{-H}^{H} |x|^r \cdot \frac{1}{2H}\, dx = \frac{1}{H} \cdot \int_0^H |x|^r\, dx \\ &= \frac{1}{H} \cdot \frac{H^{r+1}}{r+1} \\ &= \frac{H^r}{r+1} \end{aligned}$$

entonces $m_r = \frac{H^r}{r+1}$, y por lo tanto $\forall r > 0,\ m_r < \infty$.
Haciendo tender r a $+\infty$ se tiene

$$\begin{aligned}
E[\log \kappa(A)] &\leq \lim_{r\uparrow\infty} \left\{ \left(1+\frac{2}{r}\right)\log n + \right. \\
&\qquad \left. +\frac{1}{r} + \left[3\log n + \frac{1}{r}\cdot\log\left(\frac{H^r}{r+1}\right) + \log(2\cdot f(0))\right]^+ + 1 \right\} \\
&= \lim_{r\uparrow\infty} \left\{ \log n + \left[3\log n - \frac{1}{r}\log(r+1)\right]^+ + 1 \right\} \\
&\leq 4\log n + 1,
\end{aligned}$$

entonces

$$E[\log \kappa(A)] \leq \log n^4 + 1.$$

Ejemplo 4.4 *(Concentración cerca de la media).* Si consideramos la familia de distribuciones con soporte en el intervalo $[-1,1]$ dadas por

$$f(x) = \frac{1}{2}\cdot\frac{\gamma}{|x|^{1-\gamma}}\cdot Ind_{[-1,1]}(x) \ \text{ con } \ 0<\gamma<1.$$

Calculemos m_r:

$$\begin{aligned}
m_r &= 2\int_0^1 \frac{1}{2}\cdot\frac{\gamma}{|x|^{1-\gamma}}\cdot|x|^r\,dx = \gamma\int_0^1 |x|^{r+\gamma-1}\,dx \\
&= \gamma\cdot\frac{|x|^{r+\gamma}}{r+\gamma}\Big|_0^1 \\
&= \frac{\gamma}{\gamma+r},
\end{aligned}$$

entonces

$$m_r = \frac{\gamma}{\gamma+r}\ \forall r > 0.$$

Veamos que tomando en consideración el *Ejemplo* 4.2 caemos en las hipótesis del *Teorema* 5.

$$\begin{aligned}
\hat{P}([-\alpha,\alpha]) &\leq \int_{-\alpha}^{\alpha} m_r^{1/r} \cdot f(m_r^{1/r}x)\,dx \\
&= \int_{-m_r^{1/r}\alpha}^{m_r^{1/r}\alpha} f(u)\,du \\
&= 2\int_0^{m_r^{1/r}} \frac{1}{2}\frac{\gamma}{|u|^{1-\gamma}}\,du \\
&= \gamma \cdot \frac{|u|^{\gamma}}{\gamma}\Big|_0^{m_r^{1/r}} \\
&= (m_r^{1/r}\alpha)^{\gamma} \\
&= \left(\frac{\gamma}{\gamma+r}\right)^{\gamma/r} \cdot \alpha^{\gamma}.
\end{aligned}$$

Entonces tomando

$$C = \left(\frac{\gamma}{\gamma+r}\right)^{\gamma/r} = m_r^{\gamma/r}$$

y γ entre 0 y 1 se tiene

$$E[\log\kappa(A)] \leq \left(1+\frac{2}{r}\right)\log n + \frac{1}{r} + \left(\left[(2+\gamma)\log n + \frac{\gamma}{r}\log\left(\frac{\gamma}{\gamma+r}\right)\right]^{+} + 1\right),$$

tomando $r \uparrow +\infty$ se tiene

$$E[\log\kappa(A)] \leq \log n + \left(\frac{2}{\gamma}+1\right)\log n + \frac{1}{\gamma}$$

entonces

$$E[\log\kappa(A)] \leq \left(2+\frac{2}{\gamma}\right)\log n + \frac{1}{\gamma}\,.$$

4.5. Distribución de la cola de $\kappa(A)$

En esta sección enunciaremos y demostraremos el resultado más importante de esta monografía. El resultado nos muestra como se comporta la cola de la distribución de $\kappa(A)$ cuando A es una matriz aleatoria Gaussiana. Pasemos a enunciarlo

Teorema 6. *Sea $A = ((a_{ij})_{i,j=1\ldots m})$ con $m \geq 3$ y $a_{ij} \sim \mathcal{N}(0,1)$.*
Entonces existen constantes universales c y C tales que si $x > 1$

$$\frac{c}{x} < P(\kappa(A) > mx) < \frac{C}{x} \tag{4.7}$$

con $c = 0{,}13$ y $C = 5{,}60$.

Para demostrar éste teorema es necesario pasar por algunos resultados técnicos que abordaremos en la siguiente sección.

4.5.1. Preliminares técnicos

Lema 6 (Lema de Bulinskaia). *Sea $\{X_t : \ t \in I\}$ un proceso estocástico, con trayectorias C^1 (es decir $t \mapsto X_t$ es C^1 c.s.), $I \subset \mathbb{R}$ un intervalo. Sea $u \in \mathbb{R}$, y supongamos que $\forall t \in I$ X_t tiene densidad $p_{X_t}(x)$ y está acotada para t variando en compactos de I y x variando en un entorno de u. Entonces*

$$P(T_u^X \neq \phi) = 0$$

siendo $T_u^X = \{t \in I : \ X_t = u, X_t' = 0\}$ o sea los puntos críticos del proceso estocástico X que toman el valor u.

Demostración. Por el Teorema de Lindelöff, basta probar que

$$P(T_u^X \cap J \neq \phi) = 0 \quad \forall J \subset I \quad \text{intervalo compacto.}$$

Sea $l = long(J)$, C una cota de $p_{X_t}(x)$ en J, y $t_0 < t_1 < \cdots < t_m$ partición uniforme de J, con $t_{j+1} - t_j = l/m$ con $j = 0 \ldots (m-1)$.
Sea $\omega_{X'}(\eta, J)$ el módulo de continuidad de X' en J, y

$$E_{\eta,\varepsilon} = \{\omega_{X'}(\eta, J) \geq \varepsilon\}.$$

Dado $\varepsilon > 0$, podemos elegir $\eta > 0$ tal que $P(E_{\eta,\varepsilon}) < \varepsilon$ como consecuencia de la continuidad de la X', y elegimos m tal que $l/m < \eta$ se tiene

$$\begin{aligned} P(T_u^X \cap J \neq \phi) &\leq P(E_{\eta,\varepsilon}) + \sum_{j=0}^{m-1} P\big(\{T_u^X \cap [t_j, t_{j+1}] \neq \phi\} \cap E_{\eta,\varepsilon}^c\big) \\ &< \varepsilon + \sum_{j=0}^{m-1} P\big(|X_{t_j} - u| \leq \varepsilon \cdot l/m\big) \\ &= \varepsilon + \sum_{j=0}^{m-1} \int\limits_{\{|x-u| \leq \varepsilon \cdot l/m\}} p_{X_{t_j}}(x)\, dx \\ &< \varepsilon + \sum_{j=0}^{m-1} C \cdot \varepsilon \cdot l/m \\ &= \varepsilon + C\varepsilon l \end{aligned}$$

con C y l constantes, y como ε es arbitrario se obtiene el resultado. □

Lema 7. *Sean $0 \leq \mu_1 \leq \ldots \leq \mu_m$ los valores propios de A^tA, entonces casi seguramente los valores propios son dos a dos diferentes.*

Demostración. Que los valores propios son positivos es consecuencia de que la matriz A^tA es definida positiva (los detalles están explicados en el Lema 8).

Sea $\chi_c(\mu) = \mu^m + c_1\mu^{m-1} + \cdots + c_m$ el polinomio característico de A^tA con $c = (c_1, \ldots, c_m)$.
Queremos probar que

$$P(\{c : \chi_c(\mu) \text{ tiene raiz doble}\}) = 0,$$

o sea

$$P(\{c : \exists\mu \text{ tal que } \chi_c(\mu) = 0,\ \chi'_c(\mu) = 0\}) = 0$$

y para eso basta probar que para n fijo

$$P(\{c : \exists\mu,\ |\mu| \leq n \text{ tal que } \chi_c(\mu) = 0,\ \chi'_c(\mu) = 0\}) = 0.$$

Asumiendo que en compactos el proceso estocástico χ_c tiene densidad acotada, el resultado lo obtenemos por el Lema de Bulinskaia.

Para no entrar en detalle y probar cada una de las hipótesis realizadas (que requiere de cierto trabajo técnico) optamos por transitar otro camino.

Existe un resultado de Fisher en que da explícitamente la densidad conjunta de los valores propios. Eso quiere decir que la densidad conjunta de los valores propios es absolutamente continua respecto a la medida de Lebesgue. Es decir que

$$P((\mu_1, \ldots, \mu_m) \in A) = \int_A f_{(\mu_1,\ldots,\mu_m)}(x_1, \ldots, x_m)\, dx_1 \ldots dx_m \tag{4.8}$$

pero es fácil ver que la medida de Lebesgue en $\mathbb{R}^m$ de $\{\mu_i = \mu_j\}$ es 0 por ser un subespacio de dimensión menor a m. Entonces por (4.8)

$$P(\mu_1 = \mu_2) = 0.$$

En general

$$\lambda_m(\cup_{i\neq j}\{\mu_i = \mu_j\}) = 0 \implies P(\mu_i = \mu_j \,\mathrm{con}\, i \neq j \,\mathrm{e}\, i, j = 1 \ldots m) = 0$$

□

Se pueden caracterizar el mayor y el menor valor propio de la matriz A^tA. Para caracterizarlos se puede usar un resultado conocido de los cursos de Análisis Funcional:
si T es un operador normal en un espacio de Hilbert entonces se tiene que

$$\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| \leq 1\}.$$

Pero por razones de completitud y auto-contenido haremos una prueba sin usar argumentos muy sostificados.

Lema 8. *Sea A una matriz real $m \times m$, y $0 \leq \lambda_1 \leq \ldots \leq \lambda_m$ los valores propios de la matriz A^tA. Sea $X : S^{m-1} \to \mathbb{R}$ el polinomio cuadrático asociado a A^tA dado por $X(x) = x^tA^tAx$.*
Entonces

(i) $\lambda_m = \|A\|^2 = \text{máx}_{x\in S^{m-1}} X(x)$.

(ii) Si $\lambda_1 > 0$, $\lambda_1 = \frac{1}{\|A^{-1}\|^2} = \text{mín}_{x\in S^{m-1}} X(x)$.

Demostración. Observar que $X(x) = x^tA^tAx = \langle A^tAx, x\rangle = \langle Ax, Ax\rangle = \|Ax\|^2$ entonces

$$X(x) \geq 0, \tag{4.9}$$

por tanto si $\hat{x}$ es vector propio de A^tA con valor propio asociado $\hat{\lambda}$ se tiene

$$\begin{aligned} X(\hat{x}) &= \langle A^tA\hat{x}, \hat{x}\rangle & (4.10)\\ &= \langle \hat{\lambda}\hat{x}, \hat{x}\rangle & (4.11)\\ &= \hat{\lambda}\|\hat{x}\|^2, & (4.12)\end{aligned}$$

entonces de (4.9) y (4.12) se tiene que los valores propios de A^tA tienen que ser positivos o cero.
Sea $\tilde{X} : \mathbb{R}^m \to \mathbb{R}$ la extensión de X a todo $\mathbb{R}^m$, o sea $\tilde{X}(x) = \langle A^tAx, x\rangle$ con $x \in \mathbb{R}^m$. Su derivada usual es $\tilde{X}'(x) = 2A^tAx$ (derivada libre).

Queremos estudiar los máximos y mínimos de $\tilde{X}$ con la condición subsidiaria $g(x_1, \ldots, x_m) = 0$ siendo

$$g(x_1, \ldots, x_m) = x_1^2 + \cdots + x_m^2 - 1 \quad (g^{-1}(\{0\}) = S^{m-1} \subset \mathbb{R}^m).$$

Entonces para buscar los máximos y mínimos con esa condición subsidiaria basta considerar los vectores $x \in S^{m-1}$ tales que $\tilde{X}'(x) = \alpha x$ es decir que el gradiente sea perpendicular al espacio tangente de x, que no es otra cosa que decir que sea paralelo al vector x. Pero como $\tilde{X}'(x) = 2A^t Ax$ entonces estamos buscando los vectores propios de $A^t A$ que son $\lambda_1, \ldots, \lambda_m$.
Sean $s_1 \ldots, s_m$ los vectores propios respectivos en S^{m-1}. Entonces

$$\begin{aligned} X(s_i) &= \langle A^t A s_i, s_i \rangle \\ &= \lambda_i \langle s_i, s_i \rangle \\ &= \lambda_i \|s_i\|^2 \\ &= \lambda_i \end{aligned}$$

y por lo tanto

$$\max_{x \in S^{m-1}} X(x) = \lambda_m \quad \text{y} \quad \min_{x \in S^{m-1}} X(x) = \lambda_1.$$

Es claro que $\lambda_m = \|A\|^2$ pues

$$\begin{aligned} \|A\|^2 &\stackrel{def}{=} \max_{\|x\|=1} \|Ax\|^2 \\ &= \max_{\|x\|=1} X(x) \\ &= \lambda_m. \end{aligned}$$

Veamos ahora que si $\lambda_1 > 0$ entonces $\lambda_1 = \frac{1}{\|A^{-1}\|^2}$.
Es fácil ver que si A^{-1} es no singular (o sea $\lambda_1 > 0$) entonces los valores propios de $A^t A$ son los mismos que los de AA^t. Ésto es consecuencia de la siguiente igualdad:

$$AA^t - \lambda I = A(A^t A - \lambda I)A^{-1}.$$

Como $(A^{-1})^t A^t = (AA^{-1})^t = I$ se tiene que $(A^{-1})^t = (A^t)^{-1}$, y por lo tanto

$$(A^{-1})^t A^{-1} = (A^t)^{-1} A^{-1} = (AA^t)^{-1}.$$

De la anterior igualdad se concluye que los valores propios de $(A^{-1})^t A^{-1}$ son exactamente los inversos de los valores propios de $A^t A$, o sea $\lambda_1^{-1}, \ldots, \lambda_m^{-1}$, y si aplicamos el resultado anterior obtenemos que $\|A^{-1}\|^2$ es igual al mayor valor propio asociado a $(A^{-1})^t A^{-1}$, es decir

$$\|A^{-1}\|^2 = \frac{1}{\lambda_1},$$

lo que concluye la prueba. □

Observación 5. Si continuamos con esta notación, podemos reescribir la definición de número de condición dada en (1.1)

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| = \frac{\sqrt{\lambda_m}}{\sqrt{\lambda_1}} = \sqrt{\frac{\text{máx}_{x\in S^{m-1}} X(x)}{\text{mín}_{x\in S^{m-1}} X(x)}}$$

siendo $X(x) = x^t A^t A x$ $(x \in \mathbb{R}^m)$.
De esta manera para nuestros fines de demostrar el Teorema 6 nos basaremos en las distribución conjunta de (λ_1, λ_m).

4.5.2. Demostración

En esta monografía sólo incluiremos la demostración de la cota superior de la distribución de la cola de $\kappa(A)$ en el Teorema 6.

Prueba del Teorema 6. Probaremos que

$$P(\kappa(A) > x) < \frac{Cm}{x} \qquad \text{para } x > m.$$

La prueba se dividirá en varios pasos. Nuestro objetivo principal es estimar la densidad conjunta de (λ_m, λ_1) como ya observamos.

Paso 1:

Recordemos la notación, $X : S^{m-1} \to \mathbb{R}$, $X(x) = x^t A^t A x$ es decir $X(x) = \|Ax\|^2$. Notaremos $A^t A$ como B.

