

TESIS DE MAESTRIA

Sistemas Polinomiales Aleatorios

Diego Armentano

Orientador: Mario Wschebor

Maestria en Matemática (PEDECIBA)
Facultad de Ciencias
Universidad de la República
Montevideo, Uruguay

30 de Agosto del 2007

Índice general

1. Introducción	6
1.1. Breve Reseña Histórica	7
1.2. Notación	8
1.3. Objetivo	8
I Sistemas Polinomiales	10
2. Teorema de Bézout	11
2.1. Espacio Proyectivo	11
2.2. Polinomios Homogéneos	14
2.3. Estructura Hermitiana de \mathcal{H}_d e Invarianza Unitaria	16
2.4. Teorema de Bézout	21
II Sistemas Polinomiales Aleatorios	28
3. Teorema de Shub-Smale	29
3.1. Introducción	29
3.2. Teorema de Shub-Smale	31
3.2.1. Complementos Técnicos	38
4. Nuevo Enfoque	42
4.1. Introducción	42
4.2. Formalización del Nuevo Enfoque y Extensiones	44
4.2.1. Ejemplo de nuevo producto interno $\mathcal{O}(n)$ -invariante	47
4.2.2. Unicidad del producto interno de Weyl en el caso complejo	48
4.3. Herramienta principal	49
4.4. Nuevo Teorema de Shub-Smale y Extensiones	51
5. Smooth Analysis	60
5.1. Introducción	60
5.2. Teorema de Perturbación Aleatoria	62
5.3. Ejemplos	71
5.3.1. Shub-Smale	71
5.3.2. $Q^{(i)} = Q$, con raíces solamente reales	72

5.3.3. $Q^{(i)} = Q^{l_i}$ con $i = 1, \dots, m.$	73
5.4. Complementos Técnicos	75
6. Apéndice	83
6.1. Teorema Principal de Teoría de la Eliminación	83
6.2. Procesos Estocásticos	85
6.3. El conjunto Σ tiene medida nula en $\mathbb{P}(\mathcal{H}_{(d)})$	86
6.4. Fórmula de Rice, Fórmula del Área y Complementos	88
6.5. Fórmula de Co-Área	92

Resumen

El tema principal de este trabajo es el estudio del número de soluciones de un sistema de m ecuaciones polinomiales en m incógnitas reales, cuando los coeficientes son tomados al azar. Mostraremos dos enfoques diferentes de como atacar el problema. El primero es utilizado por M. Shub y S. Smale, que introducen una medida particular en el espacio de coeficientes con cierta propiedad geométrica (invarianza bajo la acción del grupo ortogonal) y mediante la Fórmula de co-área logran computar el promedio del número de soluciones. El segundo enfoque se debe a J.M. Azaïs y M. Wschebor, en donde utilizan la fórmula de Rice que expresa los momentos del número de pre-imágenes de un campo aleatorio $f : M \rightarrow \mathbb{R}^d$, donde M es un subconjunto de \mathbb{R}^d , mediante una integral. Este nuevo enfoque permite generalizar el primero a otras medidas en el espacio de coeficientes. En particular mostraremos como se puede atacar con estos métodos un nuevo problema. Dicho problema consiste en estudiar que le sucede al número de soluciones de un sistema polinomial determinístico cuando a éste se lo perturba aleatoriamente. Más precisamente, consideraremos sistemas polinomiales aleatorios de la forma $P_i(t) + X_i(t) = 0$, $t \in \mathbb{R}^m$, $i = 1, \dots, m$, donde los P_i 's (*señal*) son polinomios no aleatorios y X_i 's (*ruido*) son polinomios aleatorios independientes con distribución Gaussiana centrada invariantes bajo transformaciones ortogonales del espacio. Para cada i fijo, los polinomios P_i y X_i tienen grado efectivo d_i . Probaremos que bajo ciertas hipótesis en lo referente a la relación entre la señal y el ruido, el cociente entre el número esperado de soluciones del sistema perturbado y el número esperado de soluciones del sistema centrado, tiende geoméricamente a cero cuando el tamaño del sistema crece. Esto significa que el comportamiento del valor esperado del número de soluciones es dominado por el ruido.

Además de establecer los pre-requisitos geométricos necesarios para estudiar el primer enfoque, incluiremos una demostración del teorema de Bézout. También estudiaremos la conexión entre los dos enfoques, el cual nos ayudará a entender más en profundidad el problema planteado.

Palabras claves: Sistemas polinomiales homogéneos, Teorema de Bézout, Sistemas polinomiales aleatorios, Fórmula de co-área, Fórmula de Rice.

Abstract

The main topic of this work, is the study of the number of solutions of a polynomial system with m equations and m unknowns when the coefficients are taken at random. We show two different approaches to this problem. The first, due to M. Shub and S. Smale, introduce a particular measure on the coefficients having some geometric properties (invariance under the action of the orthogonal group). By means of the co-area formula they are able to compute the average of solutions. The second approach, due to J.M. Azaïs and M. Wschebor, use the Rice Formula, which expresses as an integral the moments of the number of inverse images of a random field $f : M \rightarrow \mathbb{R}^d$, where M is a subset of \mathbb{R}^d . This second approach, permits to extend the results of the first one to different measures on the coefficients. It also allows to consider the following problem: studying the behaviour of the number of roots of a deterministic polynomial system when it is perturbed by a random system of polynomial equations. Precisely, we consider random systems of equations over the reals, with m equations and m unknowns $P_i(t) + X_i(t) = 0$, $t \in \mathbb{R}^m$, $i = 1, \dots, m$, where the P_i 's are non random (the "signal"), and the X_i 's (the "noise") are independent real-valued Gaussian centered random polynomial field, with a probability law invariant under the action of the orthogonal group of \mathbb{R}^m . For fixed i , the polynomials P_i 's and X_i 's have effective degree d_i . We prove that under certain conditions on the relation signal over noise, it follows that the quotient between the expected value of the number of roots of the perturbed system and the expected value corresponding to the centered case, tends to zero geometrically fast as m tends to infinity. This means that the behaviour of the number of roots (under these conditions) is governed by the noise part.

Beside establishing geometric requirements needed to study the first approach, we include a proof of Bézout's theorem. Also, we study the connection between these approaches and we show that this helps to get a deeper understanding of the problem.

Key words: System of homogeneous polynomials, Bézout's theorem, System of random polynomials, Co-area formula, Rice formula.

Capítulo 1

Introducción

Los sistemas de ecuaciones polinomiales son un tema de mucha importancia dentro de la ciencia en general. El estudio de los sistemas polinomiales ha sido un tema de gran interés dentro de la matemática tanto pura como aplicada.

Cuando queremos concentrarnos en el estudio de sus soluciones nos enfrentamos a un problema conocido desde hace tiempo. Genéricamente, es imposible encontrar un método de resolución explícita de un sistema polinomial. Como consecuencia de esto, los métodos aproximados toman un papel categórico. Dichos métodos se centran principalmente en el antiguo y conocido *Método de Newton*. Este método consiste en partir de un punto arbitrario en el espacio \mathbb{R}^m , luego considerar un proceso iterativo (partiendo del punto) que intente aproximarse a una raíz del sistema, si ese no es el caso, repetir el procedimiento partiendo de un nuevo punto. Con suerte, si el punto inicial pertenece a lo que se denomina *cuenca de una raíz* el método de Newton converge (ver Figura 1.1 para el caso uni-dimensional).

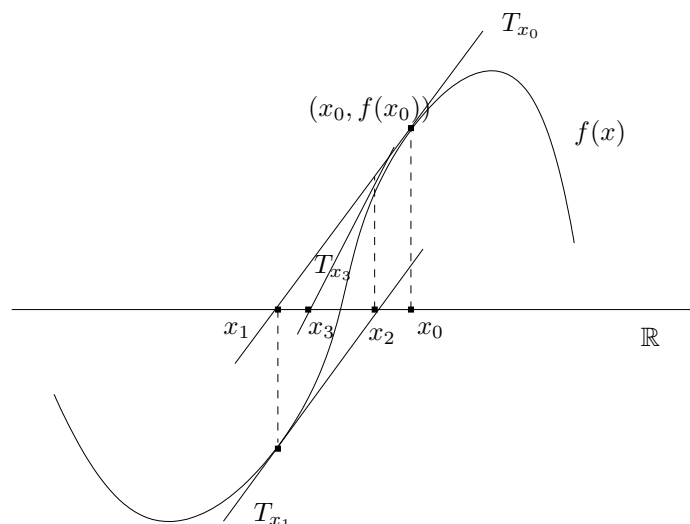


Figura 1.1: Método de Newton en \mathbb{R} , donde T_x indica la tangente por $(x, f(x))$.

Siendo este el método existente para buscar soluciones al sistema, es de vital importancia (a efectos de acelerar el método iterativo) saber donde es más “conveniente” tomar el punto de

partida, problema especialmente difícil cuando m es grande. En los últimos años se ha concentrado mucho esfuerzo en tratar de entender este problema (ver [23]).

En este trabajo nos centraremos en el estudio del número de soluciones de un sistema polinomial de m ecuaciones y m incógnitas sobre los reales. En el caso complejo, el *Teorema de Bézout* da una solución a este problema: genéricamente hay \mathcal{D} soluciones sobre \mathbb{C}^m donde \mathcal{D} es el producto de los grados de los polinomios (también llamado *número de Bézout del sistema*). En el caso real, el problema es mucho más complicado.

Una primera aproximación para abordar este problema, es randomizar los coeficientes del sistema y de esa manera poder estudiar la distribución de las soluciones del sistema. En general, éste es un problema muy ambicioso. Los resultados más abordados hoy en día se relacionan con el estudio del valor esperado del número de soluciones, pero para clases muy particulares de distribuciones, como veremos a lo largo de este texto.

Vale la pena mencionar que hasta el momento se sabe muy poco sobre el tema.

1.1. Breve Reseña Histórica

El primer resultado sobre el tema que aparece en la literatura es un artículo de Bloch y Pólya [7] que estudian el caso de un polinomio en que sus coeficientes son variables aleatorias discretas independientes tomando valores dentro del conjunto $\{-1, 0, 1\}$. Más tarde Littlewood y Offord en [18] y [19] estimaron el promedio de raíces de un polinomio respecto a diferentes distribuciones en los coeficientes.

Luego Marc Kac [15] por primera vez obtuvo resultados asintóticos sobre el valor esperado del número de raíces de un polinomio con coeficientes aleatorios (ver el libro de Barucha-Reid y Sambandham [6]). Más precisamente, mostró que un polinomio de grado d , donde sus coeficientes son variables aleatorias independientes normales estándar tiene asintóticamente en el grado $\frac{2}{\pi} \ln d$ raíces en promedio.

En 1992, Shub y Smale [22] computaron la esperanza del número de soluciones reales de un sistema de ecuaciones polinomiales donde los coeficientes se rigen por cierta ley particular. En el trabajo se utilizó una medida de probabilidad que implicaba una propiedad de invarianza geométrica del problema, muy natural desde el punto de vista geométrico. Más precisamente, dicha ley era invariante por la acción del grupo ortogonal del espacio Euclídeo en el cual se trabaja.