Dados a, $b \in \mathbb{R}$ con $a > b$ queremos calcular la probabilidad del conjunto

$$\{\lambda_m \in (a, a + da),\ \lambda_1 \in (b, b + db)\}. \tag{4.13}$$

Para eso tratemos de describir dicho conjunto en función de los vectores propios. Tratemos de identificar los vectores propios de norma uno asociados a λ_m y λ_1 que los llamaremos s y t respectivamente. Por el Lema 7 sabemos que casi seguramente los valores propios son todos distintos. Entonces con probabilidad uno es fácil probar que en el caso de que la matriz sea autoadjunta, se tiene que s y t son ortogonales:
-Sean v y w vectores propios con valores propios λ y μ respectivamente. Entonces usando el hecho de que B es simétrica se tiene

$$\langle v, w\rangle = \frac{1}{\lambda}\langle Bv, w\rangle = \frac{1}{\lambda}\langle v, B^t w\rangle = \frac{1}{\lambda}\langle v, Bw\rangle = \frac{\mu}{\lambda}\langle v, w\rangle$$

y si $\lambda \neq \mu$ se tiene que $\langle v, w\rangle = 0$.

Sea $\pi_u : \mathbb{R}^m \to \mathbb{R}^m$ la proyección ortogonal sobre el subespacio $\{u\}^\perp$, es decir $\pi_u(w) = w - \langle w, u\rangle\, u$ con u de norma uno, entonces si $\pi_u(Bu) = 0$ se tiene que $Bu - \langle Bu, u\rangle u = 0$ o sea $Bu = \langle Bu, u\rangle u$ que no es otra cosa que u sea un vector propio de B. Por lo tanto debemos pedir que $\pi_s(B_s) = 0$ $\pi_t(B_t) = 0$. Lo único que nos esta faltando es que los vectores s y t tengan los valores propios máximos y mínimos respectivamente. Y para eso solo hay que chequear que son los vectores que verifican $X''(s) \prec 0$ y $X''(t) \succ 0$.
Entonces hemos probado que casi seguramente (4.13) es igual a

$$\Big\{\exists\, s,\, t \in S^{m-1},\ \langle s, t\rangle = 0,\ X(s) \in (a, a+da),\ X(t) \in (b, b+db),$$

$$\pi_s(B_s) = 0,\ \pi_t(B_t) = 0,\ X''(s) \prec 0 \text{ y } X''(t) \succ 0\Big\}. \tag{4.14}$$

Si llamamos $N_{a,b,da,db}$ a la variable aleatoria que cuenta el número de pares (s, t) que pertenecen a (4.14), podemos observar que casi seguramente solo puede tomar los valores 0 o 4, ya que si existen, aparecen de a pares por ser X simétrica. Entonces

$$E\big(N_{a,b,da,db}\big) = 4\, P\big(\{\lambda_m \in (a, a+da),\ \lambda_1 \in (b, b+db)\}\big),$$

o sea

$$P\big(\{\lambda_m \in (a, a+da),\ \lambda_1 \in (b, b+db)\}\big) = \frac{1}{4}\, E\big(N_{a,b,da,db}\big). \tag{4.15}$$

Paso 2:

Acá daremos una cota para $E\big(N_{a,b,da,db}\big)$.
Sea

$$V = \{(s, t) :\ s,\, t \in S^{m-1},\ \langle s, t\rangle = 0\} \subset \mathbb{R}^{2m}.$$

Los puntos de V, verifican las siguientes condiciones:

$$(A) = \begin{cases} \|s\| = 1 \\ \|t\| = 1 \\ \langle s, t\rangle = 0 \end{cases}$$

y si consideramos las funciones

$$\begin{aligned} g_1(s,t) &= s_1^2 + \cdots + s_m^2 + 0 + \cdots + 0 \\ g_2(s,t) &= \ 0 + \cdots + \ 0 \ + t_1^2 + \cdots t_m^2 \\ g_3(s,t) &= s_1 t_1 + \cdots s_m t_m \end{aligned}$$

el sistema de ecuaciones (A) lo podemos escribir como

$$(A) = \begin{cases} g_1(s,t) = 1 \\ g_2(s,t) = 1 \\ g_3(s,t) = 0. \end{cases}$$

Sea $G : \mathbb{R}^{2m} \to \mathbb{R}^3$ definida por,

$$G(s,t) = \begin{pmatrix} g_1(s,t) \\ g_2(s,t) \\ g_3(s,t) \end{pmatrix}$$

entonces la matriz asociada en la base canónica del diferencial de G en el punto (s,t) es

$$dG_{(s,t)} = \begin{pmatrix} 2s_1 & \dots & 2s_m & 0 & \dots & 0 \\ 0 & \dots & 0 & 2t_1 & \dots & 2t_m \\ t_1 & \dots & t_m & s_1 & \dots & s_m \end{pmatrix}$$

que es una matriz de $3 \times 2m$. Veamos que el rango es 3 y por tanto de rango máximo. Para eso podemos observar que basta probar que el det $\left(dG_{(s,t)}\big(dG_{(s,t)}\big)^t\right) \neq 0$. Notemos $G'(s,t)$ a la matriz asociada en la base canónica a $dG_{(s,t)}$ o sea basta probar que det $\left(G'(s,t)\big(G'(s,t)\big)^t\right) \neq 0$ donde $G'(s,t)\big(G'(s,t)\big)^t$ es una matriz 3×3.

$$\det\left(G'(s,t)\big(G'(s,t)\big)^t\right) = \left| \begin{pmatrix} 2s & 0 \\ 0 & 2t \\ t & s \end{pmatrix} \cdot \begin{pmatrix} 2s & 0 & t \\ 0 & 2t & s \end{pmatrix} \right|$$

$$= \left| \begin{pmatrix} 2^2\langle s,s\rangle & 0 & 2\langle s,t\rangle \\ 0 & 2^2\langle t,t\rangle & 2\langle t,s\rangle \\ 2\langle t,s\rangle & 2\langle t,s\rangle & \langle t,t\rangle + \langle s,s\rangle \end{pmatrix} \right| = \left| \begin{pmatrix} 2^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \right|$$

entonces

$$\det\left(G'(s,t)\big(G'(s,t)\big)^t\right) = 32 \neq 0.$$

Resumiendo, tenemos una función $G : \mathbb{R}^{2m} \to \mathbb{R}^3$ donde $G'(s,t)$ tiene rango máximo y

$$V = G^{-1}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

entonces por el teorema de la función implícita, se tiene que V es una variedad diferenciable de dimensión $2m-3$, de clase C^∞ (ya que G lo es), encajada en $\mathbb{R}^{2m}$ y sin borde.

Llamaremos $\tau = (s,t)$ un punto genérico en V y $\sigma_V(d\tau)$ la medida geométrica en V. Se prueba que

$$\sigma_V(V) = \sqrt{2}\sigma_{m-1}\sigma_{m-2} \tag{4.16}$$

donde σ_{m-1} es el área de la superficie $S^{m-1} \subset \mathbb{R}^m$, es decir $\sigma_{m-1} = \frac{2\pi^{m/2}}{\Gamma(m/2)}$. Definimos en V el siguiente campo aleatorio $Y : V \to \mathbb{R}^{2m}$, como

$$Y(s,t) = \begin{pmatrix} \pi_s(Bs) \\ \pi_t(Bt) \end{pmatrix}.$$

Dado $\tau = (s,t) \in V$, tenemos que $Y(\tau) \in \left\{ \{s\}^\perp \times \{t\}^\perp \right\}$ y como

$$\begin{aligned}
&\langle Y(s,t),(t,-s)\rangle_{\mathbb{R}^{2m}} \\
&\quad = \langle \pi_s(Bs), t\rangle_{\mathbb{R}^m} - \langle \pi_t(Bt), s\rangle_{\mathbb{R}^m} \\
&\quad = \langle Bs - \langle Bs, s\rangle_{\mathbb{R}^m} s, t\rangle_{\mathbb{R}^m} - \langle Bt - \langle Bt, t\rangle_{\mathbb{R}^m} t, s\rangle_{\mathbb{R}^m} \\
&\quad = \langle Bs, t\rangle_{\mathbb{R}^m} - \langle Bs, s\rangle_{\mathbb{R}^m} \cdot \overbrace{\langle s, t\rangle_{\mathbb{R}^m}}^{=0} - \langle Bt, s\rangle_{\mathbb{R}^m} + \langle Bt, t\rangle_{\mathbb{R}^m} \cdot \overbrace{\langle t, s\rangle_{\mathbb{R}^m}}^{=0} \\
&\quad = \langle Bs, t\rangle_{\mathbb{R}^m} - \langle Bt, s\rangle_{\mathbb{R}^m} \\
&\quad = \langle s, B^t t\rangle_{\mathbb{R}^m} - \langle s, Bt\rangle_{\mathbb{R}^m} \\
&\quad = \langle s, Bt\rangle_{\mathbb{R}^m} - \langle s, Bt\rangle_{\mathbb{R}^m} \\
&\quad = 0,
\end{aligned}$$

entonces

$$Y(\tau) \in \{(t,-s)\}^\perp.$$

Sea

$$W_\tau = \left\{(t,-s)\right\}^\perp \cap \left\{ \{s\}^\perp \times \{t\}^\perp \right\} \subset \mathbb{R}^{2m}.$$

Como $\{s\}^\perp \subset \mathbb{R}^m$ tiene dimensión $m-1$ entonces $\{s\}^\perp \times \{t\}^\perp$ es un subespacio vectorial de $\mathbb{R}^{2m}$ de dimensión $2(m-1) = 2m-2$ y como $(t,-s) \in \{s\}^\perp \times \{t\}^\perp$ tenemos que $\left\{(t,-s)\right\}^\perp \cap \left\{ \{s\}^\perp \times \{t\}^\perp \right\}$ tiene dimensión $2m-3$ o sea

$$\dim(W_\tau) = 2m-3.$$

Definamos ahora

$$\Delta(\tau) = \left(\det\left[(Y'(\tau))^t \cdot Y'(\tau)\right]\right)^{1/2} \quad \text{y} \quad N = \#\{\tau : \tau \in V,\ Y(\tau) = 0\}$$

y para $\tau \in V$

$$F_\tau = \{X(s) \in (a, a+da),\ X(t) \in (b, b+db),\ X''(s) \prec 0,\ X''(t) \succ 0\}. \tag{4.17}$$

También definamos $p_{Y(\tau)}(\cdot)$ como la densidad del vector aleatorio $Y(\tau)$ en el subespacio vectorial $W_\tau \subset \mathbb{R}^{2m}$ de dimensión $2m-3$.
Veremos más adelante, cuando computemos $Y'(s,t)$, que podemos asumir que 0 es un valor regular de Y, pues lo es con probabilidad uno.
Como

$$V = G^{-1}\begin{pmatrix}1\\1\\0\end{pmatrix} \subset S^{m-1}\times S^{m-1},$$

se tiene que V es compacto (cerrado en un compacto), entonces eso implica que $N<\infty$. De lo contrario acumularían y 0 dejaría de ser un valor regular[3].
Sean $\tau_1,\dots,\tau_N$ las raíces de $Y(\tau)=0$ en V. Por el Teorema de la Función Inversa podemos encontrar en V entornos abiertos $U_1,\dots,U_N$ de $\tau_1,\dots,\tau_N$ respectivamente, tales que, para $\delta>0$ suficientemente pequeño se verifican:

- Y es un difeomorfismo local de U_j en $Y(V)\cap B_{2m}(0,\delta)$ siendo $B_{2m}(0,\delta)$ la bola Euclideana en $\mathbb{R}^{2m}$ de centro 0 y radio δ,
- $U_1,\dots,U_N$ son dos a dos disjuntos,
- si $\tau\in V\setminus\bigcup_{j=1}^N U_j$ entonces $Y(\tau)\notin B_{2m}(0,\delta)$.

Usando la fórmula de cambio de variable, tenemos

$$\begin{aligned}\int_V \Delta(\tau)\, Ind_{\{\|Y(\tau)\|<\delta\}}\,\sigma_V(d\tau) &= \sum_{i=1}^N\int_{U_i}\Delta(\tau)\,\sigma_V(d\tau) &(4.18)\\ &= \sum_{i=1}^N \mu(Y(U_i)) &(4.19)\end{aligned}$$

donde $\mu(Y(U_i))$ es la medida geométrica $2m-3$ dimensional de $Y(U_i)$ (que es una variedad diferenciable de dimensión $2m-3$ como ya observamos).

Si tomamos $\delta\downarrow 0$ entonces $\mu(Y(U_i))\sim|B_{2m-3}(0,\delta)|$ siendo $|B_{2m-3}(0,\delta)|$ la medida de Lebesgue de la bola en $\mathbb{R}^{2m-3}$.
Entonces dividendo por $|B_{2m-3}(0,\delta)|$ en (4.18) y tomando límite se sigue que:

$$N=\lim_{\delta\downarrow 0}\frac{1}{|B_{2m-3}(0,\delta)|}\int_V\Delta(\tau)\,Ind_{\{\|Y(\tau)\|<\delta\}}\,\sigma_V(d\tau).$$

De la misma manera se prueba que:

$$N_{a,b,da,db}=\lim_{\delta\downarrow 0}\frac{1}{|B_{2m-3}(0,\delta)|}\int_V\Delta(\tau)\,Ind_{F_\tau}\,Ind_{\{\|Y(\tau)\|<\delta\}}\,\sigma_V(d\tau). \tag{4.20}$$

[3]Argumento ya usado en la prueba de la Fórmula de Rice.

Aplicando Fatou en (4.20) se tiene

$$E\big(N_{a,b,da,db}\big) \leq \liminf_{\delta \downarrow 0} \frac{1}{|B_{2m-3}(0,\delta)|} E\left(\int_V \Delta(\tau)\, Ind_{F_\tau}\, Ind_{\{\|Y(\tau)\|<\delta\}}\, \sigma_V(d\tau) \right)$$

y luego usando Fubini

$$E\big(N_{a,b,da,db}\big) \leq \liminf_{\delta \downarrow 0} \frac{1}{|B_{2m-3}(0,\delta)|} \int_V E\left(\Delta(\tau)\, Ind_{F_\tau}\, Ind_{\{\|Y(\tau)\|<\delta\}}\right)\, \sigma_V(d\tau),$$

o sea

$$\begin{aligned}
&E\big(N_{a,b,da,db}\big) \\
&\leq \liminf_{\delta \downarrow 0} \frac{1}{|B_{2m-3}(0,\delta)|} \int_V E\left(\Delta(\tau)\, Ind_{F_\tau}\, Ind_{\{\|Y(\tau)\|<\delta\}}\right)\, \sigma_V(d\tau) \\
&= \liminf_{\delta \downarrow 0} \frac{1}{|B_{2m-3}(0,\delta)|} \cdot \\
&\qquad \int_V \left[\int\limits_{B_{2m-3}(0,\delta)\cap W_\tau} E\left(\Delta(\tau)\, Ind_{F_\tau} \mid Y(\tau)=y\right)\, p_{Y(\tau)}(y)\, dy \right] \sigma_V(d\tau) \\
&= \liminf_{\delta \downarrow 0} \int_V \sigma_V(d\tau) \left(\frac{1}{|B_{2m-3}(0,\delta)|} \cdot \right. \\
&\qquad \left. \int\limits_{B_{2m-3}(0,\delta)\cap W_\tau} E\left(\Delta(\tau)\, Ind_{F_\tau} \mid Y(\tau)=y\right)\, p_{Y(\tau)}(y)\, dy \right) \\
&= \int_V E\left(\Delta(\tau)\, Ind_{F_\tau} \mid Y(\tau)=0\right)\, p_{Y(\tau)}(0)\, \sigma_V(d\tau)
\end{aligned}$$

donde la última igualdad se debe a que la segunda integral es una función continua del par (τ, y) (más adelante quedara claro la continuidad).[4] Por lo tanto por la definición de F_τ se tiene

$$\begin{aligned}
E\big(N_{a,b,da,db}\big) \leq \int_a^{a+da} dx \int_b^{b+db} dy \int_V E\big(\Delta(s,t)\, Ind_{\{X''(s)\prec 0,\, X''(t)\succ 0\}} \mid \\
X(s)=x,\ X(t)=y,\ Y(s,t)=0\big)\, p_{(X(s),X(t),Y(s,t))}(x,y,0)\, \sigma_V(d(s,t)).
\end{aligned} \tag{4.21}$$

Más tarde veremos que el integrando en (4.21) es constante (no depende de (s,t))

[4]En realidad lo que acabamos de hacer es usar una fórmula de Rice modificada, pues calculamos la esperanza del número de raíces de Y en V ($N_0^Y(V)$) que verifícan el evento F_τ. La prueba de esta nueva versión de la fórmula de Rice no asume cambios significativos en la demostración ya hecha.

como consecuencia de que la matriz A es invariante bajo isometrías de $\mathbb{R}^m$, entonces

$$E\big(N_{a,b,da,db}\big) \leq \int_a^{a+da} dx \int_b^{b+db} dy \int_V E\big(\Delta(e_1,e_2)\, Ind_{\{X''(e_1)\prec 0,\, X''(e_2)\succ 0\}} \mid \\ X(e_1)=x,\ X(e_2))=y,\ Y(e_1,e_2)=0\big)\, p_{(X(e_1),X(e_2),Y(e_1,e_2))}(x,y,0)\, \sigma_V(d(e_1,e_2)). \tag{4.22}$$

De esta manera por (4.15), (4.22), y (4.16) hemos probado que la densidad conjunta de λ_m y λ_1 que llamaremos $h(a,b)$ con $a > b$ verifica la siguiente desigualdad:

$$h(a,b) \leq \frac{\sqrt{2}}{4}\sigma_{m-1}\sigma_{m-2}E\left(\Delta(e_1,e_2)\, Ind_{\{X''(e_1)\prec 0,\, X''(e_2)\succ 0\}} \mid X(e_1)=a,\right. \\ \left. X(e_2))=b,\ Y(e_1,e_2)=0\right)\, p_{(X(e_1),X(e_2),Y(e_1,e_2))}(a,b,0). \tag{4.23}$$

Paso 3:

En lo que sigue vamos a computar cada factor del lado derecho de (4.23). Antes de seguir necesitamos explicitar algunos cálculos, que con el fin de mantener una línea en la demostración los hacemos en forma de lema.