Razones para utilizar leyes con este tipo de invarianzas, provienen de la tradición de la física teórica, en la cual la invarianza unitaria (u ortogonal) juega un papel central. Grosso modo, los físicos teóricos están interesados en estudiar los polinomios pero independientemente del sistema de coordenadas en el cual están dados.

Un trabajo similar pero para el caso en que todos los polinomios tengan el mismo grado fue realizado por Kostlan [17].

Más tarde en el 2005, Azaïs y Wschebor [3] dieron una nueva prueba del teorema de Shub-Smale. El método de prueba se basa en la *Fórmula de Rice* de la teoría de campos aleatorios. Este nuevo enfoque les permitió obtener fórmulas generales para la esperanza del número de raíces cuando hay independencia estocástica entre los polinomios del sistema y la ley es centrada e invariante por isometrías del espacio.

En Wschebor [25] por primera vez se estudia la varianza del número de raíces reales.

En el 2007, Armentano-Wschebor [1] estudian el caso no centrado, en otras palabras, estudian que sucede cuando se hacen perturbaciones aleatorias de un sistema determinístico. En dicho trabajo se prueba que el número de soluciones decrece geoméricamente respecto al caso del sistema con una ley centrada, cuando se hace crecer el tamaño del sistema. El resultado es válido cuando se verifica cierta relación entre el sistema al cual se perturba y el “tamaño” de la perturbación.

Para explicar en detalle algunos de estos temas es necesario comenzar con la notación que usaremos a lo largo de este texto.

1.2. Notación

Consideremos $f = (f_1, \dots, f_m)$,

$$f_i(t) := \sum_{\|j\| \leq d_i} a_j^{(i)} t^j \quad (i = 1, \dots, m) \quad (1.1)$$

un sistema de m polinomios en m variables reales. La notación que usamos es la siguiente:

- $t := (t_1, \dots, t_m)$ denota un punto en el espacio \mathbb{R}^m ,
- $j := (j_1, \dots, j_m)$ es un multi-índice de enteros no negativos,
- $\|j\| = \sum_{h=1}^m j_h$,
- $t^j := t^{j_1} \dots t^{j_m}$,
- $a_j^{(i)} = a_{j_1, \dots, j_m}^{(i)}$,
- $a_j^{(i)} \in \mathbb{R}$,

El grado efectivo del polinomio f_i es d_i ($a_j^{(i)} \neq 0$ para algún j de norma d_i).

Notaremos por $N^f(V)$ al número de soluciones del sistema de ecuaciones

$$f_i(t) = 0 \quad (i = 1, \dots, m),$$

que se encuentran en el subconjunto $V \subset \mathbb{R}^m$, es decir,

$$N^f(V) = \#\{t \in V : f_i(t) = 0 \forall i = 1, \dots, m\}.$$

1.3. Objetivo

La idea es randomizar los coeficientes $\{a_j^{(i)}\}$, y estudiar la variable aleatoria $N^f(V)$. Poco se sabe sobre el comportamiento de $N^f(V)$, aún en el caso de elecciones simples en la distribución de los coeficientes.

El objetivo de este trabajo es tratar de mostrar dos diferentes ataques al problema. El primero es el utilizado por Shub y Smale en [22]. En éste, Shub y Smale probaron que si los coeficientes son variables aleatorias Gaussianas centradas independientes teniendo varianzas:

$$E \left[(a_j^{(i)})^2 \right] = \frac{d_i!}{j_1! \dots j_m! (d_i - \|j\|)!}. \quad (1.2)$$

entonces

$$E(N^f) = (d_1 \cdots d_m)^{1/2}. \quad (1.3)$$

Trataremos de introducir las ideas principales usadas en dicho artículo, mostrando la utilidad de considerar la varianza (1.2).

Luego mostraremos como atacar el problema usando la *Fórmula de Rice*, enfoque realizado por Azaïs-Wschebor en [3]. Mostraremos como desde el punto de vista de los campos aleatorios es posible extender los resultados de Shub y Smale. Más aún, mostraremos como este enfoque permite dar algunas respuestas al caso no centrado (a lo que se le denomina regularmente como “smooth analysis”).

Aquí la idea es remover la hipótesis sobre el centramiento de los coeficientes. Una forma de ver este problema es el siguiente. Supongamos que partimos de un sistema polinomial determinístico

$$P_i(t) = 0 \quad (i = 1, \dots, m)$$

al cual llamaremos “*señal*” del sistema. Luego lo perturbamos aleatoriamente con un “*ruido*” polinomial $\{X_i(t) : i = 1, \dots, m\}$. De esta manera consideramos el nuevo sistema de ecuaciones polinomiales aleatorias

$$P_i(t) + X_i(t) = 0 \quad (i = 1, \dots, m)$$

y nos preguntamos que sucede con la cantidad de soluciones reales del nuevo sistema. Para atacar este problema primero es claro que hay que imponer algunas hipótesis sobre la señal y sobre el ruido, y también sobre la relación entre ambos.¹ Estos detalles que son de carácter puramente técnico quedaran claros más adelante.

En [1] se probó que si la relación de la señal sobre el ruido no es ni muy pequeña ni muy grande, entonces existen constantes positivas $C, \theta, 0 < \theta < 1$ tales que

$$E(N^{P+X}) \leq C \theta^m E(N^X). \quad (1.4)$$

En una primera lectura de la desigualdad (1.4) es interesante observar que si el sistema original tiene un gran número de raíces, posiblemente infinitas, y m es grande entonces el efecto de la perturbación es que el número de raíces se reduce geoméricamente comparado con el caso centrado (i.e. el caso en que los P_i son idénticamente nulos).

En la demostración de este resultado se observa que las constantes C, θ pueden ser estimadas por medio de las hipótesis.

Observar que en el caso que el ruido sea del tipo de Shub-Smale la fórmula (1.4) toma una agradable forma:

$$E(N^{P+X}) \leq C \theta^m \sqrt{d_1 \cdots d_m}. \quad (1.5)$$

También es el caso, cuando se pueda obtener formulas generales para $E(N^X)$ como en [3].

¹A primera vista, si el sistema original tiene infinitas raíces, es de esperar que $E(N^{P+X})$ tienda a infinito a medida que el ruido sea arbitrariamente pequeño.

Parte I

Sistemas Polinomiales

Capítulo 2

Teorema de Bézout

Este capítulo tiene dos finalidades. Una de ellas es demostrar el teorema de Bézout. La otra es dotar al espacio de sistemas polinomiales de una estructura de producto interno que sea invariante bajo transformaciones ortogonales (o transformaciones unitarias en el caso complejo) para luego definir una forma de volumen sobre dicho espacio. Para estos fines, va a ser necesario exponer rápidamente conceptos básicos sobre espacios proyectivos y homogenización de polinomios. La referencia básica para este capítulo es [5].

2.1. Espacio Proyectivo

La idea de esta sección es dar las definiciones que usaremos en las secciones siguientes cuando hablemos de espacios proyectivos. Sea W un \mathbb{C} -espacio vectorial de dimensión finita. Definimos el espacio proyectivo $\mathbb{P}(W)$ como el conjunto formado por todos los subespacios uni-dimensionales de W , es decir, el conjunto de todas las direcciones (ver Figura 2.1)

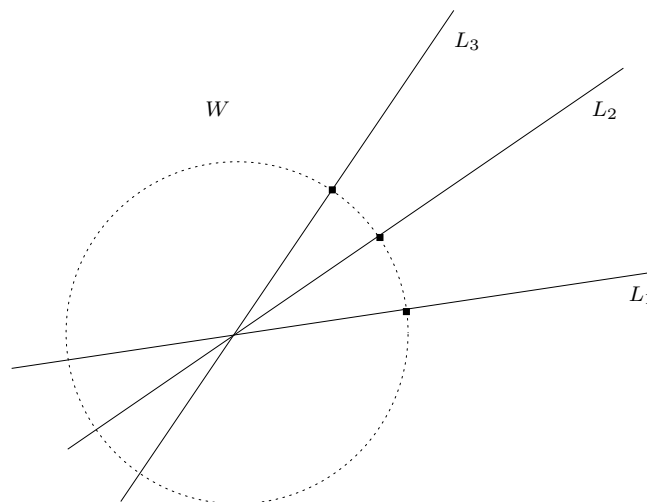


Figura 2.1: Identificación de los Subespacios L_1, L_2, L_3 .

La buena forma de definirlo es considerarlo como un espacio cociente:
-dados $v, w \in W^* := W - \{0\}$ decimos que v es equivalente a w ($v \sim w$) si y solo si existe $\lambda \in \mathbb{C}^*$

tal que $\lambda v = w$. Definimos $\mathbb{P}(W) = W^* / \sim$.

Notaremos por $Q : W^* \rightarrow \mathbb{P}(W)$ el mapa cociente, donde $Q^{-1}(x)$ es una recta que pasa por el origen (pero omitiéndolo) en la dirección de x .

Supongamos que W viene provisto de un producto interno Hermitiano $\langle \cdot, \cdot \rangle$. Entonces definimos para todo $v \in W^*$

$$T_v := \{w \in W : \langle v, w \rangle = 0\}$$

el sub-espacio ortogonal a $v \in W$, quien tiene codimensión uno en W (ver Figura 2.2).

También definimos

$$S_1(W) := \{w \in W : \|w\| = 1\}$$

como la esfera unitaria en W (que es homeomorfa a la esfera unitaria $S^{2 \dim W - 1} \subset \mathbb{C}^{\dim W}$). Observar que de considerar la restricción de $Q|_{S_1(W)} : S_1(W) \rightarrow \mathbb{P}(W)$ resulta la compacidad de $\mathbb{P}(W)$. Además observando que $(Q|_{S_1(W)})^{-1}(x)$ se puede identificar con

$$S^1 := \{\lambda \in \mathbb{C} : |\lambda| = 1\}$$

para cualquier $x \in \mathbb{P}(W)$, entonces podemos recuperar $\mathbb{P}(W)$ como el cociente de $S_1(W)$ por la acción de S^1 (con la acción $\lambda \cdot w := \lambda w$ para todo $\lambda \in S^1$ y $w \in S_1(W)$).

Es conocido que podemos dotar $\mathbb{P}(W)$ de una estructura de variedad diferenciable. Una manera de hacerlo es establecer un sistema de coordenadas en W y luego proseguir igual que la prueba de que $\mathbb{P}(\mathbb{C}^n)$ es una variedad diferenciable. Dicha prueba se puede encontrar en cualquier libro de Geometría (Diferencial o Algebraica) como por ejemplo [16], [8] o [21]. Aquí no tomaremos ese camino, sino que haremos otra prueba que resulta más geométrica.

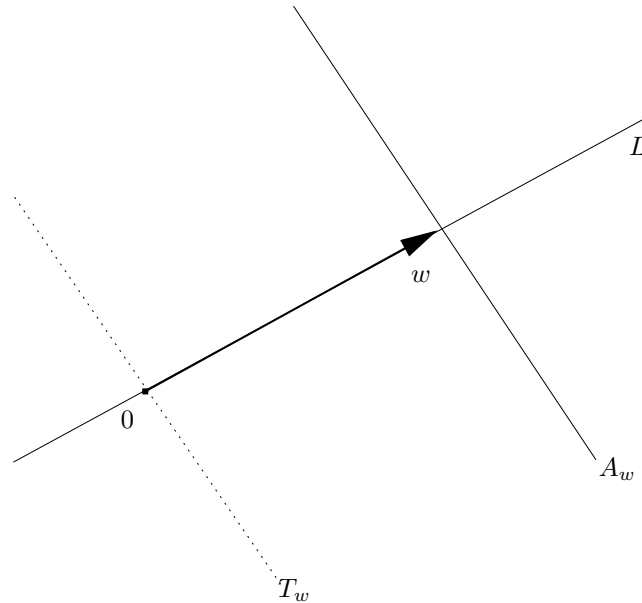


Figura 2.2: Subespacio afín A_w

Sea L una recta por el origen y tomemos $0 \neq w \in L$. Definimos el subespacio afín

$$A_w := w + T_w$$

(ver Figura 2.2)

Si tomamos L' otra recta por el origen que no sea ortogonal a L obtenemos que L' corta al subespacio afín A_w en un único punto. Eso resulta de tomar $0 \neq w' \in L'$ y observar que existe un único $\lambda_0 \in \mathbb{C}$ tal que $\lambda_0 w' \in A_w$ (basta tomar $\lambda_0 = \langle w, w' \rangle / \|w'\|^2$, ver Figura 2.3)

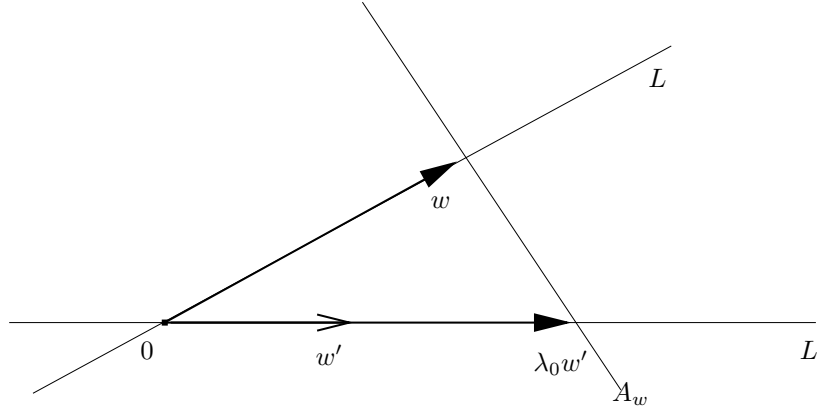


Figura 2.3: Intersección de L' con A_w

Entonces podemos definir un mapa biyectivo

$$\varphi_w : \mathbb{P}(W) \setminus Q(T_w) \rightarrow A_w.$$

(que de hecho es un homeomorfismo considerando en $\mathbb{P}(W)$ la topología cociente) que manda una dirección L' en el único punto de corte de L' con A_w . Llamando $\psi_w : A_w \rightarrow \mathbb{P}(W)$ el mapa inverso de φ_w obtenemos un sistema de cartas que definen una estructura diferenciable en $\mathbb{P}(W)$ (considerando $\{(A_w, \psi_w) : w \in W^*\}$ como el atlas diferenciable).

Definimos el fibrado tangente $T\mathbb{P}(W)$ como el cociente de $W^* \times W^*$ por la siguiente relación:

$$(x, w) \sim (x', w') \iff \exists \lambda \in \mathbb{C}^* \text{ tal que } x' = \lambda x, w' = \lambda w. \tag{2.1}$$

Una forma de convencerse de porqué usar el escalado en la segunda coordenada es observar la Figura 2.4:

Podemos definir una estructura producto interno en $T\mathbb{P}(W)$ de la siguiente manera. Dado $x \in W$ definimos

$$\langle v, w \rangle_x := \frac{\langle v, w \rangle}{\|x\|^2} \quad \forall v, w \in T_x W. \tag{2.2}$$

Veamos que la definición (2.2) baja al cociente $T\mathbb{P}(W)$ definido en (2.1). Para eso basta observar:

$$\langle \lambda v, \lambda w \rangle_{\lambda x} = \frac{\langle \lambda v, \lambda w \rangle}{\|\lambda w\|^2} = \frac{\langle v, w \rangle}{\|w\|^2} = \langle v, w \rangle_x.$$

Por lo tanto dotamos el fibrado $T\mathbb{P}(W)$ con una estructura Riemanniana.

Es posible bajar los mapas lineales de W en W como funciones en el proyectivo ya que los mapas lineales llevan rectas por el origen en rectas por el origen. De esta manera el espacio lineal de transformaciones que preservan la estructura Hermitiana de W (al cual notaremos por

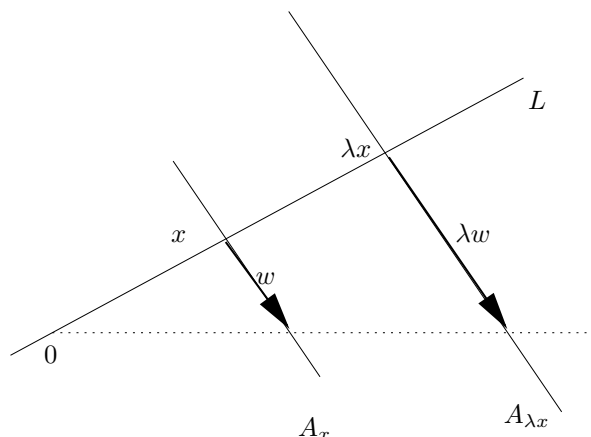


Figura 2.4: Identificación del espacio tangente $T_L \mathbb{P}(W)$ a L .

$\mathcal{U}(W)$) actúa de manera natural sobre $\mathbb{P}(W)$. Es fácil ver que si $\sigma \in \mathcal{U}(W)$ entonces el mapa $\hat{\sigma} : \mathbb{P}(W) \rightarrow \mathbb{P}(W)$ es una isometría (de variedades Riemannianas) de $\mathbb{P}(W)$. Eso se deduce fácilmente mostrando que el diferencial $D_w \hat{\sigma}$ en el punto $w \in \mathbb{P}(W)$ coincide con la restricción de σ a T_w .

Observación

Podemos definir el espacio proyectivo de igual manera en el caso de que W sea un espacio vectorial sobre \mathbb{R} . De esta manera obtenemos que $\mathbb{P}(W)$ es una variedad Riemanniana donde la métrica está definida en (2.2). Es fácil ver que la métrica definida en (2.2) coincide con la métrica en $\mathbb{P}(W)$ inducida por el mapa de cubrimiento $Q|_{S_1(W)} : S_1(W) \rightarrow \mathbb{P}(W)$.

2.2. Polinomios Homogéneos

Decimos que un polinomio f de grado d en n variables complejas es homogéneo si satisface

$$f(\lambda z_1, \dots, \lambda z_n) = \lambda^d f(z_1, \dots, z_n) \quad \forall z \in \mathbb{C}^n, \lambda \in \mathbb{C}.$$

Es fácil ver que si $f : \mathbb{C}^n \rightarrow \mathbb{C}$ es homogéneo de grado d , entonces podemos escribir

$$f(z_1, \dots, z_n) = \sum_{\|j\|=d} a_j z^j$$

(ver notación en la página 8). Observar que si $\zeta \in \mathbb{C}^n$ es raíz de f entonces toda la recta por el origen que contiene a ζ también es solución. Esta propiedad de los polinomios homogéneos resulta de mucha ayuda a la hora de estudiar las soluciones de $f(z) = 0$ con $z \in \mathbb{C}^n$, pues podemos restringir nuestro estudio a S^{2n-1} (esfera unitaria de \mathbb{C}^n) o $\mathbb{P}(\mathbb{C}^n)$ (espacio proyectivo de \mathbb{C}^n) que son variedades compactas. De esta manera, si f varía en alguna familia de polinomios homogéneos, sus raíces no se “escapan” a infinito.

Sea $\mathcal{P}_d = \mathcal{P}_d^{\mathbb{C}}$ el espacio lineal formado por los polinomios en n variables complejas de grado menor o igual a d . Es claro que podemos encontrar polinomios en \mathcal{P}_d tales que sus raíces se

encuentren arbitrariamente lejos del origen (propiedad no muy deseable). Una forma de intentar solucionar este tipo de problema es lo que se denomina *homogeneizar* el polinomio. La idea es agregar una nueva variable de manera tal que el polinomio resultante sea homogéneo. Más precisamente, si

$$f(z_1, \dots, z_n) = \sum_{\|j\| \leq d} a_j z_1^{j_1} \cdots z_n^{j_n}$$

definimos

$$\hat{f}(z_0, z_1, \dots, z_n) = \sum_{\|j\| \leq d} a_j z_0^{d-\|j\|} z_1^{j_1} \cdots z_n^{j_n}. \quad (2.3)$$

De esta manera, si $(z_1, \dots, z_n) \in \mathbb{C}^n$ es raíz de f entonces $(1, z_1, \dots, z_n) \in \mathbb{C}^{n+1}$ es raíz de \hat{f} , y por lo tanto toda la recta por el punto $(1, z_1, \dots, z_n)$ y el origen es solución (ver Figura 2.5) y

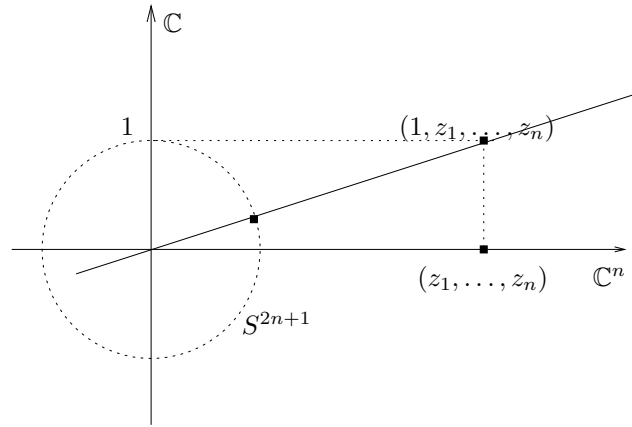


Figura 2.5: Homogeneización

podemos estudiar las raíces en S^{2n+1} o $\mathbb{P}(\mathbb{C}^{n+1})$ como comentamos antes.

De (2.3) resulta que el polinomio homogeneizado, agrega la nueva solución $(1, 0, \dots, 0)$ sobre $\mathbb{P}(\mathbb{C}^{n+1})$ si y solamente si el polinomio original tiene grado efectivo estrictamente menor que d .

En lo que sigue, notaremos por $\mathcal{H}_d = \mathcal{H}_d^{\mathbb{C}}$ al conjunto de polinomios homogéneos de grado d en $n + 1$ variables complejas. Asumimos que $0 \in \mathcal{H}_d$, y por lo tanto podemos darle a \mathcal{H}_d una estructura de espacio lineal de manera natural.