Lema 9.

$$X''(e_1) = 2(B_1 - b_{11}I_{m-1})$$
$$X''(e_2) = 2(B_2 - b_{22}I_{m-1}).$$

Demostración. Sea $Z : \mathbb{R}^m \to \mathbb{R}$ dado por $Z(x) = x^tBx = \langle Bx, x\rangle$. La derivada libre $Z'(x) : \mathbb{R}^m \to \mathbb{R}$ ($Z' : \mathbb{R}^m \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}) = (\mathbb{R}^m)^*$) (que es $\nabla_x Z$)

$$Z'(x) = 2(\langle Be_1, x\rangle, \ldots, \langle Be_m, x\rangle) \tag{4.24}$$
$$= 2(b_{11}x_1 + b_{21}x_2 + \ldots + b_{m1}x_m, \ldots, b_{1m}x_1 + \ldots + b_{mm}x_m) \tag{4.25}$$

y $Z''(x) : \mathbb{R}^m \to \mathbb{R}^m$ ($Z'' : \mathbb{R}^m \to \mathscr{L}(\mathbb{R}^m, \mathbb{R}^m)$) esta dada por $Z''(x) = 2B$.
Por definición $X = Z|_{S^{m-1}}$.
Sea $x = (x_1, \ldots, x_m) \in S^{m-1}$ entonces $x_1^2 + \cdots + x_m^2 = 1$ y por tanto podemos escribir x_1 en función de las demás variables o sea

$$x_1 = \sqrt{1 - (x_2^2 + \cdots + x_m^2)} \quad \text{para} \quad (x_2, \ldots .x_m) \sim (0, \ldots, 0).$$

Sea $\varphi : U \to S^{m-1}$ con $U \subset \mathbb{R}^{m-1}$ una parametrización local de $e_1 \in S^{m-1}$ dada por

$$\varphi(x_2, \ldots, x_m) = \left(\sqrt{1 - (x_2^2 + \cdots + x_m^2)}, x_2, \ldots, x_m\right)$$

$$\begin{tikzcd} S^{m-1} \arrow[r, "X"] & \mathbb{R} \\ U \arrow[u, "\varphi"] \arrow[ur, "\eta"'] \arrow[ur, phantom, "\circlearrowleft", shift left=3] & \end{tikzcd}$$

siendo $\eta = X \circ \varphi$.
Sea $u \in U$ entonces se tiene

$$d_u\eta \overset{def}{=} d_{\varphi^{-1}(u)}(X \circ \varphi) = d_{\varphi^{-1}(u)}(Z \circ \varphi)$$

con $Z \circ \varphi : \overbrace{U}^{\subset \mathbb{R}^{m-1}} \to \mathbb{R}$ definido como

$$Z \circ \varphi(x_2, \ldots, x_m) = Z\left(\sqrt{1-(x_2^2+\cdots+x_m^2)}, x_2, \ldots, x_m\right).$$

Sea $u = (x_2, \ldots, x_m)$ entonces

$$d_u(Z \circ \varphi) = d_{\varphi(u)}Z \circ d_u\varphi$$

Notemos por $x = \varphi(u)$. Tenemos que $d_xZ = \nabla_x Z$ es una matriz $1 \times m$ como vimos en (4.24) y $d_u\varphi$ es una matriz $m \times (m-1)$ entonces $d_u(Z \circ \varphi) = d_xZ \circ d_u\varphi$ es una matriz $1 \times (m-1)$.
Calculemos $d_u\varphi$:

$$d_u\varphi = \left(\begin{array}{ccc} -\frac{x_2}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} & \cdots & -\frac{x_m}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} \\ \hline & & \\ & I_{m-2} & \\ & & \end{array}\right)$$

siendo I_{m-2} la matriz identidad de dimension $m-2$. Entonces se tiene que $d_xZ \circ d_u\varphi$ es igual a

$$2(\langle Be_1, x\rangle, \ldots, \langle Be_m, x\rangle) \cdot \left(\begin{array}{ccc} -\frac{x_2}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} & \cdots & -\frac{x_m}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} \\ \hline & & \\ & I_{m-2} & \\ & & \end{array}\right)$$

$$\begin{aligned} = 2\Bigg(&\langle b_1, x\rangle \frac{-x_2}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} + \langle b_2, x\rangle, \ldots, \\ &\ldots, \langle b_1, x\rangle \frac{-x_m}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} + \langle b_m, x\rangle\Bigg), \end{aligned} \tag{4.26}$$

llamando a

$$b_i = Be_i = \begin{pmatrix} b_{1i} \\ \vdots \\ b_{mi} \end{pmatrix}.$$

Ordenando en (4.26) se obtiene

$$2\frac{\langle b_1, x\rangle}{\sqrt{1-(x_2^2+\cdots+x_m^2)}}(-x_2,\ldots,-x_m) + (\langle b_2, x\rangle, \ldots, \langle b_m, x_m\rangle)$$

y entonces se tiene

$$\begin{aligned} X'(x) = 2\frac{\langle b_1, x\rangle}{\sqrt{1-(x_2^2+\cdots+x_m^2)}}&(-x_2,\ldots,-x_m)+ \\ &2\left(\langle b_2, x\rangle, \ldots, \langle b_m, x_m\rangle\right) \end{aligned} \tag{4.27}$$

con $x = \left(\sqrt{1-(x_2^2+\cdots+x_m^2)}, x_2, \ldots, x_m\right)$.
Las derivadas parciales son

$$\frac{\partial X}{\partial x_j}(x) = \frac{\langle b_1, x\rangle}{\sqrt{1-(x_2^2+\cdots+x_m^2)}}(-x_j) + 2\langle b_j, x\rangle,$$

es decir

$$\frac{\partial X}{\partial x_j}(x) = 2\left(\sum_{k=1}^{m} b_{k1}x_k\right)\frac{(-x_j)}{\sqrt{1-(x_2^2+\cdots+x_m^2)}} + \sum_{k=1}^{m} b_{kj}x_k \tag{4.28}$$

$$= 2\left[-b_{11}x_j + \left(\sum_{k=2}^{m} b_{k1}x_k\right)\frac{(-x_j)}{\sqrt{1-(x_2^2+\cdots+x_m^2)}}\right. \tag{4.29}$$

$$\left. + \sum_{k=1}^{m} b_{kj}x_k\right] \tag{4.30}$$

con $j = 2\ldots m$. Entonces si derivamos (4.28) respecto a x_i se tiene

$$\frac{\partial^2 X}{\partial x_i \partial x_j}(e_1) = 2(-b_{11}+b_{ij}) \quad \text{con} \quad i,j = 2,\ldots,m$$

y por lo tanto se tiene

$$X''(e_1) = 2\left(B_1 - b_{11}I_{m-1}\right),$$

análogamente para e_2 se obtiene

$$X''(e_2) = 2\left(B_2 - b_{22}I_{m-1}\right)$$

siendo B_1 (respectivamente B_2) es la matriz $(m-1)\times(m-1)$ que resulta de suprimir la primera (segunda) fila y columna de B. $\square$

Tomemos la siguiente base ortonormal del subespacio W_τ con $\tau = (e_1, e_2)$

$$\mathscr{B}_1 = \{(e_3, 0), \ldots, (e_m, 0), (0, e_3), \ldots, (0, e_m), \frac{1}{\sqrt{2}}(e_2, e_1)\}$$

(es fácil chequear que es una base ortonormal).
Tenemos que $X(e_1) = \langle Be_1, e_1\rangle = b_{11}$ y $X(e_2) = \langle Be_2, e_2\rangle = b_{22}$. Entonces tenemos lo siguiente:

$$\begin{cases} X(e_1) &= b_{11} \\ X(e_2) &= b_{22} \\ X''(e_1) &= 2\,(B_1 - b_{11}I_{m-1}) \\ X''(e_2) &= 2\,(B_2 - b_{22}I_{m-1}) \end{cases}$$

por lo tanto podemos re-escribir la esperanza condicional en (4.23) como

$$E\left(\Delta(e_1, e_2)\, Ind_{\{B_1 - b_{11}I_{m-1}\prec 0,\, B_2 - b_{22}I_{m-1}\succ 0\}} \,\middle|\, b_{11} = a,\, b_{22} = b,\, Y(e_1, e_2) = 0\right). \quad (4.31)$$

Recordando la definición de $Y(s, t)$, se tiene que

$$Y(e_1, e_2) = (\overbrace{0, b_{21}, \ldots, b_{m1}}^{\pi_{e_1}(Be_1)}, \overbrace{b_{12}, 0, b_{32}, \ldots, b_{m2}}^{\pi_{e_2}(Be_2)})^t$$

por lo tanto $Y(e_1, e_2)$ tiene una expresión en la base ortonormal $\mathscr{B}_1$:

$$Y(e_1, e_2) = \sum_{i=3}^{m}(b_{i1}(e_i, 0) + b_{i2}(0, e_i)) + \sqrt{2}\left(\frac{1}{\sqrt{2}}(e_2, e_1)\right).$$

Se sigue que la densidad conjunta de $X(e_1), X(e_2), Y(e_1, e_2)$ que aparece en (4.23) en el espacio $\mathbb{R} \times \mathbb{R} \times W_{(e_1, e_2)}$ es la densidad conjunta de las variables aleatorias

$$\overbrace{b_{11}}^{X(e_1)}, \overbrace{b_{22}}^{X(e_2)}, \overbrace{\sqrt{2}b_{12}, b_{31}, \ldots, b_{m1}, b_{32}, \ldots, b_{m2}}^{Y(e_1,e_2)}$$

tomada en el punto $(a, b, 0, \ldots\ldots, 0)$.
Primero computemos la densidad conjunta de

$$b_{31}, \ldots, b_{m1}, b_{32}, \ldots, b_{m2}$$

dados a_1, a_2, donde a_1 (respectivamente a_2) es la primera (segunda) columna de la matriz A. Llamaremos q a dicha densidad condicional.
Como A es Gaussiana, entonces a_1 y a_2 son vectores aleatorios Gaussianos estándares en $\mathbb{R}^m$.

Sean $X = (X_3, \ldots, X_m)$ y $Y = (Y_3, \ldots, Y_m)$ con $X_i = b_{i1}$ y $Y_i = b_{i2}$.
Como $b_{ij} = \langle a_i, a_j \rangle$ se tiene

$$\begin{array}{rcl} X_i & = & a_{1i}\,\underline{\underline{a_{11}}} + \cdots + a_{mi}\,\underline{\underline{a_{m1}}} \\ Y_i & = & a_{1i}\,\underline{\underline{a_{12}}} + \cdots + a_{mi}\,\underline{\underline{a_{m2}}} \end{array}$$

donde el doble subrayado nos indica que están dados.
Es claro que la distribución de X_i dado a_1 con $i = 3, \ldots, m$ es la distribución de una variable aleatoria con distribución $\mathcal{N}(0, \sum_{j=1}^{m} a_{j1}^2) = \mathcal{N}(0, \|a_1\|^2)$ y análogamente la distribución de Y_i dado a_2 con $i = 3, \ldots, m$ es la distribución de una variable aleatoria con distribución $\mathcal{N}(0, \|a_2\|^2)$. Entonces se concluye que q es una densidad normal en $\mathbb{R}^{2(m-2)}$ con esperanza 0 y matriz de covarianza $\Sigma = E\left[(X,Y)^t \cdot (X,Y)\right]$

$$\Sigma = \begin{pmatrix} X_1 \\ \vdots \\ X_{m-2} \\ Y_1 \\ \vdots \\ Y_{m-2} \end{pmatrix} \cdot \Big(X_1, \ldots, X_{m-2}, Y_1, \ldots, Y_{m-2} \Big)$$

$$= E \left[\begin{array}{ccc|ccc} X_1^2 & \cdots & X_1 X_{m-2} & X_1 Y_1 & \cdots & X_1 Y_{m-2} \\ \vdots & \ddots & & \vdots & \ddots & \\ X_{m-2} X_1 & \cdots & X_{m-2}^2 & X_{m-2} Y_1 & \cdots & X_{m-2} Y_{m-2} \\ \hline Y_1 X_1 & \cdots & Y_1 X_{m-2} & Y_1^2 & \cdots & Y_1 Y_{m-2} \\ \vdots & \ddots & & \vdots & \ddots & \\ Y_{m-2} X_1 & \cdots & Y_{m-2}^2 & Y_{m-2} Y_1 & \cdots & Y_{m-2}^2 \end{array} \right]$$

entonces

$$\Sigma = \left[\begin{array}{c|c} \|a_1\|^2 I_{m-2} & E(X^t Y) \\ \hline E(X^t Y) & \|a_2\|^2 I_{m-2} \end{array} \right].$$

Calculemos $E(X^t Y)$:

$$\begin{array}{rcl} E(X_i Y_j) & = & E\left((a_{1i}\, a_{11} + \cdots + a_{mi}\, a_{m1})(a_{1j}\, a_{12} + \cdots + a_{mj}\, a_{m2}) \right) \\ & = & \delta_{ij} E(a_{1i}^2\, a_{11}\, a_{12} + \cdots + a_{mi}^2\, a_{m1}\, a_{m2}) \\ & = & \delta_{ij} \langle a_1, a_2 \rangle \end{array}$$

donde las igualdades resultan de la independencia de los coeficientes de A y de que tienen distribución normal estándar, entonces

$$E(X^t Y) = \langle a_1, a_2 \rangle\, I_{m-2}. \tag{4.32}$$

Concluimos de (4.32) que

$$\Sigma = \left[\begin{array}{c|c} \|a_1\|^2 I_{m-2} & \langle a_1, a_2\rangle\, I_{m-2} \\ \hline \langle a_1, a_2\rangle\, I_{m-2} & \|a_2\|^2 I_{m-2} \end{array}\right].$$

Hemos concluido que la densidad conjunta q de $(b_{31}, \ldots, b_{m1}, b_{32}, \ldots, b_{m2})$ dados a_1 y a_2 es la densidad de una normal con esperanza 0 y varianza Σ.

Calculemos ahora la densidad de $\big(b_{11}, b_{22}, b_{12}\big) = (\|a_1\|^2, \|a_2\|^2, \|a_1\|\,\|a_2\|\,\langle \widetilde{a_1}, \widetilde{a_2}\rangle)$ en el punto $(a, b, 0)$ siendo $\widetilde{a_i} = \frac{a_i}{\|a_i\|}$.
Para eso probemos lo siguiente:

Lema 10. *Sea ξ un vector aleatorio con distribución $\mathcal{N}(0, I_m)$ en $\mathbb{R}^m$ entonces*

$$\|\xi\| \perp \frac{\xi}{\|\xi\|}$$

donde el símbolo $\perp$ quiere decir que son independientes.

Demostración. Calculemos $P\left(\|\xi\| \le x,\ \frac{\xi}{\|\xi\|} \in A\right)$ con $A \subset S^{m-1}$.

$$\begin{aligned} P\left(\|\xi\| \le x,\ \frac{\xi}{\|\xi\|} \in A\right) &= P\left(\xi \in K\right) \\ &= \int_K \frac{1}{(2\pi)^{m/2}}\, e^{-\frac{1}{2}\|z\|^2}\, dz \\ &= \underbrace{\int_A \sigma_{m-1}(du)}_{\text{uniforme en la esfera}} \quad \overbrace{\int_0^x \rho^{m-1} \frac{1}{(2\pi)^{m/2}}\, e^{-\frac{1}{2}\rho^2}\, d\rho}^{\sqrt{\chi^2_m(\rho)}} \end{aligned}$$

entonces se tiene que

$$P\left(\|\xi\| \le x,\ \frac{\xi}{\|\xi\|} \in A\right) = P\left(\frac{\xi}{\|\xi\|} \in A\right) \cdot P(\|\xi\| \le x).$$

□

Con este lema tenemos que $(\|a_1\|, \|a_2\|)$ es independiente de $(\widetilde{a_1}, \widetilde{a_2})$, entonces $(\|a_1\|, \|a_2\|) \perp \langle \widetilde{a_1}, \widetilde{a_2}\rangle$. Entonces, notando por $f(W)$ la densidad de la variable W obtenemos

$$f\big(b_{11}, b_{22}, b_{12}\big) = f(b_{11}) f(b_{22}) f\big(b_{12} \mid b_{11}, b_{22}\big),$$

donde

$$f(\|a_1\| \cdot \|a_2\| \langle \widetilde{a_1}, \widetilde{a_2} \rangle \mid \|a_1\|^2, \|a_2\|^2) =$$
$$f(\sqrt{ab}\,\langle \widetilde{a_1}, \widetilde{a_2} \rangle \mid \|a_1\|^2 = a,\ \|a_2\|^2 = b) = f(\sqrt{ab}\,\langle \widetilde{a_1}, \widetilde{a_2} \rangle)$$

por la independencia.
Si observamos que

$$P\left(\sqrt{ab}Z \leq x\right) = P\left(Z \leq \frac{x}{\sqrt{ab}}\right) = \int_{-\infty}^{\frac{x}{\sqrt{ab}}} f_Z(y)\,dy$$

se tiene

$$f_{\sqrt{ab}Z}(x) = \frac{1}{\sqrt{ab}} f_Z(\frac{x}{\sqrt{ab}}),$$

y como $\|a_1\|^2 \sim \chi_m^2$ siendo χ_m^2 la distribución de Fisher con m grados de libertad, y $X = \langle \widetilde{a_1}, \widetilde{a_2} \rangle$ entonces

$$p_{(b_{11},b_{22},b_{12})}(a,b,0) = \chi_m^2(a)\chi_m^2(b)\frac{1}{\sqrt{ab}} p_{\langle \widetilde{a_1}, \widetilde{a_2} \rangle}(0). \tag{4.33}$$

Sea

$$\xi = (\xi_1, \ldots, \xi_m)$$

un vector aleatorio Gaussiano en $\mathbb{R}^m$ estándar. Como $\langle \widetilde{a_1}, \widetilde{a_2} \rangle$ es invariante por rotaciones (por ser producto interno de vectores con distribución normal, que son invariantes por movimientos rígidos) se tiene que $\langle \widetilde{a_1}, \widetilde{a_2} \rangle$ tiene igual distribución que $\frac{\xi_1}{\|\xi\|}$, entonces:

$$\eta(t) = \frac{1}{2t} P(|\langle \widetilde{a_1}, \widetilde{a_2} \rangle| \leq t) = \frac{1}{2t} P\left(\frac{\xi_1^2}{\|\xi\|^2} \leq t^2\right),$$

y como

$$\|\xi\|^2 = \xi_1^2 + \cdots + \xi_m^2 = \xi_1^2 + \zeta \qquad \text{con} \quad \zeta \sim \chi_{m-1}^2$$

se tiene entonces

$$\frac{\xi_1^2}{\xi_1^2 + \chi_{m-1}^2} \leq t^2 \Longrightarrow \frac{\xi_1^2}{\chi_{m-1}^2} \leq \frac{t^2}{1-t^2}.$$

Por lo tanto

$$\eta(t) = \frac{1}{2t} P\left(\frac{\xi_1^2}{\chi_{m-1}^2} \leq \frac{t^2}{1-t^2}\right) \tag{4.34}$$
$$= \frac{1}{2t} P\left(F_{1,m-1} \leq \frac{t^2(m-1)}{1-t^2}\right) \tag{4.35}$$
$$= \frac{1}{2t} \int_0^{\frac{t^2(m-1)}{1-t^2}} f_{1,m-1}(u)\,du \tag{4.36}$$

siendo $F_{1,m-1}$ la distribución de Fisher con $(1, m-1)$ grados de libertad. Se tiene entonces que

$$f_{1,m-1}(u) = C_{m-1}\,\frac{1}{\sqrt{u}}\left(1+\frac{1}{m-1}u\right)^{-m/2}$$

siendo

$$C_{m-1} = \frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{m-1}{2})}\,\frac{1}{\sqrt{m-1}}.$$

Por lo tanto cuando $t \approx 0$

$$\begin{aligned}\eta(t) \approx \frac{1}{2t}\int_0^{\frac{t^2(m-1)}{1-t^2}} C_{m-1}\,\frac{1}{\sqrt{u}}\,du &= \frac{C_{m-1}}{t}\sqrt{\frac{t^2(m-1)}{1-t^2}}\\ &= C_{m-1}\,\frac{\sqrt{m-1}}{\sqrt{1-t^2}},\end{aligned}$$

entonces

$$\begin{aligned}\lim_{t\downarrow 0}\eta(t) &= C_{m-1}\sqrt{m-1}\\ &= \frac{1}{\sqrt{\pi}}\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m-1}{2})}\\ &= p_{\langle\widetilde{a_1},\widetilde{a_2}\rangle}(0)\end{aligned}$$

por lo tanto se tiene

$$p_{\langle\widetilde{a_1},\widetilde{a_2}\rangle}(0) = \frac{1}{\sqrt{\pi}}\frac{\Gamma(\frac{m}{2})}{\Gamma(\frac{m-1}{2})}. \tag{4.37}$$

Concluyendo tenemos que:
-La densidad conjunta de (b_{11}, b_{22}, b_{12}) en el punto $(a, b, 0)$ es

$$\begin{aligned}&\left[\frac{1}{\Gamma(\frac{m}{2})}\left(\frac{1}{2}\right)^{m/2} a^{m/2-1}\,e^{-a/2}\right]\cdot\left[\frac{1}{\Gamma(\frac{m}{2})}\left(\frac{1}{2}\right)^{m/2} b^{m/2-1}\,e^{-b/2}\right]\cdot\\ &\quad\cdot\left[\frac{1}{\sqrt{ab}}\,\frac{1}{\sqrt{\pi}}\,\frac{\Gamma(\frac{m}{2})}{\frac{m-1}{2}}\right]\\ &= \frac{1}{\sqrt{\pi}}\left(\frac{1}{2}\right)^{m/2}(ab)^{m/2-1-1/2}\,e^{-(a+b)/2}\,\frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})},\end{aligned}$$

es decir, la densidad conjunta de (b_{11}, b_{22}, b_{12}) es:

$$\frac{1}{\sqrt{\pi}}\left(\frac{1}{2}\right)^{m/2}\frac{(ab)^{(m-2)/2}}{\sqrt{ab}}\,e^{-(a+b)/2}\,\frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})},$$

y por lo tanto la densidad conjunta de $(b_{11}, b_{22}, \sqrt{2}b_{12})$ es:

$$p_{(b_{11},b_{22},\sqrt{2}b_{12})}(a,b,0) = \frac{1}{\sqrt{2}\sqrt{\pi}} \left(\frac{1}{2}\right)^{m/2} \frac{(ab)^{(m-2)/2}}{\sqrt{ab}} e^{-(a+b)/2} \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})}. \tag{4.38}$$

Lo que estamos buscando es $p_{(X(e_1),X(e_2),Y(e_1,e_2))}(a,b,\ldots\ldots,0)$, es decir

$$p_{(b_{11},b_{22},\sqrt{2}b_{12},b_{31},\ldots,b_{m1},b_{32},\ldots,b_{m2})}(a,b,0,\ldots\ldots,0), \tag{4.39}$$

pero (4.39) es igual al producto de las densidades

$$p_{(b_{11},b_{22},\sqrt{2}b_{12})}(a,b,0) \cdot \underbrace{p_{(b_{31},\ldots,b_{m1},b_{32},\ldots,b_{m2})}(0,\ldots,0 \mid b_{11}=a,\ b_{22}=b,\ b_{12}=0)}_{q(a,b,0)}$$

donde q era una densidad normal con esperanza 0 y varianza Σ dada anteriormente, es decir

$$q(a,b,0) = \frac{1}{(2\pi)^{2(m-2)/2}\left(\det(\Sigma')\right)^{1/2}}$$

siendo

$$\Sigma' = \begin{pmatrix} aI_{m-2} & 0 \\ 0 & bI_{m-2} \end{pmatrix}.$$

Entonces se tiene que (4.39) es igual a

$$\left[\frac{1}{\sqrt{2}\sqrt{\pi}} \left(\frac{1}{2}\right)^{m/2} \frac{(ab)^{(m-2)/2}}{\sqrt{ab}} e^{-(a+b)/2} \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})}\right] \cdot \left[\frac{1}{(2\pi)^{m-2}(ab)^{(m-2)/2}}\right] =$$

$$= \frac{1}{\sqrt{2}\sqrt{\pi}} \frac{1}{(2\pi)^{m-2}} \frac{1}{2^m} \frac{1}{\sqrt{ab}} e^{-(a+b)/2} \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})}.$$

Hemos probado que

$$p_{(X(e_1),X(e_2),Y(e_1,e_2))}(a,b,\ldots\ldots,0) = \tag{4.40}$$
$$\frac{1}{\sqrt{2}\sqrt{\pi}} \frac{1}{(2\pi)^{m-2}} \frac{1}{2^m} \frac{1}{\sqrt{ab}} e^{-(a+b)/2} \frac{1}{\Gamma(\frac{m}{2})\Gamma(\frac{m-1}{2})} \tag{4.41}$$

Paso 4:

Falta acotar la esperanza condicional en (4.23), y para eso calculemos $\Delta(e_1, e_2)$.

Sea $T_{(s,t)}V$ el espacio tangente a V en el punto (s,t). Como la dimensión del espacio tangente a una variedad diferenciable es igual a la dimensión de la variedad, se tiene que $\dim(T_{(s,t)}V) = 2m-3$. Es claro que $T_{(s,t)}V$ es paralelo al complemento ortogonal en $\mathbb{R}^m \times \mathbb{R}^m$ de los vectores $(s,0),(0,t),(t,s)$.
Sea

$$\mathscr{B}_2 = \{(e_3,0),\ldots,(e_m,0),(0,e_3),\ldots,(0,e_m),\frac{1}{\sqrt{2}}(e_2,-e_1)\}$$

base ortonormal de $T_{(e_1,e_2)}V$ y calculemos $Y'(e_1,e_2)$ en esa base. Como $Y : V \to \mathbb{R}^{2m}$ se tiene que $Y'(e_1,e_2) : T_{(e_1,e_2)}V \to \mathbb{R}^{2m}$, por lo tanto podemos ver el diferencial como un operador de $\mathbb{R}^{2m-3}$ en $\mathbb{R}^{2m}$.
Escribimos

$$Y(s,t) = \begin{pmatrix} y_1(s,t) \\ \vdots \\ y_m(s,t) \\ z_1(s,t) \\ \vdots \\ z_m(s,t) \end{pmatrix}$$

donde

$$\begin{aligned} y_i(s,t) &= \sum_{j=1}^{m} b_{ij}s_j - \left(\sum_{j,k=1}^{m} b_{jk}s_js_k\right) s_i \\ z_l(s,t) &= \sum_{j=1}^{m} b_{lj}t_j - \left(\sum_{j,k=1}^{m} b_{jk}t_jt_k\right) t_l \end{aligned}$$

recordando que $\pi_u(Bu) = Bu - \langle Bu, u\rangle\, u$.
Entonces si calculamos las derivadas parciales, obtenemos

$$\frac{\partial y_i}{\partial s_k}(s,t) = \begin{cases} b_{ik} - 2\left(\sum_{k=1}^{m} b_{hk}s_k\right) s_i & i \neq k \\ b_{ik} - 2\left(\sum_{k=1}^{m} b_{hk}s_k\right) s_i - \left(\sum_{j,k=1}^{m} b_{jk}s_js_k\right) & i = k, \end{cases}$$

y si evaluamos en (e_1, e_2) se tiene

$$\frac{\partial y_i}{\partial s_k}(e_1,e_2) = \begin{cases} b_{ik} & i \neq k,\ i > 1 \\ b_{ii} - b_{11} & i = k \end{cases}$$

$$\frac{\partial y_1}{\partial s_k}(e_1,e_2) = \begin{cases} b_{1k} - 2b_{1k} = -b_{1k} & k \neq 1 \\ -2b_{11} & k = 1. \end{cases}$$

Análogamente para las derivadas parciales de las coordenadas en z obtenemos

$$\frac{\partial z_i}{\partial t_k}(e_1,e_2)=\begin{cases} b_{ik} & i\neq k,\ i\neq 2\\ b_{ii}-b_{22} & i=k,\ i\neq 2\end{cases}$$

$$\frac{\partial z_2}{\partial t_k}(e_1,e_2)=\begin{cases} -b_{1k} & k\neq 1\\ -2b_{22} & k=1.\end{cases}$$

Para encontrar $Y'(e_1,e_2)$ en la base $\mathscr{B}_2$, basta tomar como matriz la que tiene como columnas a las derivadas en las direcciones de los respectivos vectores de la base de $\mathscr{B}_2$. Es decir

$$\begin{aligned}Y'(e_1,e_2)=\Big(&d_{(e_1,e_2)}Y\cdot(e_3,0),\ldots\ldots,d_{(e_1,e_2)}Y\cdot(e_m,0),\\ &d_{(e_1,e_2)}Y\cdot(0,e_3),\ldots\ldots,d_{(e_1,e_2)}Y\cdot(0,e_m),d_{(e_1,e_2)}Y\cdot(e_2,-e_1)\frac{1}{\sqrt{2}}\Big),\end{aligned}$$

luego

$$Y'(e_!,e_2)=\left[\begin{array}{ccc|ccc|c} \frac{\partial y_1}{\partial s_3} & \cdots & \frac{\partial y_1}{\partial s_m} & 0 & \cdots & 0 & *\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ \frac{\partial y_m}{\partial s_3} & \cdots & \frac{\partial y_m}{\partial s_m} & 0 & \cdots & 0 & *\\ \hline 0 & \cdots & 0 & \frac{\partial z_1}{\partial t_3} & \cdots & \frac{\partial z_1}{\partial t_m} & *\\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & 0 & \frac{\partial z_m}{\partial t_3} & \cdots & \frac{\partial z_m}{\partial t_m} & * \end{array}\right],$$

siendo la última columna igual a $u=d_(e_1,e_2)Y\cdot(e_2,-e_1)\frac{1}{\sqrt{2}}$

$$=\left[\begin{array}{cccc|cccc|c} -b_{13} & -b_{14} & \cdots & -b_{1m} & 0 & 0 & \cdots & 0 & *\\ b_{23} & b_{24} & \cdots & b_{2m} & 0 & 0 & \cdots & 0 & *\\ b'_{33} & b_{34} & \cdots & b_{3m} & 0 & 0 & \cdots & 0 & *\\ b_{43} & b'_{44} & \cdots & b_{4m} & 0 & 0 & \cdots & 0 & *\\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ b_{m3} & b_{m4} & \cdots & b'_{mm} & 0 & 0 & \cdots & 0 & *\\ \hline 0 & 0 & \cdots & 0 & b_{13} & b_{14} & \cdots & b_{1_m} & *\\ 0 & 0 & \cdots & 0 & -b_{23} & -b_{24} & \cdots & -b_{2m} & *\\ 0 & 0 & \cdots & 0 & b''_{33} & b_{34} & \cdots & -b_{3m} & *\\ 0 & 0 & \cdots & 0 & b_{43} & b''_{44} & \cdots & b_{3m} & *\\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \cdots & \vdots & \vdots\\ 0 & 0 & \cdots & 0 & b_{m3} & b_{m4} & \cdots & b''_{mm} & * \end{array}\right]$$

con $b'_{jj}=b_{jj}-b_{11}$ y $b''_{ll}=b_{ll}-b_{22}$.