Notaremos por

$$\Lambda : \mathcal{P}_d \rightarrow \mathcal{H}_d, \quad \text{al mapa } f \mapsto \hat{f}. \quad (2.4)$$

Es fácil ver que Λ es un isomorfismo entre espacios vectoriales, donde la inversa esta dada por

$$\Lambda^{-1}g(z_0, \dots, z_n) = g(1, z_1, \dots, z_n) \quad \forall g \in \mathcal{H}_d.$$

Dada la n -upla $d = (d_1, \dots, d_n)$ de enteros positivos definimos $\mathcal{P}_{(d)} = \mathcal{P}_{(d)}^{\mathbb{C}}$ como el espacio lineal de los sistemas polinomiales $f : \mathbb{C}^n \rightarrow \mathbb{C}^n$ de la forma $f = (f_1, \dots, f_n)$ donde $f_i \in \mathcal{P}_{d_i}$ para $i = 1, \dots, n$.

Análogamente definimos $\mathcal{H}_{(d)} = \mathcal{H}_{(d)}^{\mathbb{C}}$ como el espacio lineal de sistemas polinomiales $f : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^n$ de la forma $f = (f_1, \dots, f_n)$ donde cada $f_i : \mathbb{C}^{n+1} \rightarrow \mathbb{C}$ es un polinomio homogéneo

de grado d_i ($f_i \in \mathcal{H}_{d_i}$) para $i = 1, \dots, n$. Es decir, $\mathcal{P}_{(d)} = \mathcal{P}_{d_1} \times \dots \times \mathcal{P}_{d_n}$ y $\mathcal{H}_{(d)} = \mathcal{H}_{d_1} \times \dots \times \mathcal{H}_{d_n}$ donde $\mathcal{H}_{d_i} = \Lambda(\mathcal{P}_{d_i})$ definido en (2.4).

Definimos de manera análoga los espacio $\mathcal{H}_d^{\mathbb{R}}$, $\mathcal{P}_d^{\mathbb{R}}$, $\mathcal{H}_{(d)}^{\mathbb{R}}$, $\mathcal{P}_{(d)}^{\mathbb{R}}$ cuando la variable y los coeficientes son reales.

2.3. Estructura Hermitiana de \mathcal{H}_d e Invarianza Unitaria

Aquí \mathcal{H}_d indica el espacio de polinomios homogéneos en k variables complejas.

Teorema 1. *Existe un producto interno Hermitiano en \mathcal{H}_d invariante por el grupo unitario $\mathcal{U}(k)$. Es decir,*

$$\langle f \circ U, g \circ U \rangle = \langle f, g \rangle \quad \forall f, g \in \mathcal{H}_d \text{ y } U \in \mathcal{U}(k).^1$$

Demostración. La idea de la prueba es mostrar que el espacio vectorial \mathcal{H}_d es isomorfo a un subespacio lineal de las transformaciones d -multilineales sobre el espacio \mathbb{C}^k , donde se puede definir un producto interno (conocida como *producto interno de Frobenius*) que al enviarla por el isomorfismo cumple lo requerido.

Sean V y W espacios vectoriales Hermitianos de dimensión finita, $\mathcal{V} = \{v_1, \dots, v_m\}$ y $\mathcal{W} = \{w_1, \dots, w_n\}$ bases ortonormales respectivas. Notamos por $\mathcal{L}(V, W)$ al espacio lineal formado por todos los operadores lineales de V en W . Dados $L \in \mathcal{L}(V, W)$ y $M \in \mathcal{L}(V, W)$ notamos por $[L]$ e $[M]$ sus matrices asociadas en las bases \mathcal{V} , \mathcal{W} . Definimos entonces el producto interno de Frobenius en $\mathcal{L}(V, W)$ como

$$\langle L, M \rangle = \text{tr}([L][M]^*).$$

donde $[M]^*$ es la matriz adjunta de $[M]$. Observar que si $[L] = ((a_{ji}))$ y $[M] = ((b_{ji}))$ para $j = 1, \dots, m$; $i = 1, \dots, n$ entonces

$$\langle L, M \rangle = \sum_{\substack{j=1, \dots, m \\ i=1, \dots, n}} a_{ji} \overline{b_{ji}}.$$

En otras palabras, si definimos $L_{ij} \in \mathcal{L}(V, W)$ como la transformación que manda v_i en w_j , y los demás vectores de la base al 0, se tiene que $\{L_{ij} : i = 1, \dots, m; j = 1, \dots, n\}$ es una base de $\mathcal{L}(V, W)$, y si escribimos $L = \sum_{i,j} a_{ij} L_{ij}$ e $M = \sum_{i,j} b_{ij} L_{ij}$ obtenemos

$$\langle L, M \rangle = \sum_{\substack{i=1, \dots, m \\ j=1, \dots, n}} a_{ij} \overline{b_{ji}} \tag{2.5}$$

Con esta definición hacemos que $\{L_{ij}\}$ sea una base ortonormal de $\mathcal{L}(V, W)$.

Sea $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ el espacio vectorial de las funciones d -multilineales sobre \mathbb{C}^k , es decir, si $f \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ entonces

$$f : \underbrace{\mathbb{C}^k \times \dots \times \mathbb{C}^k}_d \rightarrow \mathbb{C}$$

¹En la sección *Formalización del Nuevo Enfoque y Extensiones* del capítulo 4, incluimos una demostración más corta de este hecho.

y es lineal en cada una de las coordenadas. Observar que $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ se puede identificar de manera natural como $\mathcal{L}(\mathbb{C}^k, \mathcal{L}_{d-1}(\mathbb{C}^k, \mathbb{C}))$. Con esta identificación podemos definir un producto interno Hermitiano en $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$.

Sea $X_j^{(i)} \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ con $i = 1, \dots, m; j = 1, \dots, d$ definido por

$$X_j^{(i)}(v_1, \dots, v_d) = v_j^{(i)} \quad \text{para } (v_1, \dots, v_d) \in \mathbb{C}^k \times \dots \times \mathbb{C}^k$$

donde notamos por $v_j^{(i)}$ la i -ésima coordenada de $v_j \in \mathbb{C}^k$ para $j = 1, \dots, d$. Es fácil ver que

$$\tau_{i_1, \dots, i_d}^{(d)} := \prod_{j=1}^d X_j^{(i_j)} \quad (2.6)$$

(producto puntual) define una base de $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ al variar (i_1, \dots, i_d) en $k^d := \{1, \dots, k\}^d$. Ahora podemos definir un producto interno Hermitiano en $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ declarando al conjunto $\{\tau_{i_1, \dots, i_d}^{(d)} : (i_1, \dots, i_d) \in k^d\}$ como una base ortonormal. Veamos que este producto interno coincide con el definido en (2.5) cuando identificamos $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ con $\mathcal{L}(\mathbb{C}^k, \mathcal{L}_d(\mathbb{C}^k, \mathbb{C}))$ (notaremos por $\Lambda : \mathcal{L}_d(\mathbb{C}^k, \mathbb{C}) \rightarrow \mathcal{L}(\mathbb{C}^k, \mathcal{L}_d(\mathbb{C}^k, \mathbb{C}))$ al isomorfismo inducido). Para eso definamos $L_{l, i_1, \dots, i_{d-1}} : \mathbb{C}^k \rightarrow \mathcal{L}_{d-1}(\mathbb{C}^k, \mathbb{C})$ como

$$L_{l, i_1, \dots, i_{d-1}}(e_j) = \begin{cases} 0 & j \neq l \\ \tau_{i_1, \dots, i_d}^{(d-1)} & j = l, \end{cases} \quad (2.7)$$

es decir, como el elemento de la base que mapea $e_l \in \mathbb{C}^k$ en $\tau_{i_1, \dots, i_d}^{(d-1)}$ y los demás vectores de la base al 0 (como antes definido). Observar que $\Lambda(\tau_{i_1, \dots, i_d}^{(d)}) = L_{i_d, i_1, \dots, i_{d-1}}$ (basta ver que $\Lambda(\tau_{i_1, \dots, i_d}^{(d)})(e_j) = \tau_{i_1, \dots, i_d}^{(d)}(\cdot, \dots, \cdot, e_j)$ y comparar con (2.7)).

Lema 1. Sean V y W dos espacios vectoriales Hermitianos, $U_1 : V \rightarrow V$, $U_2 : W \rightarrow W$ operadores unitarios en V y en W respectivamente. Entonces los mapas

$$L \mapsto L \circ U_1 \quad \text{y} \quad L \mapsto U_2 \circ L \quad \forall L \in \mathcal{L}(V, W)$$

preservan la estructura Hermitiana en $\mathcal{L}(V, W)$.

Dem. de Lema 1. Eligiendo bases ortonormales en V y en W se tiene

$$\begin{aligned} \langle U_2 \circ L, U_2 \circ M \rangle &= \text{tr}([U_2 \circ L][U_2 \circ M]^*) \\ &= \text{tr}([U_2][L][M]^*[U_2]^*) \\ &= \text{tr}(\underbrace{[U_2]^*[U_2]}_{Id_W}[L][M]^*) \\ &= \text{tr}([L][M]^*) = \langle L, M \rangle, \end{aligned}$$

donde $[M]^*$ indica la matriz asociada al operador autoadjunto de M . Análogamente se hace para $L \circ U_1$ pero usando ahora que $U_1 \circ U_1^* = Id_V$ que vale pues U_1 es unitario y por lo tanto $U_1^{-1} = U_1^*$. \square

Sea $\mathcal{U}(k)^d$ el producto cartesiano de d copias de $\mathcal{U}(k)$. Definimos la acción de $\mathcal{U}(k)^d$ sobre $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ de la siguiente manera:

-dado (U_1, \dots, U_d) en $\mathcal{U}(k)^d$, definimos

$$(U_1, \dots, U_d) \cdot L(v_1, \dots, v_d) := L(U_1 v_1, \dots, U_d v_d). \quad (2.8)$$

Proposición 1. *La acción dada en (2.8) preserva la estructura Hermitiana de $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$.*

Dem. de Proposición 1. Hagámoslo por inducción. Para $d = 1$ resulta directamente del Lema 1. Supongamos que vale para $d - 1$. Identifiquemos $L \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ con $\mathcal{L}(\mathbb{C}^k, \mathcal{L}_d(\mathbb{C}^k, \mathbb{C}))$ a través del mapa Λ . Veamos como es ahora la acción de $\mathcal{U}(k)^d$ en $\mathcal{L}(\mathbb{C}^k, \mathcal{L}_d(\mathbb{C}^k, \mathbb{C}))$.

$$\begin{aligned} (U_1, \dots, U_d) \cdot [\Lambda(L)](e_j)(v_1, \dots, v_{d-1}) &= L(U_1 v_1, \dots, U_d v_d) \\ &= \Lambda(L)(U_d e_j)(U_1 v_1, \dots, U_{d-1} v_{d-1}) \\ &= (U_1, \dots, U_{d-1}) \cdot \underbrace{[\Lambda(L)(U_d e_j)]}_{\in \mathcal{L}_{d-1}(\mathbb{C}^k, \mathbb{C})}(v_1, \dots, v_{d-1}). \end{aligned}$$

es decir

$$(U_1, \dots, U_d) \cdot [\Lambda(L)] = (U_1, \dots, U_{d-1}) \cdot [\Lambda(L) \circ U_d].$$

Sean $L, M \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$, entonces

$$\begin{aligned} \langle \hat{U} \cdot L, \hat{U} \cdot M \rangle &:= \langle (U_1, \dots, U_d) \cdot [\Lambda(L)], (U_1, \dots, U_d) \cdot [\Lambda(M)] \rangle \\ &= \langle (U_1, \dots, U_{d-1}) \cdot [\Lambda(L) \circ U_d], (U_1, \dots, U_{d-1}) \cdot [\Lambda(M) \circ U_d] \rangle \\ &= \langle \Lambda(L) \circ U_d, \Lambda(M) \circ U_d \rangle \\ &= \langle \Lambda(L), \Lambda(M) \rangle = \langle L, M \rangle \end{aligned}$$

donde la tercera y cuarta igualdad resultan del Lema 1. □

Decimos que $L \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ es *simétrica* si

$$L(v_1, \dots, v_d) = L(v_{\sigma(1)}, \dots, v_{\sigma(d)})$$

para toda permutación σ del conjunto $\{1, \dots, d\}$.