Estudiemos la última columna, que esta formada por las coordenadas del vector u:

$$u = \frac{1}{\sqrt{2}} \left[\begin{array}{ccc|ccc} \frac{\partial y_1}{\partial s_1} & \dots & \frac{\partial y_1}{\partial s_m} & 0 & \dots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial s_1} & \dots & \frac{\partial y_m}{\partial s_m} & 0 & \dots & 0 \\ \hline 0 & \dots & 0 & \frac{\partial z_1}{\partial t_1} & \dots & \frac{\partial z_1}{\partial t_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial z_m}{\partial t_1} & \dots & \frac{\partial z_m}{\partial t_m} \end{array} \right] \cdot \begin{pmatrix} e_2 \\ e_1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{\partial y_1}{\partial s_2} \\ \vdots \\ \frac{\partial y_m}{\partial s_2} \\ -\frac{\partial z_1}{\partial t_1} \\ \vdots \\ -\frac{\partial z_m}{\partial t_1} \end{pmatrix}$$

y sustituyendo las derivadas parciales, se obtiene

$$u = \frac{1}{\sqrt{2}} \left(b_{12}, b_{22} - b_{11}, \underbrace{b_{32}, \dots, b_{m2}}_{w^t}, b_{22} - b_{11}, -b_{12}, \underbrace{-b_{31}, \dots, -b_{m1}}_{-v^t} \right)^t .$$

Definamos

$$w^t = (b_{32}, \dots, b_{m2})$$

y

$$v^t = (b_{31}, \dots, b_{m1}),$$

entonces se tiene

$$Y'(e_1, e_2) = \left[\begin{array}{c|c|c} -v^t & 0_{1\times(m-2)} & -\frac{1}{\sqrt{2}} b_{12} \\ \hline w^t & 0_{1\times(m-2)} & \frac{1}{\sqrt{2}}(b_{22} - b_{11}) \\ \hline B_{12} - b_{11} I_{m-2} & 0_{(m-2)\times(m-2)} & \frac{1}{\sqrt{2}} w \\ \hline 0_{1\times(m-2)} & v^t & \frac{1}{\sqrt{2}}(b_{22} - b_{11}) \\ \hline 0_{1\times(m-2)} & -w^t & \frac{1}{\sqrt{2}} b_{12} \\ \hline 0_{(m-2)\times(m-2)} & B_{12} - b_{22} I_{m-2} & -\frac{1}{\sqrt{2}} v \end{array} \right] .$$

Luego si condicionamos $b_{11} = a$, $b_{22} = b$, $v = w = 0$ se tiene que

$$Y'(e_1, e_2) = \left[\begin{array}{c|c|c} 0_{1\times(m-2)} & 0_{1\times(m-2)} & 0 \\ \hline 0_{1\times(m-2)} & 0_{1\times(m-2)} & \frac{1}{\sqrt{2}}(b - a) \\ \hline B_{12} - a I_{m-2} & 0_{(m-2)\times(m-2)} & 0_{(m-2)\times 1} \\ \hline 0_{1\times(m-2)} & 0_{1\times(m-2)} & \frac{1}{\sqrt{2}}(b - a) \\ \hline 0_{1\times(m-2)} & 0_{1\times(m-2)} & 0 \\ \hline 0_{(m-2)\times(m-2)} & B_{12} - b I_{m-2} & 0_{(m-2)\times 1} \end{array} \right] .$$

Para completar el cálculo de $\Delta(e_1, e_2)$ es necesario calcular $(Y'(e_1, e_2))^t \cdot Y'(e_1, e_2)$ que es igual:

$$(Y'(e_1, e_2))^t \cdot Y'(e_1, e_2) = \left[\begin{array}{c|c|c} C^t C & 0_{m-2} & 0_{(m-2)\times 1} \\ \hline 0_{m-2} & D^t D & 0_{(m-2)\times 1} \\ \hline 0_{1\times(m-2)} & 0_{1\times(m-2)} & 2\left(\frac{1}{\sqrt{2}}(b - a)\right)^2 \end{array} \right] ,$$

donde $C = B_{12} - aI_{m-2}$ y $D = B_{12} - bI_{m-2}$ entonces por la definición de $\Delta(e_1, e_2)$ se tiene

$$\Delta(e_1,e_2)^2 = \left|\det\left[\begin{array}{c|c|c} C^tC & 0_{m-2} & 0_{(m-2)\times 1} \\ \hline 0_{m-2} & D^tD & 0_{(m-2)\times 1} \\ \hline 0_{1\times(m-2)} & 0_{1\times(m-2)} & (b-a)^2 \end{array}\right]\right|,$$

y desarrollando el determinante obtenemos que

$$\begin{aligned} \Delta(e_1,e_2)^2 &= \left|\det(C)^2\det(D)^2(b-a)^2\right| \\ &= \left|\det(B_{12}-aI_{m-2})^2\det(B_{12}-bI_{m-2})^2(b-a)^2\right| \end{aligned}$$

por lo tanto

$$\Delta(e_1,e_2) = \left|\det(B_{12}-aI_{m-2})\right| \cdot \left|\det(B_{12}-bI_{m-2})\right| \cdot (a-b). \tag{4.42}$$

Notar que si una matriz simétrica M de $n \times n$ es definida positiva ($M \succ 0$) entonces $\langle Mx, x\rangle > 0$ para todo $x = (x_1, \ldots, x_n) \in \mathbb{R}^n - \{0\}$. Luego si consideramos $\hat{x} = (0, x_2, \ldots, x_n) \neq (0, \ldots, 0)$ se tiene que $\langle M\hat{x}, \hat{x}\rangle > 0$ y por lo tanto $\langle M_1 y, y\rangle > 0$ para todo $y = (y_1, \ldots, y_{n-1}) \in \mathbb{R}^{n-1} \setminus \{0\}$. De esta manera se concluye lo siguiente: si $M \succ 0$ entonces $M_1 \succ 0$. Análogo para el caso en que M sea definida negativa.
Por lo tanto si $B_1 - aI_{m-1} \prec 0$ implica que $B_{12} - aI_{m-2} \prec 0$ y análogamente si $B_2 - bI_{m-1} \succ 0$ implica que $B_{12} - bI_{m-2} \succ 0$.
Si $M \succ 0$ entonces $\lambda_1, \ldots, \lambda_n$ son todos mayores que 0 y $M - aI \prec 0$ entonces es claro que $a > \lambda_i$ para $i = 1 \ldots n$. Esto se ve claramente ya que

$$0 > \langle (M-aI)v_i, v_i\rangle = \langle \lambda_i v_i - av_i, v_i\rangle = \langle \lambda_i v_i, v_i\rangle - \langle av_i, v_i\rangle = \lambda_i - a$$

siendo v_i vector propio de norma 1 asociado al valor propio λ_i para $i = 1 \ldots n$.
Entonces:

$$\begin{aligned} |\det(M-aI)| &= \left|\prod_{i=1}^{k}(a-\lambda_i)\right| \\ &= \prod_{i=1}^{k}(a-\lambda_i) \\ &\leq \prod_{i=1}^{k} a = a^k. \end{aligned}$$

Por tanto hemos concluido que:
Si $M \succ 0$ y $M - aI \prec 0 \quad \Longrightarrow \quad |\det(M-aI)| \leq a^k$.

Entonces retomando la esperanza condicional en (4.31)

$$E\left(\Delta(e_1,e_2)\, Ind_{\{B_1 - b_{11}I_{m-1} \prec 0,\, B_2 - b_{22}I_{m-1} \succ 0\}} \,\middle|\, b_{11} = a,\ b_{22} = b,\ b_{12} = 0,\ v = w = 0\right)$$

y observando (4.42) podemos usar la anterior conclusión para obtener que la esperanza condicional esta acotada por

$$a^{m-2}(a-b)E\big(|\det(B_{12}-bI_{m-2})|\, Ind_{\{B_2-bI_{m-1}\succ 0\}} \mid b_{11}=a,\ b_{22}=b,\\ b_{12}=0,\ v=w=0\big),$$

acotando $a-b$ por a y que $B_2-bI_{m-1}\succ 0$ implica que $B_{12}-bI_{m-2}\succ 0$ se tiene que la esperanza condicional en (4.31) es acotada por

$$a^{m-1}E\left(|\det(B_{12}-bI_{m-2})|\, Ind_{\{B_{12}-bI_{m-2}\succ 0\}} \mid v=w=0\right)$$

donde la nueva condición se debe a que las variables b_{11}, b_{22} y b_{12} ya no juegan ningún papel en la esperanza condicional.
Ahora si condicionamos por a_1 y a_2 tenemos que la esperanza condicional tiene la siguiente forma

$$E\Big(f(a_3,\dots,a_m) \mid a_1,\ a_2,\ \langle a_i,a_1\rangle=0,\ \langle a_i,a_2\rangle=0,\ i=3\dots m\Big)$$

donde los condicionantes a_1 y a_2 son vectores de $\mathbb{R}^m$ no colineales (con probabilidad 1). Llamando V al subespacio generado por a_1 y a_2, o sea

$$V=\Big\langle a_1,a_2\Big\rangle$$

y tomando como base ortonormal de $\mathbb{R}^m$ a $\{v_1,\dots,v_m\}$ siendo $\{v_1,v_2\}\subset V$ se tiene que para $i=3\dots m$

$$a_i=\sum_{j=1}^{m}\lambda_{ij}\,v_j=\sum_{j=3}^{m}\lambda_{ij}\,v_j$$

por ser perpendiculares a a_1 y a_2 ($a_i\in V^\perp$ para $i=3\dots m$). Por la invariancia por isometrías de la distribución normal en $\mathbb{R}^m$, tenemos que $\lambda_{ij}\sim\mathcal{N}(0,1)$ y son independientes en i. Luego la esperanza condicional es igual que la esperanza incondicional de $f(\tilde a_1,\dots,\tilde a_m)$ siendo $\tilde a_i$ variables aleatorias con distribución estándar en $\mathbb{R}^{m-2}$.
De esta manera, se tiene

$$E\Big(|\det(B_{12}-bI_{m-2})|\, Ind_{\{B_2-bI_{m-1}\succ 0\}} \mid a_1,\ a_2,\ \langle a_i,a_k\rangle=0,\ k=1,2,\ i=3\dots m\Big)$$

es igual a

$$E\Big(|\det(M)-bI_{m-2})|\, Ind_{\{M-bI_{m-2}\succ 0\}}\Big)$$

siendo $M=((M_{ij}))$ con $M_{ij}=\langle v_i,v_j\rangle$ y los v_i son normales estándar en $\mathbb{R}^{m-2}$ e independientes para $i,j=3,\dots,m$, entonces la esperanza condicional nos queda acotada por

$$\begin{aligned} a^{m-1}E\left(|\det(M)-bI_{m-2}) \mid Ind_{\{M-bI_{m-2}\succ 0\}}\right) &\le a^{m-1}E\left(\det M\right)\\ &= a^{m-1}(m-2)! \end{aligned}$$

donde la igualdad es consecuencia del *Lema 11* que demostraremos a continuación. Luego tenemos acotada la esperanza condicional en (4.31) por

$$a^{m-1}(m-2)!.$$

Lema 11. *Sean $\xi_1, \dots, \xi_m$ vectores aleatorios i.i.d en $\mathbb{R}^d$, con $d \geq m$ y distribución común Gaussiana estándar con matriz de varianza I_d. Si llamamos $W_{m,d}$ a la matriz $m \times m$ dada por*

$$W_{m,d} = ((\langle \xi_i, \xi_j \rangle))_{i,j=1\dots m},$$

y sea $D(\lambda)$ su polinomio característico, es decir

$$D(\lambda) = \det(W_{m,d} - \lambda I_d).$$

Entonces se prueba

(i)

$$E(\det(W_{m,d})) = d(d-1)\cdots(d-m+1)$$

(ii)

$$E\big(D(\lambda)\big) = \sum_{k=0}^{m} (-1)^k \, C_k^m \, \frac{d!}{(d-m+k)!} \, \lambda^k$$

siendo C_k^m combinaciones de m en k, ($C_k^m = \frac{m!}{k!(m-k)!}$).

Demostración. Los vectores son linealmente independientes con probabilidad uno. Fijemos los valores de $\xi_1, \dots, \xi_{m-1}$, y sea $\{v_1, \dots, v_d\}$ una base ortonormal de $\mathbb{R}^d$ tal que $v_1, \dots, v_{m-1}$ pertenecen al subespacio V_{m-1} generado por $\xi_1, \dots, \xi_{m-1}$.
Es conocido que el volumen del paralelepípedo formado por los vectores $\xi_1, \dots, \xi_m$, es decir el volumen de P siendo

$$P = \{\sum_{k=1}^{m} a_k \xi_k \; : \; 0 \leq a_k \leq 1, \; k = 1 \dots m \; \}$$

es igual al determinante en $\mathbb{R}^m$ de los vectores ξ_i expresados en la base $\{v_1, \dots, v_m\}$. Sea $A = ((a_{ij}))_{i,j=1\dots m}$ tal que $\xi_i = \sum_{j=1}^m a_{ij} v_j$. Entonces

$$\langle \xi_i, \xi_j \rangle = \left\langle \sum_{k=1}^{m} a_{ik} v_k, \sum_{h=1}^{m} a_{jh} v_h \right\rangle = \sum_{k=1}^{m} a_{ik} a_{jk} = (AA^t)_{ij}$$

siendo $(AA^t)_{ij}$ el coeficiente (i,j) de la matriz AA^t. Por lo tanto usando el hecho conocido de que si $T : \mathbb{R}^m \to \mathbb{R}^m$ es una transformación lineal con matriz asociada A en $\{e_1, \dots, e_m\}$ base ortonormal de $\mathbb{R}^m$ entonces

$$\det(T(e_1), \dots, T(e_m)) = \det(A) \det(e_1, \dots, e_m) = \det(A),$$

se concluye que

$$Vol(P) = \det(\xi_1, \dots, \xi_m) = \det(A) = \sqrt{\det(AA^t)} = \sqrt{\det(W_{m,d})},$$

es decir, el cuadrado del volumen del paralelepípedo P es igual al determinante de la matriz $W_{m,d}$. Entonces es claro

$$\det(W_{m,d}) = d^2(\xi_m, V_{m-1}) \det(W_{m-1,d})$$

siendo $d(\xi_m, V_{m-1})$ la distancia Euclídea entre ξ_m y V_{m-1}. Ahora, la distribución condicional de la distancia $d^2(\xi_m, V_{m-1})$ dados $\xi_1, \dots, \xi_{m-1}$ por la invariancia por rotaciones de la distribución Gaussiana es igual a la distancia de ξ_m al subespacio $\{0\}^{m-1} \times \mathbb{R}^{d-m+1}$. Pero nuevamente por la invariancia por transformaciones ortogonales de la distribución Gaussiana, esta distancia es igual a la distancia al 0 en $\mathbb{R}^{d-m+1}$ de un vector Gaussiano en $\mathbb{R}^{d-m+1}$. Por lo tanto $d^2(\xi_m, V_{m-1})$ tiene distribución χ^2_{d-m+1}.
Entonces

$$\begin{aligned}
E\big(\det(W_{m,d})\big) &= E\big(d^2(\xi_m, V_{m-1}) \det(W_{m-1,d})\big) \\
&= E\Big(E\big(d^2(\xi_m, V_{m-1}) \det(W_{m-1,d}) \,\big|\, \xi_1, \dots, \xi_{m-1}\big)\Big) \\
&= E\Big(d^2(\xi_m, V_{m-1}) E\big(\det(W_{m-1,d}) \,\big|\, \xi_1, \dots, \xi_{m-1}\big)\Big) \\
&= E\big(d^2(\xi_m, V_{m-1})\big)\, E\big(\det(W_{m-1,d})\big) \\
&= (d-m+1)\, E\big(\det(W_{m-1,d})\big),
\end{aligned}$$

y si iteramos obtenemos (i)

$$E\big(\det(W_{m,d})\big) = (d-m+1)\,(d-m+2)\dots(d-1)d.$$

Probemos (ii).
Para eso escribamos $D(\lambda)$ como

$$D(\lambda) = \sum_{k=0}^{m} \frac{D^{(k)}(0)}{k!} \lambda^k. \tag{4.43}$$

Diferenciando k veces la función determinante se obtiene

$$D^{(k)}(\lambda) = (-1)^k \cdot \sum \det(W^{i_1,\dots,i_k}_{m,d} - \lambda I_{m-k})$$

donde las sumas se obtienen en todas las k-uplas distintas de $1, \dots, m$ de elementos distintos, y siendo $W^{i_1,\dots,i_k}_{m,d}$ la matriz $(m-k) \times (m-k)$ que resulta de eliminar las

filas y columnas indexadas en i_1 hasta i_k.
Entonces

$$D^{(k)}(0) = (-1)^k \cdot \sum \det(W_{m,d}^{i_1,\ldots,i_k}),$$

tomando esperanza se obtiene

$$E\big(D^{(k)}(0)\big) = (-1)^k \cdot \sum E\big(\det(W_{m,d}^{i_1,\ldots,i_k})\big)$$

y considerando lo hecho en la parte (i), y que la cantidad de sumandos es arreglos de m elementos en k lugares, se tiene

$$E\big(D^{(k)}(0)\big) = (-1)^k\, d(d-1)\ldots(d-(m-k)+1)\, \frac{m!}{(m-k)!}.$$

Entonces tomando esperanza en (4.43) se tiene

$$E\big(D(\lambda)\big) = \sum_{k=0}^{m} \frac{(-1)^k}{k!}\, C_k^m \frac{d!}{(d-m)!} \frac{m!}{(m-k)!}\, \lambda^k,$$

o sea

$$E\big(D(\lambda)\big) = \sum_{k=0}^{m} (-1)^k\, C_k^m \frac{d!}{(d-m+k)!}\, \lambda^k.$$

□

Entonces juntando esto con el cálculo hecho para la densidad en (4.40) se obtiene que

$$h(a,b) \leq C_m \frac{e^{-(a+b)/2}}{\sqrt{ab}} a^{m-1} \tag{4.44}$$

donde

$$C_m = \frac{1}{4(m-2)!}.$$

Paso 6:

Ahora probaremos la cota superior en (4.13). Para probar dicha cota es necesario hacer un trabajo previo.
El siguiente lema es extraído de [17].

Lema 12. *Sea $\overline{A}$ una matriz cuadrada arbitraria determinística en $\mathbb{R}^{n\times n}$, y A una matriz de variables aleatorias independientes centrada en $\overline{A}$, con distribuciones normales de varianza σ^2. Sea $v \in S^{m-1}$,
entonces*

$$P\left(\|A^{-1}v\| > x\right) < \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{\sigma x}.$$

Demostración. Llamemos $A = ((a_{ij}))_{i,j=1\ldots n}$, donde los a_{ij} tienen distribuciones $\mathscr{N}(\overline{a_{ij}}, \sigma^2)$ respectivamente, donde $\overline{A} = ((\overline{a_{ij}}))_{i,j=1\ldots n}$. Sea $U = ((u_{ij}))_{i,j=1\ldots n}$ una matriz ortogonal, y sea $B = UA$.
Los elementos de B son

$$b_{ij} = \sum_{k=1}^{n} u_{ik}\, a_{kj}.$$

Entonces la esperanza de b_{ij} es:

$$E(b_{ij}) = \sum_{k=1}^{n} u_{ik}\, \overline{a_{kj}},$$

y por lo tanto la nueva matriz B esta centrada en $U\overline{A}$.
Calculemos la matriz de covarianza:

$$\begin{aligned}
cov(b_{ij}, b_{i'j'}) &= E\Big(\big(\sum_{k=1}^{n} u_{ik}\,(a_{kj} - \overline{a_{kj}})\big) \cdot \big(\sum_{k'=1}^{n} u_{i'k'}\,(a_{k'j'} - \overline{a_{k'j'}})\big)\Big) \\
&= \sum_{k,k'=1}^{n} u_{ik}u_{i'k'}\; \underbrace{E\big\{(a_{kj} - \overline{a_{kj}})(a_{k'j'} - \overline{a_{k'j'}})\big\}}_{\sigma^2\,\delta_{kk'}\,\delta_{jj'}} \\
&= \sum_{k=1}^{n} u_{ik}\, u_{i'k}\, \sigma^2\, \delta_{jj'} \\
&= \sigma^2\, \delta_{jj'} \sum_{k=1}^{n} u_{ik}u_{i'k} \\
&= \sigma^2\, \delta_{jj'}\, \delta_{ii'},
\end{aligned}$$

entonces obtuvimos

$$cov(b_{ij}, b_{i'j'}) = \sigma^2\, \delta_{jj'}\, \delta_{ii'}.$$

Es claro que que los coeficientes de la matriz UA son independientes y tienen distribución normal con esperanza $\sum_{k=1}^{n} u_{ik}\overline{a_{kj}}$ y varianza σ^2.

Tomemos U ortogonal tal que $U^{-1}e_1 = v$. Entonces

$$\|A^{-1}v\| = \|A^{-1}U^{-1}e_1\| = \|(UA)^{-1}e_1\|.$$

Sea $B = UA$ (donde U es ortogonal por serlo U^{-1}). Entonces basta probar que

$$P\left(\|B^{-1}e_1\| > x\right) < \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{\sigma x}$$

donde B es una matriz de variables aleatorias independientes centrada en $U\overline{A}$, con distribuciones normales de varianza σ^2.

Es evidente que $\|B^{-1}e_1\|$ es la norma de la primera columna de B^{-1}, es decir $\|B^{-1}e_1\|$ es la longitud del vector $(B^{-1})_{:1}$, siendo $(B^{-1})_{:j}$ el vector que tiene como coeficientes la j-ésima columna de la matriz B^{-1}, y $(B^{-1})_{j:}$ el vector que tiene como coeficientes la j-ésima fila.
Como $BB^{-1} = Id$ es claro que las filas $B_{2:}, \dots, B_{n:}$ de B son ortogonales a $(B^{-1})_{:1}$, y el producto interno entre $B_{1:}$ con $(B^{-1})_{:1}$ es 1.
Entonces se tiene que la norma de la proyección de $B_{1:}$ sobre $(B^{-1})_{:1}$ es el inverso de la norma de $(B^{-1})_{:1}$. Veamos esto:
-Definiendo por w al vector de norma 1 en la dirección de $(B^{-1})_{:1}$, y V al subespacio generado por las filas $B_{2:}, \dots, B_{n:}$, o sea

$$w = \frac{(B^{-1})_{:1}}{\|(B^{-1})_{:1}\|} \quad \text{y} \quad V = < B_{2:}, \dots, B_{n:} >$$

se observa que la proyección π_w sobre el subespacio generado por w de $B_{1:}$ es

$$\pi_w(B_{1:}) = \langle B_{1:}, w \rangle \, w,$$

entonces la norma de la proyección es

$$\begin{aligned} \|\pi_w(B_{1:})\| &= |\langle B_{1:}, w \rangle| \\ &= \left| \left\langle B_{1:}, \frac{(B^{-1})_{:1}}{\|(B^{-1})_{:1}\|} \right\rangle \right| \\ &= |\langle B_{1:}, (B^{-1})_{:1} \rangle| \cdot \frac{1}{\|(B^{-1})_{:1}\|} \\ &= \frac{1}{\|(B^{-1})_{:1}\|}. \end{aligned}$$

Como las filas de B son independientes, la distribución condicional de la primer fila dadas las demás es la distribución incondicional. Entonces fijando una base ortonormal de $\mathbb{R}^n$ formada por $\{w, v_2, \dots, v_m\}$ se tiene

$$B_{1:} = \lambda_1 w + \sum_{i=2}^{n} \lambda_i v_i$$

donde $\pi_w(B_{1:}) = \lambda_1 w$, entonces $\|\pi_w(B_{1:})\| = |\lambda_1|$ que es la distribución del módulo de una variable aleatoria normal $\mathcal{N}(\overline{\mu}, \sigma^2)$ para algún $\overline{\mu}$.

Entonces

$$\begin{aligned} P\left(\|A^{-1}v\| > x\right) &= P\left(\frac{1}{\|A^{-1}v\|} < \frac{1}{x}\right) \\ &= P\left(|\lambda_1| < \frac{1}{x}\right) \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-1/x}^{+1/x} e^{-(u-\overline{\mu})^2/2\sigma^2}\, du \\ &\leq \sqrt{\frac{2}{\pi}} \frac{1}{x\sigma}, \end{aligned}$$

lo que completa la prueba. □

Corolario 6.1. *Asumimos que $A = (a_{ij})$ con $a_{ij} = m_{ij} + g_{ij}$ para $i, j = 1 \ldots m$, donde g_{ij} son variables aleatorias i.i.d. normales estándar. Sea $v \in S^{m-1}$. Entonces*

$$P\left(\|A^{-1}v\| > x\right) < \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x}.$$

Lema 13. *Sea $U = (U_1, \ldots, U_m)$ un vector aleatorio con distribución uniforme en S^{m-1} y sea t_{m-1} una variable aleatoria con distribución Student de $m-1$ grados de libertad.*
Entonces, si $c \in (0, m)$ se tiene

$$P\left(U_1^2 > \frac{c}{m}\right) = P\left(t_{m-1}^2 > \frac{m-1}{m-c} c\right).$$

Demostración. Como ya vimos en la demostración del Lema 10, podemos asumir que

$$U = \frac{V_1}{\|V\|}$$

con $V = (V_1, \ldots, V_m)$ con distribución Gaussiana en $\mathbb{R}^m$. Entonces se tiene

$$\left\{U_1^2 > \frac{c}{m}\right\} = \left\{\frac{V_1^2}{\|V\|^2} > \frac{c}{m}\right\} = \left\{\frac{V_1^2}{V_2^2 + \cdots + V_m^2} > \frac{c}{m-c}\right\}$$

entonces

$$\begin{aligned} P\left(U_1^2 > \frac{c}{m}\right) &= P\left(\frac{(m-1)V_1^2}{V_2^2 + \cdots + V_m^2} > \frac{m-1}{m-c} c\right) \\ &= P\left(t_{m-1}^2 > \frac{m-1}{m-c} c\right) \end{aligned}$$

siendo t_{m-1} una variable aleatoria con distribución Student de $m-1$ grados de libertad. □

Proposición 6.1. *Asumimos que* $A = (a_{ij})$ *con* $a_{ij} = m_{ij} + g_{ij}$ *para* $i, j = 1 \dots m$, *donde* g_{ij} *son variables aleatorias i.i.d. normales estándar y* $M = (m_{ij})_{i,j=1\dots m}$ *una matriz determinística.*
Entonces

$$P(\|A^{-1}\| \geq x) \leq C_2(m)\frac{\sqrt{m}}{x}$$

con $C_2(m)$ *una constante que depende de* m *pero esta acotada por* 2,34737 *para todo* m.

Demostración. Sea U con distribución uniforme en S^{m-1} (como en el Lema 13) independiente de A. Condicionando se tiene

$$P\big(\|A^{-1}U\| > x\big) = E\Big[P\big(\|A^{-1}U\| > x \mid U\big)\Big],$$

luego si aplicamos el Corolario 6.1 tenemos

$$P\big(\|A^{-1}U\| > x\big) \leq \left(\frac{2}{\pi}\right)^{1/2}\frac{1}{x}. \tag{4.45}$$

Sea v_A donde A^{-1} alcanza su norma como operador, es decir, sea v_A de norma uno tal que $\|A^{-1}v_A\| = \|A^{-1}\|$ y sea $u \in \mathbb{R}^m$ tal que $\|u\| = 1$. Llamando C a A^{-1}, se tiene que $\|Cv_A\| = \|C\|$, entonces v_A es el vector propio de C^tC con mayor valor propio igual a $\|C\|^2$ (ver Lema 8). Es decir $\|C^tCv_A\| = \|C\|^2$, y como $\langle C^tCv_A, u\rangle = \|C\|^2 \langle v_A, u\rangle$ se tiene

$$|\|C\|^2 \langle v_A, u\rangle| = |\langle Cv_A, Cu\rangle| \leq \|Cv_A\|\,\|Cu\| = \|C\|\,\|Cu\|,$$

y cambiando C por A^{-1} concluimos que

$$\|A^{-1}u\| \geq \|A^{-1}\| \cdot |\langle v_A, u\rangle|. \tag{4.46}$$

Tomando $c \in (0, m)$ se tiene usando (4.46) que

$$\begin{aligned}
P\left(\|A^{-1}U\| > x\left(\frac{c}{m}\right)^{1/2}\right) &\geq P\left(\big\{\|A^{-1}\| > x\big\} \cap \left\{|\langle v_A, U\rangle| \geq \left(\frac{c}{m}\right)^{1/2}\right\}\right)\\
&= E\left[P\left(\big\{\|A^{-1}\| > x\big\} \cap \left\{|\langle v_A, U\rangle| \geq \left(\frac{c}{m}\right)^{1/2}\right\}\Bigg| A\right)\right]\\
&= E\left[Ind_{\{\|A^{-1}\|>x\}}P\left(|\langle v_A, U\rangle| \geq \left(\frac{c}{m}\right)^{1/2}\Bigg| A\right)\right]\\
&\overset{Lema13}{=} E\left[Ind_{\{\|A^{-1}\|>x\}}P\left(t^2_{m-1} > \frac{m-1}{m-c}c\right)\right]\\
&= P\left(t^2_{m-1} > \frac{m-1}{m-c}c\right)E\left[Ind_{\{\|A^{-1}\|>x\}}\right]\\
&= P\left(t^2_{m-1} > \frac{m-1}{m-c}c\right)P\big(\|A^{-1}\| \geq x\big).
\end{aligned}$$