Sea $\mathcal{S}_d(\mathbb{C}^k, \mathbb{C})$ el subespacio de transformaciones d -multilineales simétricas en $\mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$.

Proposición 2. *Existe un isomorfismo lineal $\theta : \mathcal{S}_d(\mathbb{C}^k, \mathbb{C}) \rightarrow \mathcal{H}_d$.*

Dem. de Proposición 2. Dado $L \in \mathcal{S}_d(\mathbb{C}^k, \mathbb{C})$ definimos

$$\theta(L)(x) := L(x, \dots, x) \quad \forall x \in \mathbb{C}^k. \quad (2.9)$$

Observar que (2.9) define un polinomio homogéneo de grado d pues

$$\theta(L)(\lambda x) = L(\lambda x, \dots, \lambda x) = \lambda^d L(x, \dots, x), \quad \forall \lambda \in \mathbb{C}.$$

Definimos $\rho : \mathcal{H}_d \rightarrow \mathcal{S}_d(\mathbb{C}^k, \mathbb{C})$ como

$$\rho(f) := \frac{1}{d!} D_0^d f \quad \forall f \in \mathcal{H}_d$$

donde $D_0^d f$ indica la derivada de orden d del polinomio f en 0, es decir

$$D_0^d f(e_{i_1}, \dots, e_{i_d}) = \frac{\partial^d}{\partial x_{i_1} \cdots \partial x_{i_d}} f \Big|_{x=0} \quad \forall (i_1, \dots, i_d) \in k^d$$

que es naturalmente un operador d -multilineal simétrico. Es fácil ver que ρ es lineal e inyectivo. Veamos que θ y ρ son uno inverso del otro:

-Sea $L \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$, entonces podemos escribir

$$L(v_1, \dots, v_d) = \sum_{(i_1, \dots, i_d) \in k^d} v_1^{i_1} \cdots v_d^{i_d} L(e_{i_1}, \dots, e_{i_d})$$

y por lo tanto

$$\theta(L)(x) = \sum_{(i_1, \dots, i_d) \in k^d} x_{i_1} \cdots x_{i_d} L(e_{i_1}, \dots, e_{i_d}).$$

Entonces si $(i_1, \dots, i_d) \in k^d$ se tiene

$$\begin{aligned} \rho(\theta(L))(e_{i_1}, \dots, e_{i_d}) &= \frac{1}{d!} D_0^d (\theta(L))(e_{i_1}, \dots, e_{i_d}) \\ &= L(e_{i_1}, \dots, e_{i_d}) \end{aligned}$$

de donde resulta $\rho \circ \theta = Id_{\mathcal{H}_d}$.

-Resta probar la igualdad en el otro sentido. Sea $f \in \mathcal{H}_d$, entonces escribimos f de la siguiente manera

$$f(x) = \sum_{(i_1, \dots, i_d) \in k^d} a_{(i_1, \dots, i_d)} x_{i_1} \cdots x_{i_d}.$$

Entonces

$$\theta(\rho(f))(x) = \theta\left(\frac{1}{d!} D_0^d f\right)(x) = \frac{1}{d!} D_0^d f(x, \dots, x)$$

que no es otra cosa que $f(x)$ por la fórmula de Taylor aplicada a f en 0. \square

A través de este isomorfismo lineal podemos definir un producto interno en \mathcal{H}_d . Dados $f, g \in \mathcal{H}_d$, definimos

$$\langle f, g \rangle_{\mathcal{H}_d} := \langle \theta^{-1}(f), \theta^{-1}(g) \rangle.$$

Solo nos falta chequear que este producto Hermitiano en \mathcal{H}_d verifica la propiedad de invarianza bajo la acción del grupo $\mathcal{U}(k)$. Dado $U \in \mathcal{U}(k)$ definimos

$$\widehat{U} := \underbrace{(U, \dots, U)}_d.$$

Observando que si $f = \theta(L)$ y $g = \theta(M)$ con $L, M \in \mathcal{L}_d(\mathbb{C}^k, \mathbb{C})$ se tiene

$$f \circ U = \theta(\widehat{U} \cdot L)$$

ya que $\theta(\widehat{U} \cdot L)(x) = \widehat{U} \cdot L(x, \dots, x) = L(Ux, \dots, Ux) = \theta(L)(Ux) = f \circ U(x)$, entonces concluimos

$$\begin{aligned} \langle f \circ U, g \circ U \rangle_{\mathcal{H}_d} &= \langle \theta(\widehat{U} \cdot L), \theta(\widehat{U} \cdot M) \rangle_{\mathcal{H}_d} \\ &= \langle \widehat{U} \cdot L, \widehat{U} \cdot M \rangle \\ &\stackrel{Prop.1}{=} \langle L, M \rangle \\ &\stackrel{(def)}{=} \langle f, g \rangle_{\mathcal{H}_d}. \end{aligned}$$

□

Hasta ahora hemos construido un producto interno Hermitiano en \mathcal{H}_d al cual hemos notado por $\langle \cdot, \cdot \rangle_{\mathcal{H}_d}$. Ahora mostraremos explícitamente dicho producto interno.

Teorema 2. *Dados f y g en \mathcal{H}_d se tiene*

$$\langle f, g \rangle_{\mathcal{H}_d} = \sum_{\|j\|=d} a_j \bar{b}_j \binom{d}{j_1, \dots, j_d}^{-1} \quad (2.10)$$

donde $f(x) = \sum_{\|j\|=d} a_j x^j$, $g(x) = \sum_{\|j\|=d} b_j x^j$, y

$$\binom{d}{j_1, \dots, j_d} = \frac{d!}{j_1! \cdots j_d!}.$$

Demostración. Es fácil ver que solo necesitamos chequear el producto interno de f_α con f_β en \mathcal{H}_d , donde $f_\alpha(x) = x^\alpha$, $f_\beta(x) = x^\beta$ siendo $\alpha = (\alpha_1, \dots, \alpha_k)$ y $\beta = (\beta_1, \dots, \beta_k)$ multi-índices con $\|\alpha\| = \|\beta\| = d$.

Por definición $\rho(f_\alpha) = \frac{1}{d!} D_0^d f_\alpha$, entonces

$$\rho(f_\alpha)(e_{i_1}, \dots, e_{i_d}) = \frac{1}{d!} \frac{\partial^d}{\partial x_{i_1} \cdots \partial x_{i_d}} f_\alpha \Big|_{x=0}.$$

Escribimos

$$f_\alpha(x) = \frac{1}{\#\mathcal{F}(\alpha)} \sum_{(i_1, \dots, i_d) \in \mathcal{F}(\alpha)} x_{i_1} \cdots x_{i_d}$$

donde

$$\mathcal{F}(\alpha) = \{(i_1, \dots, i_d) \in k^d : \#\{j : i_j = l\} = \alpha_l, l = 1, \dots, k\}.$$

Es fácil observar que

$$\#\mathcal{F}(\alpha) = \binom{d}{j_1, \dots, j_d}.$$

Entonces tenemos que

$$\rho(f_\alpha) = \frac{1}{\binom{d}{j_1, \dots, j_d}} \sum_{(i_1, \dots, i_d) \in \mathcal{F}(\alpha)} \tau_{(i_1, \dots, i_d)}^{(d)}$$

donde $\tau_{(i_1, \dots, i_d)}^{(d)} = \prod_{j=1}^d X_j^{i_j}$ estaba definido en (2.6). Por lo tanto se tiene:

- si $\alpha \neq \beta$

$$\begin{aligned} \langle f_\alpha, f_\beta \rangle_{\mathcal{H}_d} &= \frac{1}{\binom{d}{\alpha_1, \dots, \alpha_d} \cdot \binom{d}{\beta_1, \dots, \beta_d}} \left\langle \sum_{(i_1, \dots, i_d) \in \mathcal{F}(\alpha)} \tau_{(i_1, \dots, i_d)}^{(d)}, \sum_{(i'_1, \dots, i'_d) \in \mathcal{F}(\beta)} \tau_{(i'_1, \dots, i'_d)}^{(d)} \right\rangle \\ &= \frac{1}{\binom{d}{\alpha_1, \dots, \alpha_d} \cdot \binom{d}{\beta_1, \dots, \beta_d}} \sum_{\substack{(i_1, \dots, i_d) \in \mathcal{F}(\alpha) \\ (i'_1, \dots, i'_d) \in \mathcal{F}(\beta)}} \langle \tau_{(i_1, \dots, i_d)}^{(d)}, \tau_{(i'_1, \dots, i'_d)}^{(d)} \rangle = 0 \end{aligned}$$

pues $\mathcal{F}(\alpha) \cap \mathcal{F}(\beta) = \emptyset$ si los consideramos como subconjuntos de k^d .

- si $\alpha = \beta$ entonces

$$\begin{aligned} \langle f_\alpha, f_\alpha \rangle_{\mathcal{H}_d} &= \frac{1}{\binom{d}{\alpha_1, \dots, \alpha_d}^2} \left\langle \sum_{(i_1, \dots, i_d) \in \mathcal{F}(\alpha)} \tau_{(i_1, \dots, i_d)}^{(d)}, \sum_{(i_1, \dots, i_d) \in \mathcal{F}(\alpha)} \tau_{(i_1, \dots, i_d)}^{(d)} \right\rangle \\ &= \frac{1}{\binom{d}{\alpha_1, \dots, \alpha_d}^2} \cdot \#\mathcal{F}(\alpha) \\ &= \binom{d}{\alpha_1, \dots, \alpha_d}^{-1}. \end{aligned}$$

Aplicando la linealidad (conjugada) del producto Hermitiano se prueba (2.10). \square

Observación

Hemos introducido un producto interno en $\mathcal{H}_{(d)}$ invariante por $\mathcal{U}(n+1)$. De esta manera conseguimos una métrica Hermitiana en $\mathbb{P}(\mathcal{H}_{(d)})$ que es invariante bajo la acción de $\mathbb{P}(\mathcal{U}(n+1))$ siendo éste el grupo de isometrías en $\mathbb{P}(\mathcal{H}_{(d)})$ inducido por $\mathcal{U}(n+1)$ en $\mathcal{H}_{(d)}$ (ver *Observación* en página 13). Esta métrica definida en $\mathbb{P}(\mathcal{H}_{(d)})$ es conocida en la literatura como la métrica de Fubini-Study.