Luego

$$P\big(\|A^{-1}\| \geq x\big) \leq \frac{1}{P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right)} P\left(\|A^{-1}U\| > x\left(\frac{c}{m}\right)^{1/2}\right),$$

ahora usando (4.45) se tiene

$$P\big(\|A^{-1}\| \geq x\big) \leq \frac{1}{P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right)} \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{x}\left(\frac{m}{c}\right)^{1/2}.$$

La constante a saber es

$$\frac{1}{P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right)} \left(\frac{2}{\pi}\right)^{1/2} \frac{1}{c^{1/2}},$$

por tanto solo resta encontrar $c \in (0, m)$ de manera tal que la cota sea mínima. Para eso necesitamos tomar el ínfimo en $c \in (0, m)$. Tomando

$$C_2(m) = \sqrt{\frac{2}{\pi}} \left(\sup_{c\in(0,m)} \sqrt{c} P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right)\right)^{-1}$$

obtenemos una cota mínima.
Como

$$\begin{aligned} \sup_{c\in(0,m)} c^{1/2} P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right) &\geq \sup_{c\in(0,1)} c^{1/2} P\left(t_{m-1}^2 > \frac{m-1}{m-c}c\right) \\ &\geq \sup_{c\in(0,1)} c^{1/2} P\left(t_{m-1}^2 > c\right) \end{aligned}$$

se tiene que es posible acotar la $C_2(m)$ a través de cálculos numéricos, y de esta manera llegamos al resultado de la proposición. □

Una vez terminado el trabajo preliminar podemos continuar con el *Paso* 6:

Para $x > 1$ se tiene por la *Observación* 5 en la página 43 que

$$P\big(\kappa(A) > x\big) = P\Big(\frac{\lambda_m}{\lambda_1} > x^2\Big)$$

entonces partiendo el espacio en donde λ_1 sea mayor o menor a $\frac{L^2 m}{x^2}$ se tiene

$$P\big(\kappa(A) > x\big) \leq P\left(\lambda_1 < \frac{L^2 m}{x^2}\right) + P\left(\frac{\lambda_m}{\lambda_1} > x^2,\, \lambda_1 \geq \frac{L^2 m}{x^2}\right) \tag{4.47}$$

siendo L una constante positiva a determinar después.
Recordando que $\lambda_1 = \|A^{-1}\|^{-2}$, el primer término de (4.47) se acota usando la *Proposición 6.1*

$$P\left(\lambda_1 < \frac{L^2 m}{x^2}\right) = P\left(\|A^{-1}\| > \frac{x}{L\sqrt{m}}\right) \leq C_2(m)\frac{Lm}{x}. \tag{4.48}$$

Para el segundo termino de (4.47) se tiene

$$P\left(\frac{\lambda_m}{\lambda_1} > x^2,\, \lambda_1 \geq \frac{L^2 m}{x^2}\right) = \int_{L^2 m x^{-2}}^{+\infty} db \int_{bx^2}^{+\infty} h(a,b)\, da \leq G_m(x^2) \tag{4.49}$$

siendo

$$G_m(y) = C_m \int_{L^2 m y}^{+\infty} db \int_{by}^{+\infty} C_m \frac{e^{-(a+b)/2}}{\sqrt{ab}} a^{m-1}\, da$$

(donde la desigualdad en (4.49) es consecuencia de la acotación en (4.44)).
Tratemos de derivar $G_m(y)$:
-Sea

$$G(y) = \int_{\frac{A}{y}}^{+\infty} db \int_{by}^{+\infty} g(a,b)\, da$$

con A y b constantes y g derivable.
Notando por $F(b,y) = \int_{by}^{+\infty} g(a,b)\, da$ y $f(y) = A/y$ se tiene

$$G(y) = \int_{f(y)}^{+\infty} F(b,y)\, db.$$

Sea $h > 0$ entonces usando el hecho que $f(y)$ es decreciente en y se tiene

$$\begin{aligned}
\frac{G(y+h) - G(y)}{h} &= \frac{1}{h}\left[\int_{f(y+h)}^{+\infty} F(b,y+h)\, db - \int_{f(y)}^{+\infty} F(b,y)\, db\right] \\
&= \frac{1}{h}\left[\int_{f(y+h)}^{+\infty} \big(F(b,y+h) - F(b,y)\big)\, db + \int_{f(y+h)}^{f(y)} F(b,y)\, db\right]
\end{aligned}$$

y tomando límite en $h \downarrow 0$ se tiene

$$\begin{aligned}
\lim_{h\downarrow 0} \frac{G(y+h) - G(y)}{h} &= \lim_{h\downarrow 0} \int_{f(y+h)}^{+\infty} \left(\frac{F(b,y+h) - F(b,y)}{h}\right) db + \\
&\qquad + \lim_{h\downarrow 0} \frac{1}{h} \int_{f(y+h)}^{f(y)} F(b,y)\, db \\
&= \int_{f(y)}^{+\infty} \frac{\partial F}{\partial y}(b,y)\, db + F(f(y),y) \cdot \lim_{h\downarrow 0} \frac{f(y) - f(y+h)}{h} \\
&= \int_{f(y)}^{+\infty} \frac{\partial F}{\partial y}(b,y)\, db - F(f(y),y) \cdot f'(y).
\end{aligned}$$

Análogamente para $h < 0$.
Entonces obtuvimos

$$G'(y) = \int_{\frac{A}{y}}^{+\infty} \frac{\partial F}{\partial y}(b,y)\, db - F(A/y, y) \cdot f'(y).$$

Es claro que

$$\frac{\partial F}{\partial y}(b,y) = -b \cdot g(by, b)$$

entonces sustituyendo se obtiene

$$G'(y) = -\int_{\frac{L^2 m}{y}}^{+\infty} \frac{e^{-b(y+1)/2}(by)^{m-1}}{\sqrt{y}}\, db + \frac{L^2 m}{y^2} \int_{L^2 m}^{+\infty} \frac{e^{-(a+(L^2 m)/y)/2} a^{m-1}}{\sqrt{a}\sqrt{m} L \sqrt{y}^{-1}}\, da.$$

Como $G'_m(y) = C_m \cdot G'(y)$ se tiene que

$$\begin{aligned}
-G'_m(y) &\leq C_m \int_{\frac{L^2 m}{y}}^{+\infty} e^{-b(y+1)/2}\, b^{m-1}\, y^{m-3/2}\, db \\
&= C_m \cdot y^{m-3/2} \int_{\frac{L^2 m}{y}}^{+\infty} e^{-b(y+1)/2}\, b^{m-1}\, db \\
&= \frac{1}{4(m-2)!} y^{-3/2} y^m \int_{\frac{L^2 m}{y}}^{+\infty} e^{-b(y+1)/2}\, b^{m-1}\, db,
\end{aligned}$$

haciendo el cambio de variable $z = b(y+1)/2$ se tiene

$$\begin{aligned}
-G'_m(y) &\leq \frac{y^{-3/2}}{4(m-2)!} \cdot y^m \cdot \frac{2}{y+1} \int_{\frac{L^2 m(y+1)}{2y}}^{+\infty} e^{-z}\, z^{m-1}\, 2^{m-1} \frac{1}{(y+1)^{m-1}}\, dz \\
&= \frac{y^{-3/2}}{4(m-2)!} \cdot y^m \cdot \frac{2^m}{(y+1)^m} \int_{\frac{L^2 m(y+1)}{2y}}^{+\infty} e^{-z}\, z^{m-1}\, dz \\
&= \frac{y^{-3/2}}{4(m-2)!} \cdot 2^m \cdot \left(\frac{y}{y+1}\right)^m \int_{\frac{L^2 m(y+1)}{2y}}^{+\infty} e^{-z}\, z^{m-1}\, dz
\end{aligned}$$

y como $y/(y+1) < 1$ se tiene

$$-G'_m(y) < \frac{y^{-3/2}}{4(m-2)!} 2^m \int_{\frac{L^2 m(y+1)}{2y}}^{+\infty} e^{-z}\, z^{m-1}\, dz.$$

Sea

$$I_m(a) = \int_a^{+\infty} e^{-z}\, z^{m-1}\, dz,$$

integrando por partes se tiene

$$I_m(a) = -e^{-z}z^{m-1}\Big|_a^{+\infty} + (m-1)\int_a^{+\infty} e^{-z}\, z^{m-2}\, dz = e^{-a}a^{m-1} + (m-1)I_{m-1}(a),$$

es decir

$$I_m(a) = e^{-a}a^{m-1} + (m-1)I_{m-1}(a).$$

Entonces lo que se obtuvo es una ecuación de recurrencia y si desarrollamos se tiene que

$$I_m(a) = e^{-a}\big[a^{m-1} + (m-1)a^{m-2} + (m-1)(m-2)a^{m-3} + \cdots + (m-1)!\big].$$

Tomando $a > \frac{5}{2}m$ se tiene $m < \frac{2}{5}a$, entonces

$$\begin{aligned}
I_m(a) &< e^{-a}\big[a^{m-1} + ma^{m-2} + m^2a^{m-3} + \cdots + m^m\big] \\
&\leq e^{-a}\big[a^{m-1} + \left(\frac{2}{5}\right) a^{m-1} + \left(\frac{2}{5}\right)^2 a^{m-1} + \cdots + \left(\frac{2}{5}\right)^m a^{m-1}\big] \\
&< e^{-a}\, a^{m-1}\left(\sum_{i=0}^{+\infty}\left(\frac{2}{5}\right)^i\right) \\
&= e^{-a}\, a^{m-1}\left(\frac{1}{1-\frac{2}{5}}\right) \\
&= e^{-a}\, a^{m-1}\cdot\frac{5}{3},
\end{aligned}$$

entonces

$$I_m(a) < \frac{5}{3}\, e^{-a}\, a^{m-1}.$$

Como $a = L^2m/2 > (5/2)m$, podemos tomar $L^2 > 5$ y se obtiene la siguiente cota

$$G'_m(y) \leq D_m\, y^{-3/2} \;\text{ siendo }\; D_m = \frac{2^m}{4(m-2)!}\, e^{-L^2m/2}\left(\frac{L^2m}{2}\right)^{m-1}\frac{5}{3},$$

es decir

$$D_m = \frac{5}{6}\,\frac{m^{m-1}}{(m-2)!}\, L^{2(m-1)}\, e^{-L^2m/2}.$$

Ahora si aplicamos la fórmula de Stirling

$$k! = \left(\frac{k}{e}\right)^k \sqrt{2\pi k}\, e^{\frac{A_k}{k}} \;\text{ con }\; 0 < A_k < 12$$

se tiene que

$$D_m \leq \frac{5}{6}\,\frac{m^{m-1}}{\left(\frac{m-2}{e}\right)^{m-2}}\,\frac{L^{2m}\, L^{-2}\, e^{-L^2m/2}}{\sqrt{2\pi(m-2)}e^0} = \frac{5}{6}\,\frac{m\, m^{m-2}}{(m-2)^{m-2}}\,\frac{e^{m-2}\, L^{2m}}{L^2}\,\frac{e^{-L^2m/2}}{\sqrt{2\pi(m-2)}}$$

y considerando que $1+x<e^x$ se tiene

$$\left(\frac{m}{m-2}\right)^{m-2} = \left(\frac{m-2+2}{m-2}\right)^{m-2} = \left(1+\frac{2}{m-2}\right)^{m-2} < \left(e^{2/(m-2)}\right)^{m-2} = e^2.$$

Resulta que

$$\begin{aligned} D_m &\leq \frac{5}{6}\,e^2\,m\,\frac{e^{m-2}\,L^{2m}}{L^2}\,\frac{e^{-L^2m/2}}{\sqrt{2\pi}\sqrt{(m-2)}} = \frac{5}{6}\,\frac{1}{\sqrt{2\pi}}\,\frac{m}{\sqrt{(m-2)}}\,\frac{1}{L^2}\,e^{2m\log(L)}\,e^{-L^2m/2+m} \\ &= \frac{5}{12}\,\frac{\sqrt{2}}{\sqrt{2\pi}}\,\frac{m}{\sqrt{(m-2)}}\,\frac{1}{L^2}\,e^{-m/2(L^2-4\log(L)-2)} \end{aligned}$$

y si buscamos las raíces de $L^2-4\log(L)-2$ obtendremos una cota. Existe una raíz en aproximadamente $L\approx 2{,}3145$ que es mayor a $\sqrt{5}$ entonces se tiene que

$$D_m \leq \frac{5\sqrt{2}}{12\sqrt{\pi}L^2}m.$$

De esta manera se tiene

$$0 \leq G_m(y) = \int_y^{+\infty} -G'_m(t)\,dt < D_m\int_y^{+\infty}\frac{dt}{t^{3/2}} = 2D_m\,y^{-1/2}$$

y entonces cambiando y por x^2, en (4.49) se tiene

$$P(\frac{\lambda_m}{\lambda_1} > x^2,\ \lambda_1 \geq \frac{L^2m}{x^2}) \leq 2D_m\,\frac{1}{x}. \tag{4.50}$$

Juntando (4.48) y (4.50) obtenemos la cota superior

$$P\big(\kappa(A) > x\big) \leq C_2(m)\frac{Lm}{x} + 2D_m\,\frac{1}{x} < (C_2(+\infty)\,L + 2\frac{D_m}{m})\frac{m}{x} < \frac{Cm}{x}$$

con

$$C = 5{,}60 \approx C_2(+\infty)\,L + \frac{2\,5\sqrt{2}}{12\sqrt{\pi}L^2}.$$

$\square$

Capítulo 5

Miscelánea

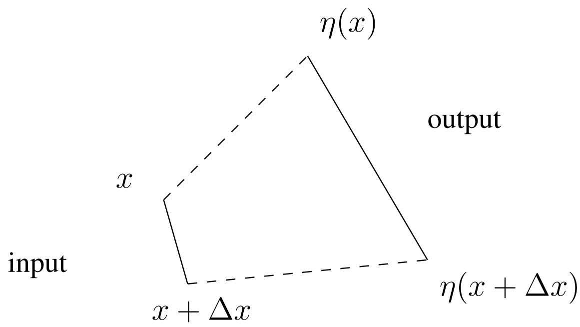
5.1. Número de condición en general

La idea de esta sección es mostrar que el concepto de número de condición se puede extrapolar a problemas más generales que los tratados en esta monografía. Las referencias para este capítulo son: [8] y [10].

El número de condición es un tema clásico en análisis numérico. Fue introducido con el propósito de obtener cotas tanto para los redondeos como para la complejidad del problema.
Es decir el número de condición fue introducido originalmente para medir la sensibilidad de un input a perturbaciones, más precisamente mide la sensibilidad del "peor" caso para pequeñas perturbaciones. Por ejemplo, si nuestro objetivo es computar una función η donde el output varia continuamente como función del input, el número de condición para el input $x \in \mathbb{R}^n$ es medir el tamaño de $\|\eta(x+\Delta x) - \eta(x)\|$ donde Δx es la perturbación.
Una manera de hacerlo es considerar

$$cond(x) = \lim_{\delta \downarrow 0} \sup_{\Delta x \leq \delta} \frac{\|\eta(x+\Delta x) - \eta(x)\| / \|\eta(x)\|}{\|\Delta x\| / \|x\|}$$

donde $\| \cdot \|$ es alguna norma en $\mathbb{R}^n$.