2.4. Teorema de Bézout

El Teorema de Bézout es una generalización del Teorema Fundamental del Álgebra. Antes de comenzar debemos establecer un poco de notación.

- Al igual que antes, \mathcal{H}_d es el espacio de polinomios homogéneos en $n+1$ variables complejas de grado d .
- Dada la k -upla de enteros positivos $(d) := (d_1, \dots, d_k)$ definimos el espacio $\mathcal{H}_{(d)}$ como el producto

$$\mathcal{H}_{(d)} := \mathcal{H}_{d_1} \times \cdots \times \mathcal{H}_{d_k}.$$

- Dado W un espacio vectorial notaremos por $\mathbb{P}(W)$ el correspondiente espacio proyectivo.
- Dado $w \in \mathbb{P}(W)$, el espacio tangente $T_w \mathbb{P}(W)$ a w lo identificamos con T_w (ver página 13)

Definamos la *variedad solución (proyectiva)* como el conjunto $V \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$ dado por

$$V := \{(f, x) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1}) : f(x) = 0\}.$$

Aquí, para no hacer muy pesada la notación, estamos notando por $f \in \mathbb{P}(\mathcal{H}_{(d)})$ y $x \in \mathbb{P}(\mathbb{C}^{n+1})$ como elementos representativos de la clase. Observar que V está correctamente definido en el espacio proyectivo pues estamos trabajando con polinomios homogéneos, entonces $f(x) = 0$ si y solo si $\lambda f(\mu x) = 0$ para todo $\lambda, \mu \in \mathbb{C}^*$.

Proposición 3. *La variedad solución V es una variedad diferenciable conexa encajada en $\mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$ de codimensión k . Además el espacio tangente al punto (f, x) (que notaremos como es usual $T_{(f,x)}V$) está dado por*

$$T_{(f,x)}V = \{(h, w) \in T_f\mathbb{P}(\mathcal{H}_{(d)}) \times T_x\mathbb{P}(\mathbb{C}^{n+1}) : h(x) + d_x f(w) = 0\}.$$

Demostración. Sea $\hat{V} \subset \mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*}$ definido por²

$$\hat{V} := \{(f, x) \in \mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*} : f(x) = 0\},$$

y $\epsilon : \mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*} \rightarrow \mathbb{C}^k$ el mapa evaluación definido por

$$\epsilon(f, x) = f(x).$$

Por lo tanto, $\hat{V} = \epsilon^{-1}(\{0\})$.³

Veamos que $0 \in \mathbb{C}^k$ es un valor regular de ϵ . Basta ver que $\forall (f, x) \in \hat{V}$ el mapa tangente

$$d_{(f,x)}\epsilon : T_{(f,x)}(\mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*}) \rightarrow \mathbb{C}^k$$

es sobreyectivo. Observando que ϵ respecto a la primer coordenada es lineal, se tiene

$$d_{(f,x)}\epsilon(h, w) = h(x) + d_x f(w) \quad \forall (h, w) \in T_f\mathcal{H}_{(d)} \times T_x\mathbb{C}^{n+1}.$$

Es fácil ver que en estas condiciones $d_{(f,x)}\epsilon \left(T_{(f,x)}(\mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*}) \right) = \mathbb{C}^k$ (basta considerar $w = 0$, y elegir adecuadamente una familia de sistemas $\mathcal{F} \subset \mathcal{H}_{(d)}$ tal que $h(x)$ recorra todo \mathbb{C}^k cuando h recorre \mathcal{F}). Entonces 0 es un valor regular y \hat{V} es una variedad diferenciable encajada en $\mathcal{H}_{(d)}^* \times \mathbb{C}^{n+1^*}$ de codimensión k . Además es claro que el espacio tangente coincide con el núcleo del mapa tangente de ϵ . Entonces

$$T_{(f,x)}\hat{V} = \{(h, w) : h(x) + d_x f(w) = 0\}.$$

Veamos que \hat{V} es conexo:

Para eso basta ver que si $(h, y) \in \hat{V}$ entonces fijado $x \in \mathbb{C}^{n+1^*}$, existe una curva $\{\Lambda_t\} =$

²Dado un espacio vectorial X notamos por $X^* := X \setminus 0$.

³Observar que en principio, V es lo que se llama una *variedad algebraica proyectiva* pues es la proyección de la intersección de los ceros de una cantidad finita de polinomios homogéneos definidos en $\mathcal{H}_{(d)} \times \mathbb{C}^{n+1}$, donde identificamos $\mathcal{H}_{(d)}$ con \mathbb{C}^N siendo N la cantidad de coeficientes de un sistema en $\mathcal{H}_{(d)}$ (ver disgresión en la página 25 sobre variedades algebraicas).

$\{(h_t, x_t)\} \subset \hat{V}$ para $t \in [0, 1]$, donde $\Lambda_0 = (h, y)$ y $x_1 = x$. En otras palabras, dado $(h, y) \in \hat{V}$ encontraremos un camino en \hat{V} que empieza en (h, y) y termina en (f, x) para algún $f \in \mathcal{H}_{(d)}$. Una vez que tenemos esto es fácil conectar (mediante un camino en \hat{V}) dos pares (h, y) y (g, x) de puntos en \hat{V} . Para eso primero transportamos el primer par (h, y) a uno de la forma (f, x) , y luego modificamos linealmente el sistema para llevarlo de f a g . Basta tomar la curva $\{(f + t(g - f), x)\}_{t \in [0, 1]}$, que esta incluida en \hat{V} ya que $(f + t(g - f))(x) = 0$.

Veamos que existe siempre tal curva:

Sea $\varphi_t : \mathbb{C}^{n+1} \rightarrow \mathbb{C}^{n+1}$ para $t \in [0, 1]$ la deformación lineal continua e invertible que verifique $\varphi_0(y) = y$ y $\varphi_1(y) = x$.⁴ Definamos h_t como la composición $h \circ \varphi_t^{-1}$ donde es fácil chequear que $h_t \in \mathcal{H}_{(d)}$ para todo $t \in [0, 1]$. Sea $x_t = \varphi_t(x)$, entonces $\Lambda_t = (h_t, x_t)$ verifica que $\Lambda_0 = (h, x)$, $\Lambda_1 = (h_1, y)$ y

$$h_t(x_t) = h \circ \varphi_t^{-1}(x_t) = h(y) = 0$$

y por lo tanto $\{\Lambda_t\}$ es un camino en \hat{V} que verifica lo deseado.

Con esto, la conexidad de V es muy simple. Consideremos la acción en \hat{V} por el grupo multiplicativo $\mathbb{C}^* \times \mathbb{C}^*$, donde la acción esta definida por:

$$(\lambda, \mu) \cdot (f, x) = (\lambda f, \mu x) \quad (\lambda, \mu) \in \mathbb{C}^* \times \mathbb{C}^* \tag{2.11}$$

Observar que el grupo actúa libremente⁵ y además deja invariante a \hat{V} . Cocientando \hat{V} por dicha acción obtenemos V , luego V es conexo.

Resta por probar que V es una variedad diferenciable. Un primer intento es considerar la acción definida en (2.11) y cocientar V por la acción identificando las órbitas. Pero el problema que surge es que cuando cocientamos una variedad diferenciable por un grupo no compacto (que actúa libremente) no implica que el cociente sea una variedad, ni siquiera un espacio topológico Hausdorff. Un ejemplo simple es considerar en el toro plano \mathbb{T}^2 el cociente por rectas de igual pendiente irracional (basta identificar las órbitas del flujo lineal en el plano generado por la ecuación diferencial $\dot{X} = (1, \sqrt{2})X$) De esa manera la órbita de un punto se enrolla densamente en el toro y es fácil ver que el cociente por esas orbitas no es Hausdorff.

Sí es verdad, que si cocientamos una variedad diferenciable por un grupo compacto que actúa libremente entonces el cociente resulte una variedad diferenciable. Intentemos por ese lado. Considerando la norma $\|\cdot\|_{d_i}$ en \mathcal{H}_{d_i} definida en (2.10), definimos la norma $\|\cdot\|$ en $\mathcal{H}_{(d)}$ inducida por el producto, es decir,

si $f = (f_1, \dots, f_k)$ entonces

$$\|f\| = \|f_1\|_{d_1} + \dots + \|f_k\|_{d_k}.$$

Más precisamente si $f_i(x) = \sum_{\|j\|=d_i} a_j^{(i)} x^j$ entonces

$$\|f\|^2 = \sum_{i=1}^k \sum_{\|j\|=d_i} |a_j^{(i)}|^2 \binom{d_i}{j}^{-1}.$$

⁴Si y no ortogonal a x con el producto interno usual de \mathbb{C}^{n+1} , basta definir φ_t como $\varphi_t(z) = A_t z$, donde A_t es la matriz $(n+1) \times (n+1)$ dada en una base ortonormal $\{y/\|y\|, v_1, \dots, v_n\}$ de \mathbb{C}^{n+1} para ciertos vectores columnas v_1, \dots, v_n , de la siguiente forma: $A_t = [y/\|y\| + t(x/\|y\| - y/\|y\|), v_1, \dots, v_n]$. Si son ortogonales, basta homotetizar la dirección que contenga a y hasta que tenga la misma norma que x , y luego rotarlo continuamente con una curva de transformaciones unitarias.

⁵Un grupo G que actúa sobre un conjunto X se dice que *actúa libremente* si $g \cdot x = g' \cdot x$ implica $g = g'$.

Sea $S_1(\mathcal{H}_{(d)})$ y $S_1(\mathbb{C}^{n+1})$ las esferas unitarias de $\mathcal{H}_{(d)}$ y \mathbb{C}^{n+1} respectivamente, y $Q : \mathcal{H}_{(d)} \times \mathbb{C}^{n+1} \rightarrow \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$ el mapa cociente. Entonces observar que

$$Q\left(\hat{V} \cap S_1(\mathcal{H}_{(d)}) \times S_1(\mathbb{C}^{n+1})\right) = V. \quad (2.12)$$

No es difícil chequear que $\hat{V} \cap S_1(\mathcal{H}_{(d)}) \times S_1(\mathbb{C}^{n+1})$ es una variedad diferenciable debido a la transversalidad de la intersección⁶.

Sea $S^1 = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$, observando (2.12) resulta que si cocientamos por la acción de S^1 obtenemos V . Pero como el cociente de una variedad diferenciable por la acción libre de S^1 (que es un grupo compacto) es una variedad diferenciable se tiene lo que buscábamos.

Observar que la codimensión se mantiene al cocientar. \square

De ahora en adelante consideraremos $k = n$, y por lo tanto $\dim V = \dim \mathbb{P}(\mathcal{H}_{(d)})$.

Sea $\pi_1 : V \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1}) \rightarrow \mathbb{P}(\mathcal{H}_{(d)})$ la proyección en la primer coordenada, y Σ' el conjunto de puntos críticos de π_1 .

Estudiemos Σ' :

$(f, x) \in V$ es un punto crítico de π_1 si el mapa diferencial $d_{(f,x)}\pi_1 : T_{(f,x)}V \rightarrow T_f\mathbb{P}(\mathcal{H}_{(d)})$ no es sobreyectivo. Observar que

$$d_{(f,x)}\pi_1(h, w) = h$$

donde

$$(h, w) \in T_{(f,x)}V = \{(h, w) \in T_f\mathbb{P}(\mathcal{H}_{(d)}) \times T_x\mathbb{P}(\mathbb{C}^{n+1}) : h(x) + d_x f(w) = 0\}.$$

Entonces, que (f, x) sea un punto crítico es equivalente a que exista $w \in T_x\mathbb{P}(\mathbb{C}^{n+1})$ con $w \neq 0$ tal que $d_x f(w) = 0$. Para convencerse razonemos por absurdo. Si no existiera tal w veamos que el mapa tiene que ser sobreyectivo. Si $h \in T_f\mathbb{P}(\mathcal{H}_{(d)}) = \{g \in \mathcal{H}_{(d)} : \langle f, g \rangle = 0\}$ y $h(x) = 0$, tomando $w = 0$ se tiene que $(h, 0) \in T_{(f,x)}V$. Si $h(x) \neq 0$ como $d_x f$ es sobreyectivo existe $w \in T_x\mathbb{P}(\mathbb{C}^{n+1})$ tal que $d_x f(w) = -h(x)$ y por lo tanto $(h, w) \in V$. Entonces se tiene $T_f\mathbb{P}(\mathcal{H}_{(d)}) \subset \text{Im } d_{(f,x)}\pi_1 \subset T_f\mathbb{P}(\mathcal{H}_{(d)})$ lo cual implica que (f, x) no es crítico.

De esta manera, concluimos que (f, x) es un punto crítico de π_1 si el diferencial $D_x f$ no tiene rango máximo. Equivalentemente, $(f, x) \in V$ es un punto crítico si el determinante de todos los menores complementarios de tamaño n es cero. Por lo tanto, Σ' es una *variedad algebraica proyectiva* de V . Basta tomar los polinomios $F_0, F_1, \dots, F_n : \mathcal{H}_{(d)} \times \mathbb{C}^{n+1} \rightarrow \mathbb{C}$ dados por $F_0(f, x) = \epsilon(f, x)$, y $F_i(f, x) = h_i(f, x)$ ($i = 1, \dots, n$) donde $\epsilon(f, x) := f(x)$ y h_i para $i = 1, \dots, n$ indica el determinante de los menores complementarios de tamaño n del diferencial $D_x f$ en alguna base del tangente.⁷ Entonces hemos concluido:

Lema 2. *El conjunto Σ' de pares (f, x) en la variedad solución V tales que x es un cero degenerado de f es una variedad algebraica proyectiva.*

⁶Recordar una condición necesaria y suficiente para que la intersección de dos variedades sea una variedad es que la intersección sea transversal.

⁷Es fácil chequear que $\Sigma' = \pi(\cap_0^n F_i^{-1}(0)) \subset V$, donde $\pi : \mathcal{H}_{(d)} \times \mathbb{C}^{n+1} \rightarrow \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$ es la proyección canónica, y que cada F_i es un polinomio homogéneo para $i = 0, \dots, n$.

Disgresión sobre variedades algebraicas:

Vale la pena mencionar que cuando hablamos de *variedades algebraicas* sobre \mathbb{C}^N para algún N , estamos hablando de la intersección de los ceros de una cantidad finita de polinomios definidos en \mathbb{C}^N . A diferencia de algunos textos, a esos conjuntos los llaman *conjuntos algebraicos* y reservan la definición de variedad algebraica a conjuntos algebraicos irreducibles (en el sentido de que no pueden descomponerse en unión disjunta de dos conjuntos algebraicos). Pero en este texto hablaremos de variedades algebraicas como conjuntos algebraicos, ya que la irreducibilidad no juega ningún papel de relevancia.

Cuando hablamos de *variedad algebraica proyectiva* nos referimos a la proyección de una variedad algebraica definida por polinomios homogéneos.

Definimos la *variedad discriminante* como $\Sigma = \pi_1(\Sigma')$. Entonces por lo razonado anteriormente es fácil ver que $f \in \Sigma$ si y solo si f tiene ceros degenerados.

Proposición 4. *El conjunto $\Sigma \subset \mathbb{P}(\mathcal{H}_d)$ de sistemas polinomiales con ceros degenerados, es una variedad algebraica proyectiva.*

Demostración. Solamente incluiremos la demostración de este lema para el caso particular $n = 1$, el cual se puede encontrar explícitamente el polinomio homogéneo que define Σ , y el caso general lo incluiremos en el apéndice.

Sea $f \in \mathcal{H}_d$, es decir, $f : \mathbb{C}^2 \rightarrow \mathbb{C}$ homogéneo de grado d . Notamos $f(z, w) = \sum_{j=0}^d a_j z^j w^{d-j}$. Sea $x \in \mathbb{C}^2 - \{0\}$ tal que $f(x) = 0$. Es fácil ver que $(f, x) \in \Sigma'$ si y solo si $\frac{\partial f}{\partial z}(x) = \frac{\partial f}{\partial w}(x) = 0$. Eso resulta de que $(f, x) \in \Sigma'$ si la derivada libre $D_x f$ no tiene rango máximo, y como la derivada libre en la dirección de x debe ser 0 (pues es un polinomio homogéneo), resulta que la derivada tangencial también debe anularse. Entonces $D_x f$ es la transformación nula. Antes de continuar debemos hacer un breve paréntesis para definir lo que es la *resultante* de dos polinomios.

Definición 2.5. *Dados dos polinomios homogéneos f y g en \mathbb{C}^2 de grados d y d' , dados por*

$$f(z, w) = \sum_{j=0}^d a_j z^j w^{d-j}, \quad y \quad g(z, w) = \sum_{j=0}^{d'} b_j z^j w^{d'-j},$$

*definimos la **resultante de f y g** ($\mathcal{R}(f, g)$) como el polinomio en los coeficientes de f y g que resulta de tomar el determinante de la siguiente matriz $(d + d') \times (d + d')$ (matriz de Sylvester):*

$$\begin{pmatrix} a_0 & a_1 & \dots & \dots & a_{d-1} & 0 & 0 & \dots & 0 \\ 0 & a_0 & \dots & \dots & a_{d-1} & a_d & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & a_0 & \dots & \dots & \dots & a_d \\ \hline b_0 & \dots & \dots & b_{d'} & 0 & \dots & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & b_0 & \dots & \dots & \dots & \dots & \dots & b_{d'} \end{pmatrix}$$

Por lo tanto, la resultante de dos polinomios es un nuevo polinomio en los coeficientes de los polinomios. La propiedad principal de la resultante esta dada en el siguiente lema:

Lema 3. f y g tienen una raíz común en $\mathbb{P}(\mathbb{C}^2)$ si y solo si la resultante $\mathcal{R}(f, g) = 0$.

Un caso particular de esto, es el *discriminante* de un polinomio.

Definición 2.6. Dado un polinomio homogéneo f en \mathbb{C}^2 , definimos el **discriminante** de f como $\mathcal{D}(f) := \mathcal{R}\left(\frac{\partial f}{\partial z}, \frac{\partial f}{\partial w}\right)$, que resulta ser un polinomio de grado $2d - 1$ en los coeficientes de f .

Para terminar la demostración de la *Proposición 4*, basta observar que Σ es la proyección mediante el mapa cociente de $\mathcal{D}(f)^{-1}(0) \cap \epsilon^{-1}(0)$. \square

Prueba del Lema 3. Observar que todo polinomio homogéneo en dos variables complejas se puede escribir como producto de factores lineales de la forma $\alpha z - \beta w$. Eso resulta de la siguiente igualdad algebraica $f(z, w) = w^d f(z/w, 1)$ (dividiendo en dos casos, si es necesario) y notando por $y := z/w$ como nueva variable, el resultado sigue de la factorización en \mathbb{C} del polinomio $f(y, 1)$.

Luego es fácil ver que si f y g tienen una raíz común en $\mathbb{P}(\mathbb{C}^2)$ se tiene que existen dos polinomios p_1 y p_2 de grados $d - 1$ y $d' - 1$ respectivamente, tales que $p_1 f + p_2 g = 0$.⁸ Sea el mapa $\Lambda : \mathcal{H}_{d-1} \times \mathcal{H}_{d'-1} \rightarrow \mathcal{H}_{d+d'-1}$ dado por $\Lambda(p_1, p_2) = p_1 f + p_2 g$. Observar que es una aplicación lineal entre espacios de igual dimensión ($d + d'$). Concluimos entonces que, f y g tienen una raíz común en $\mathbb{P}(\mathbb{C}^2)$ si y solo si, el mapa lineal Λ no es invertible. Resta observar que la matriz de Sylvester resulta ser la matriz asociada (traspuesta) a la transformación Λ en la bases de monomios ordenada según la potencia a la que esta elevado z . \square

Proposición 5. El conjunto $\mathbb{P}(\mathcal{H}_{(d)}) - \Sigma$ es conexo.

Demostración. Basta probarlo para el caso afín, es decir probar que $\mathcal{H}_{(d)} - \tilde{\Sigma}$ es conexo siendo $\tilde{\Sigma}$ la pre-imagen de Σ por el mapa cociente. Además, basta probar que dicho conjunto es arco-conexo. Sean f y g dos sistemas polinomiales en $\mathcal{H}_{(d)} - \tilde{\Sigma}$, y sea $F : \mathcal{H}_{(d)} \rightarrow \mathbb{C}$ uno de los polinomios homogéneos que define a $\tilde{\Sigma}$.⁹ Consideremos el segmento complejo $L := \{\lambda f + (1 - \lambda)g : \lambda \in \mathbb{C}, |\lambda| \leq 1\}$ que une f con g . Por lo tanto L es una superficie real (de dimensión 2) dependiendo del parámetro λ variando en el disco unidad $\{z \in \mathbb{C} : |z| \leq 1\}$. Entonces, definiendo $\varphi(z) := F(zf + (1 - z)g)$, que es un polinomio en una variable compleja, obtenemos que a lo sumo hay k raíces en \mathbb{C} de φ , siendo k el grado del polinomio F . Pero resulta entonces que el número de ceros de F restringido a L es finito, y por tanto podemos conectar por medio de una curva real f con g evitando esos ceros. De esta manera obtenemos una curva que conecta f con g , y tal que la restricción de F a la curva no se anula. Entonces hemos construido una curva totalmente incluida en $\mathcal{H}_{(d)} - \tilde{\Sigma}$, y por lo tanto resulta que $\mathcal{H}_{(d)} - \tilde{\Sigma}$ es arco-conexo (ver *Figura 2.6*). \square

⁸Si (z_0, w_0) es solución de f y g , entonces podemos escribir $f(z, w) = (w_0 z - z_0 w)f_1(z, w)$ y $g(z, w) = (w_0 z - z_0 w)g_1(z, w)$.