Esta definición deja de ser útil cuando el output toma un conjunto de valores discretos. Ejemplos clásicos de este tipo de problemas son los problemas de decisión, es decir, aquellos en que dado un input devuelven "si" o "no", o "1" o "0" (por ejemplo preguntarle a la maquina si una matriz es invertible o no). Si consideramos la definición dada arriba obtenemos que

$$cond(x) = \begin{cases} 0 & \text{si } x \notin \Sigma \\ \infty & \text{si } x \in \Sigma \end{cases}$$

siendo Σ el conjunto de input que no responde ni "si" ni "no" (o sea la frontera entre el output "si" o "no"). Por ejemplo en el caso de decir cuando una matriz es invertible o no, tenemos que Σ son las matrices no invertibles.
Una manera de atacar este tipo de problema para obtener un número de condición útil, es tomar como definición el *Teorema del Número de Condición* 3 que estudiamos para el caso de número de condición de matrices cuadradas.
Sea $\{y_j : i \in J\}$ el conjunto de outputs posibles, y Σ la frontera de los conjuntos $X_i = \{x : \eta(x) = y_i\}$.
Cualquier dato en Σ lo consideramos como ill-posed. Entonces definimos el número de condición como:

$$cond(x) = \frac{\|x\|}{d(x, \Sigma)} \tag{5.1}$$

donde $\|\cdot\|$ es alguna norma en el espacio discreto de inputs.

En la *Introducción* a esta monografía se comento como entra en juego la probabilidad en el estudio del número de condición, comentando sobre la dificultad del cálculo mismo del número de condición. Otro enfoque a considerar es estimar cuan complejo es calcular el número de condición, o sea estudiar el número de condición del número de condición.
Existe una conjetura de *Renegar* que dice que computar $cond_\Pi(x)$ para el problema Π, es igual de complejo que resolver el problema Π.
A la "condición de la condición" se le llama *número de condición grado*-2 y se nota $cond_\Pi^{[2]}(x)$ que fue introducida por *J. Demmel*. Es decir

$$cond_\Pi^{[2]}(x) = \lim_{\delta \downarrow 0} \sup_{\Delta x \leq \delta} \frac{\|cond(x+\Delta x) - cond(x)\| / \|cond(x)\|}{\|\Delta x\| / \|x\|}$$

donde usamos la definición para el caso en que el output varia continuamente en función del input, tal como lo hace el número de condición (aunque el problema sea discreto).
Demmel probó para algunos tipos de problemas que el número de condición grado-2 coincide con el número de condición salvo multiplicar por alguna constante.

F. Cucker y D.Cheung probaron en [8] el siguiente teorema:

Teorema 7. *Sea* Π *un problema en el que su número de condición se definió en términos de la distancia relativa a* Σ *como en (5.1), entonces*

$$cond_\Pi(x) - 1 \le cond_\Pi^{[2]}(x) \le cond_\Pi(x) + 1.$$

Por su simplicidad, incluiremos la prueba:

Demostración. Sea como antes Σ las ill-posed, x un input y notemos $\rho(x) = d(x, \Sigma)$.

$$\begin{aligned}
cond_\Pi^{[2]}(x) &= \lim_{\delta \downarrow 0} \sup_{\Delta x \le \delta} \frac{\|cond(x+\Delta x) - cond(x)\| / \|cond(x)\|}{\|\Delta x\| / \|x\|} \\
&= \lim_{\delta \downarrow 0} \sup_{\Delta x \le \delta} \frac{\left| \frac{\|x+\Delta x\|}{\rho(x+\Delta x)} - \frac{\|x\|}{\rho(x)} \right| \|x\|}{\frac{\|x\|}{\rho(x)} \|\Delta x\|} \\
&= \lim_{\delta \downarrow 0} \sup_{\Delta x \le \delta} \left| \frac{\|x+\Delta x\|\rho(x) - \|x\|\rho(x+\Delta x)}{\rho(x+\Delta x)\|\Delta x\|} \right|.
\end{aligned}$$

Es fácil ver que

$$\begin{aligned}
\big| \|x+\Delta x\|\rho(x) - \|x\|\rho(x+\Delta x) \big| &= \big| \big(\|x+\Delta x\| - \|x\| \big) \rho(x) + \\
&\qquad + \|x\| \big(\rho(x) - \rho(x+\Delta x) \big) \big| \\
&\le \|\Delta x\|\rho(x) + \|x\| \, \|\Delta x\|
\end{aligned}$$

usando que $|\rho(x) - \rho(x+\Delta x)| \le \|\Delta x\|$. Entonces se tiene

$$\begin{aligned}
\left| \frac{\|x+\Delta x\|\rho(x) - \|x\|\rho(x+\Delta x)}{\rho(x+\Delta x)\|\Delta x\|} \right| &\le \frac{\|\Delta x\|\rho(x) + \|x\| \, \|\Delta x\|}{\rho(x+\Delta x)\|\Delta x\|} \\
&\le \frac{\|\Delta x\|\rho(x) + \|x\| \, \|\Delta x\|}{\rho(x) - \|\Delta x\|}.
\end{aligned}$$

Por lo tanto

$$\begin{aligned}
cond_\Pi^{[2]}(x) &= \lim_{\delta \downarrow 0} \sup_{\Delta x \le \delta} \left| \frac{\|x+\Delta x\|\rho(x) - \|x\|\rho(x+\Delta x)}{\rho(x+\Delta x)\|\Delta x\|} \right| \\
&\le \lim_{\delta \downarrow 0} \sup_{\Delta x \le \delta} \left| \frac{\rho(x) + \|x\|}{\rho(x) - \|\Delta x\|} \right| \\
&= \frac{\rho(x) + \|x\|}{\rho(x)} \\
&= 1 + \frac{\|x\|}{\rho(x)} \\
&= 1 + cond_\Pi(x)
\end{aligned}$$

y esto prueba la cota superior.
Para probar la cota inferior consideremos Δx_0 tal que $\|\Delta x_0\| = \rho(x)$ y $x + \Delta x_0 \in \Sigma$.
Para cualquier $0 < \delta < \|\Delta x_0\|$ sea

$$\|\Delta x_0^\delta\| = \frac{\delta}{\rho(x)} \Delta x_0.$$

Entonces $\|\Delta x_0^\delta\| = \delta$ y $\rho(x + \Delta x_0^\delta) = \rho(x) - \|\Delta x_0^\delta\| = \rho(x) - \delta$.
De igual manera que antes se tiene

$$\begin{aligned}
\left| \frac{\|x + \Delta x_0^\delta\|\rho(x) - \|x\|\rho(x + \Delta x_0^\delta)}{\rho(x + \Delta x_0^\delta)\|\Delta x_0^\delta\|} \right| &= \left| \frac{\|x + \Delta x_0^\delta\|\rho(x) - \|x\|(\rho(x) - \delta)}{(\rho(x) - \delta)\delta} \right| \\
&\geq \left| \frac{(\|x\| - \|\Delta x_0^\delta\|)\rho(x) - \|x\|\rho(x) - \delta}{(\rho(x) - \delta)\delta} \right| \\
&= \left| \frac{(\|x\| - \delta)\rho(x) - \|x\|\rho(x) - \delta}{(\rho(x) - \delta)\delta} \right| \\
&= \frac{-\delta\rho(x) + \|x\|\delta}{(\rho(x) - \delta)\delta} \\
&= \frac{\|x\| - \rho(x)}{\rho(x) - \delta},
\end{aligned}$$

entonces

$$\begin{aligned}
cond_\Pi^{[2]}(x) &= \lim_{\delta \downarrow 0} \sup_{\Delta x \leq \delta} \left| \frac{\|x + \Delta x\|\rho(x) - \|x\|\rho(x + \Delta x)}{\rho(x + \Delta x)\|\Delta x\|} \right| \\
&\geq \lim_{\delta \downarrow 0} \frac{\|x\| - \rho(x)}{\rho(x) - \delta} \\
&= \frac{\|x\| - \rho(x)}{\rho(x)} \\
&= cond_\Pi(x) - 1,
\end{aligned}$$

lo que prueba la cota inferior. $\square$

Capítulo 6

Apéndice

6.1. Serie de Neumann

Nuestro objetivo es mostrar que si un operador lineal T en un espacio de Banach verifica que $\|T\| < 1$ entonces el operador $I - T$ es invertible y la norma de la inversa esta acotada por $\frac{1}{1-\|T\|}$.

Sea $(X, \|\cdot\|)$ un espacio normado. Diremos que una serie $\sum x_n$ en normalmente convergente cuando converge $\sum \|x_n\| < \infty$. Si el espacio es completo es fácil ver que toda serie normalmente convergente converge.
En efecto si $S_n = x_1 + \cdots + x_n$, la condición de que $\sum \|x_n\| < \infty$ nos dice que la cola de la serie tiende a cero. Por tanto dado $\varepsilon > 0$ existe n_0 tal que $\|x_{n_0+1}\| + \cdots\cdots < \varepsilon$. Pero entonces si $m > n_0$ se tiene

$$\begin{aligned} \|S_m - S_{n_0}\| = \|x_{n_0+1} + \cdots + x_m\| &\leq \|x_{n_0+1}\| + \cdots + \|x_m\| \\ &< \sum_{n=n_0+1}^{+\infty} \|x_n\| \\ &< \varepsilon \end{aligned}$$

entonces $\{S_n\}$ es de Cauchy y por tanto converge.
Sea ahora $(X, \|\cdot\|)$ un espacio de Banach (espacio normado completo), y $T \in B(X)$ siendo $B(X)$ el espacio de los operadores lineales de X en X acotados. Supongamos que $\|T\| < 1$ entonces es fácil ver que $\sum_{n=0}^{+\infty} T^n$ converge a un elemento de $B(X)$. En efecto usando la propiedad submultiplicativa de la norma de operadores $\|T^n\| \leq \|T\|^n$ se tiene

$$\sum_{n=0}^{+\infty} \|T^n\| \leq \sum_{n=0}^{+\infty} \|T\|^n = \frac{1}{1 - \|T\|} < \infty$$

entonces $\sum_{n=0}^{+\infty} T^n$ es normalmente convergente y por tanto converge, además su norma esta acotada por $\frac{1}{1-\|T\|}$. Veamos que $\sum_{n=0}^{+\infty} T^n$ es la inversa de $I - T$.

$$\begin{aligned}
\left(\sum_{n=0}^{+\infty} T^n\right) \cdot (I-T) &= \lim_N \left(\sum_{n=0}^{N} T^n\right) \cdot (I-T) \\
&= \lim_N \left(\sum_{n=0}^{N} T^n - T^{n+1}\right) \\
&= \lim_N I - T^{N+1} \\
&= I
\end{aligned}$$

pues $\lim_n \|T^n\| \leq \lim_n \|T\|^n = 0$ y entonces $\lim_n T^n = 0$.
Análogamente se prueba que $(I - T) \cdot \sum_{n=0}^{+\infty} T^n = I$.
Hemos probado que $\|T\| < 1$ implica que $I - T$ es invertible siendo

$$\sum_{n=0}^{+\infty} T^n = (I-T)^{-1}$$

donde a la anterior suma se denomina *Serie de Neumann*.

6.2. Teorema del Valor Medio para campos vectoriales

Sea $f : U \to \mathbb{R}$ continuamente diferenciable ($f \in C^1$), con $U \in \mathbb{R}^d$, $x, y \in U$ y suponemos que el segmento entre x e y este incluido en U. Sea $F : [0,1] \to \mathbb{R}$ dada por $F(t) = f(x + t(x - y))$, entonces $F \in C^1$ y vale el Teorema de Valor Medio de Lagrange para funciones de $\mathbb{R}$ en $\mathbb{R}$.

$$F(1) - F(0) = F'(\hat{t}) \quad \text{con} \quad \hat{t} \in (0,1).$$

Pero $F'(\hat{t}) = \nabla f \cdot (x - y)$ entonces si $\eta = x + \hat{t}(x - y)$ tenemos

$$f(y) - f(x) = f_{x_1}(\eta)(x_1 - y_1) + \cdots + f_{x_d}(\eta)(x_d - y_d)$$

siendo $f_{x_j} = \frac{\partial f}{\partial x_j}$. Tomando valor absoluto en ambos miembros se tiene

$$|f(y) - f(x)| \leq \sum_{j=1}^{d} |f_{x_j}(\eta)|(x_j - y_j)$$

y como f es continuamente diferenciable, $\exists\ M$ tal que $|f_{x_j}(\chi)| < M$ con χ en el segmento con extremos x e y que es un compacto. Entonces

$$\begin{aligned}|f(y)-f(x)| &\leq M\sum_{j=1}^{d}|x_j-y_j| \\ &\leq M\sqrt{d}\|x-y\|\end{aligned}$$

siendo $\|\cdot\|$ la norma Euclideana. En conclusión, obtuvimos la siguiente desigualdad:

$$|f(y)-f(x)| \leq M\sqrt{d}\|x-y\|.$$

Ahora con esta última desigualdad atacamos el caso para campos vectoriales.

Sea $F: U \to \mathbb{R}^d$ con $F \in C^1$, con $U \subset \mathbb{R}^d$, $F(x) = (F_1(x), \ldots, F_d(x))$.

$$F(y)-F(x) = \begin{pmatrix} F_1(y)-F_1(x) \\ \vdots \\ F_d(y)-F_d(x) \end{pmatrix}$$

entonces

$$\begin{aligned}\|F(y)-F(x)\| &= \left(\sum_{j=1}^{d}|F_j(y)-F_j(x)|^2\right)^{\frac{1}{2}} \\ &\leq \left(\sum_{j=1}^{d}M^2 d\|x-y\|^2\right)^{\frac{1}{2}} \\ &= \left(d^2M^2\|x-y\|^2\right)^{\frac{1}{2}} \\ &= dM\|x-y\|,\end{aligned}$$

es decir

$$\|F(y)-F(x)\| \leq dM\|x-y\|. \tag{6.1}$$

Referencias

[1] J-M. Azais; M. Wschebor, *Upper and Lower Bounds for the Tails of the Distribution of the Condition Number of a Gaussian Matrix*, SIAM Journal of Matrix Analysis and Applications, **26** (2005), 426–440.

[2] ———, *On the Distribution of the Maximum of a Gaussian Field with d Parameters*, The Annals of Applied Probability, **15** (2005), 254–278.

[3] ———, *Computing the Distribution of the Maximum of a Gaussian Process*, Prepublicaciones de Matemática de la UDELAR, 01/ 44, http://premat.fing.edu.uy, (2001).

[4] J.A. Cuesta-Albertos; M. Wschebor, *Some Remarks on the Condition Number of a Real Random Square Matrix*, Journal of Complexity, **19** (2003), 548–554.

[5] ———, *Condition Numbers and Extrema of Random Fields*, IV Ascona Seminar on Stochastic Analysis, Random Fields and Applications, Progress in Probability, Birkhaüser, (2004), 69–82.

[6] F.Cucker; M. Wschebor, *On the Expected Condition Numbers of Linear Programming Problems*, Numerische Mathematik, **94** (2003), 419–478.

[7] M. Wschebor, *One-parameter Gaussian Process: Lectures on the Distribution of the Maximum*, Notas de curso de postgrado, Mérida, Asociación Venezolana de Matemáticas, 2001.

[8] D. Cheung; F. Cucker, *A note on level-2 condition numbers*, Technical Report, City University of Hong Kong, (2004).

[9] L. Blum; F. Cucker; M. Shub; S. Smale, *Complexity and real computation*, Springer, 1998.

[10] F. Cucker; R. Wong, *The collected papers of Stephen Smale*, vol. 3, Singapore University Press; River Edge, NJ : World Scientific, Singapore, 2000.

[11] S. Smale, *Complexity Theory and Numerical Analysis*, Acta Numerica, Cambridge University Press, (1997), 523–551.

[12] K. Atkinson, *An introduction to numerical Analysis*, Wiley, University of Iowa, 1987.

[13] J.H. Wilkinson, *Rounding Errors in Algebraic Process*, Prentice-Hall, Series in automatic computation, 1963.

[14] ———, *The Algebraic Eigenvalue Problem*, Oxford Science Publications, 1992.

[15] Von Neumann; H. Goldstine, *Numerical inverting of matrices of high order*, Bull. Amer. Math. Soc. (1947), 1021–1099.

[16] A.M. Turing, *Rounding-off errors in matrix processes*, Quart. J. Mech. Appl. Math. (1948), 287–308.

[17] D. Spielman A. Sankar S. Teng, *Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices*, Preprint, http://basilo.kaist.ac.kr/papers/MIT/Spielman/ (2003).