⁹Sabemos por la *Proposición 4* que existe un conjunto finito de polinomios que define $\tilde{\Sigma}$.

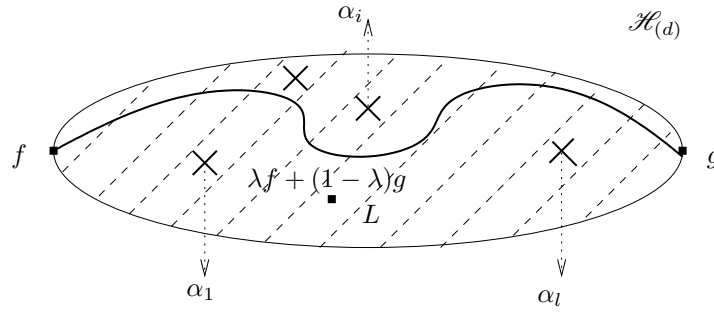


Figura 2.6: $L := \{\lambda f + (1 - \lambda)g : \lambda \in \mathbb{C}, |\lambda| \leq 1\}$, $\alpha_1, \dots, \alpha_l$ sistemas polinomiales con ceros degenerados.

Teorema 3 (Teorema de Bézout). Si $f \in \mathbb{P}(\mathcal{H}_{(d)}) - \Sigma$ entonces f tiene exactamente $\mathcal{D} := d_1 \cdots d_n$ raíces en $\mathbb{P}(\mathbb{C}^{n+1})$.

Dem. del Teorema 3. Por la conexidad de $\mathbb{P}(\mathcal{H}_{(d)}) - \Sigma$, usando el teorema de la función inversa se tiene que el número de soluciones es constante en \mathcal{D} .¹⁰ Computemos en algún caso fácil. Consideremos el sistema $f = (f_1, \dots, f_n)$ con

$$f_i(z) = z_0^{d_i} - z_i^{d_i}.$$

Entonces $f(z) = 0$ si y solo si $f_i(z) = 0$ ($i = 1, \dots, n$) si y solo si $z_0^{d_i} = z_i^{d_i}$ ($i = 1, \dots, n$). Entonces fijando $z_0 = 1$ (que lo podemos hacer pues necesariamente debe ser distinto de cero para que no nos quede la solución trivial idénticamente cero) se tiene que $f_i(1, z_1, \dots, z_n) = 0$ si z_i es raíz d_i -ésima de la unidad para $i = 1, \dots, n$. Por lo tanto $(1, z_1, \dots, z_n)$ es solución del sistema si y solo si z_i es raíz d_i -ésima para $i = 1, \dots, n$. Resultando que el número de soluciones de f es exactamente $d_1 \cdots d_n$. Resta ver que $f \in \mathbb{P}(\mathcal{H}_{(d)}) - \Sigma$. Para eso basta computar el diferencial de f en z en la base canónica y observar que si z es raíz del sistema entonces la matriz tiene rango máximo. \square

¹⁰Sea $\pi_1 : V - \Sigma' \rightarrow \mathbb{P}(\mathcal{H}_{(d)}) - \Sigma$, entonces todo elemento es un valor regular. Resulta del teorema de la función inversa que la cantidad de pre-ímagenes es localmente constante, y como cada pre-imagen es de la forma (f, ζ) con ζ raíz, se obtiene que la cantidad de soluciones es constante en un entorno de f .

Parte II

Sistemas Polinomiales Aleatorios

Capítulo 3

Teorema de Shub-Smale

3.1. Introducción

En el capítulo anterior, introducimos un producto interno en $\mathcal{H}_{(d)}$. Dicho producto interno es invariante por el grupo ortogonal $\mathcal{O}(n+1)$ en el caso real ($\mathcal{H}_{(d)}^{\mathbb{R}}$), e invariante por el grupo unitario $\mathcal{U}(n+1)$ en el caso complejo ($\mathcal{H}_{(d)}^{\mathbb{C}}$).

En esta sección nos restringiremos a estudiar el caso real, y notaremos $\mathcal{H}_{(d)} := \mathcal{H}_{(d)}^{\mathbb{R}}$.

A través de dicho producto interno, vimos que podemos dotar $\mathbb{P}(\mathcal{H}_{(d)})$ de una estructura de variedad Riemanniana de manera natural, y por lo tanto podemos definir el elemento de volúmen inducido. De esta manera se dota a $\mathbb{P}(\mathcal{H}_{(d)})$ con una estructura de espacio de medida.¹ Si $x \in \mathbb{P}(W)$ donde W es un \mathbb{R} -espacio vectorial de dimensión n con producto interno, definimos la forma de volúmen $\nu(x)$ en el punto $x \in W$ como

$$\nu(x)(w_1, \dots, w_n) := \frac{\det(x, w_1, \dots, w_n)}{\|x\|^{n+1}}$$

siendo $\{w_1, \dots, w_n\}$ vectores de $T_x\mathbb{P}(W)$ y \det es la función determinante en W . De esta manera podemos integrar funciones a valores reales en $\mathbb{P}(\mathcal{H}_{(d)})$, o en $\mathbb{P}(\mathbb{R}^{n+1})$, o por ejemplo en $V \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{R}^{n+1})$.

Recordar que si f es una función de \mathbb{R}^m en \mathbb{R}^n entonces $N_u^f(U) := \#\{x \in U : f(x) = u\}$ con $U \subset \mathbb{R}^m$ y $N_u^f := N_u^f(\mathbb{R}^m)$.

Con estas precisiones estamos en condiciones de enunciar el Teorema de Shub-Smale

Teorema 4 (Teorema de Shub-Smale [22]). *El número promedio de soluciones reales en $\mathbb{P}(\mathbb{R}^{n+1})$ de $f \in \mathbb{P}(\mathcal{H}_{(d)})$ es $\mathcal{D}^{1/2}$, donde \mathcal{D} es el número de Bézout.*

Esto significa que

$$\frac{1}{\text{Vol}(\mathbb{P}(\mathcal{H}_{(d)}))} \cdot \int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f d\mathbb{P}(\mathcal{H}_{(d)}) = \mathcal{D}^{1/2}.$$

¹Una buena referencia donde se discute sobre como inducir medidas sobre variedades diferenciables, es [14]. En el caso de que la variedad tenga una estructura Riemanniana, siempre es posible definir una medida. Basta definir una forma de volúmen (que se define en cada espacio tangente de manera natural a través de la métrica), y luego integrar respecto a ella.

En esta sección demostraremos el *Teorema 4* siguiendo los mismos pasos que en [22].

Antes de comenzar la prueba necesitamos un poco de artillería pesada, de la cual algunas demostraciones se incluirán en el *Apéndice*.

La herramienta principal que usaremos es la *Fórmula de co-área* (3.2). Primero mostremos un caso particular de la misma y luego enunciaremos el caso general.

Sea $S^{n-1} \subset \mathbb{R}^n$ la esfera unitaria y $B^k \subset \mathbb{R}^k$ la bola unitaria, sea $\pi : S^{n-1} \rightarrow B^k$ la restricción de la proyección ortogonal de \mathbb{R}^n a \mathbb{R}^k , entonces para toda función $\varphi : S^{n-1} \rightarrow \mathbb{R}$ integrable se tiene

$$\int_{S^{n-1}} \varphi(\Theta) d\Theta = \int_{x \in B^k} \int_{y \in \pi^{-1}(x)} \varphi(y) \frac{1}{(1 - \|x\|^2)^{1/2}} dy dx. \quad (3.1)$$

(ver Figura 3.1) La demostración la incluiremos en el *Apéndice*. Veamos el caso general.

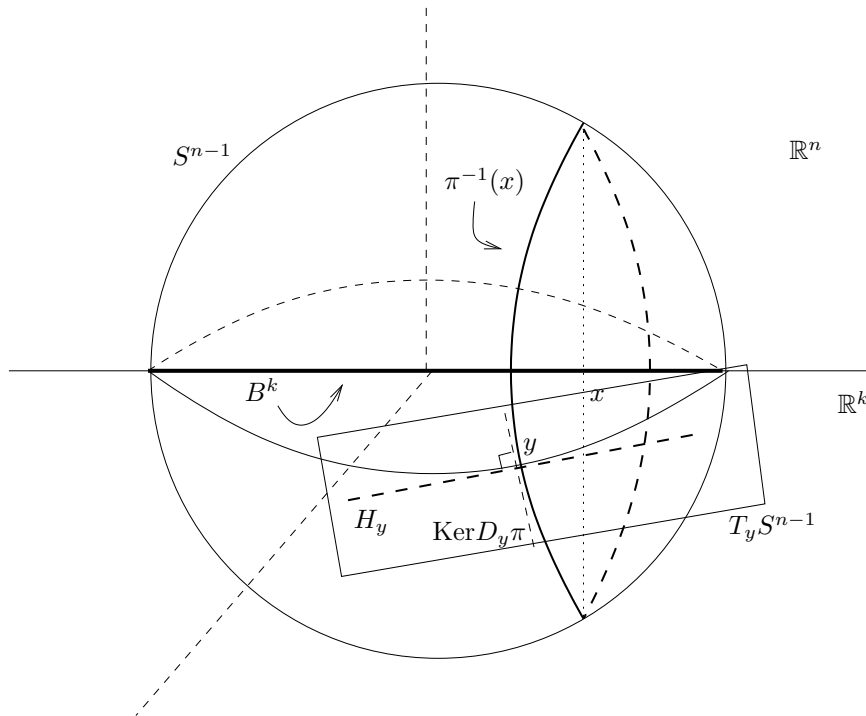


Figura 3.1: Caso particular de la fórmula de co-área

Sean X e Y variedades Riemannianas y $F : X \rightarrow Y$ un mapa sobreyectivo tal que para casi todo $x \in X$ (con la medida inducida por la fórmula de volúmen en X) existe el diferencial $D_x F$ y es sobreyectivo. Para los puntos donde existe $D_x F$ definimos el *espacio horizontal* H_x en x como el subespacio lineal de $T_x X$ complementario al núcleo de $D_x F$. También definimos para dichos puntos la *derivada normal de F en x* (o *jacobiano normal de F en x*) como

$$NJ(F(x)) := |\det(D_x F|_{H_x})|,$$

(ver Figura 3.1). Es decir, la derivada normal de F en x es el volúmen del paralelepípedo formado por los vectores $D_x F(v_1), \dots, D_x F(v_l)$ en $T_{F(x)} Y$ siendo $\{v_1, \dots, v_l\}$ una base ortonormal del espacio horizontal H_x .

Teorema 5. Sea $F : X \rightarrow Y$ en las condiciones anteriores, entonces $\forall \varphi : X \rightarrow \mathbb{R}$ integrable se tiene

$$\int_X \varphi(x) dX = \int_Y \int_{x \in F^{-1}(y)} \varphi(x) \frac{1}{NJ(F(x))} dF^{-1}(y) dx. \quad (3.2)$$

Demostración. Ver Apéndice. □

La fórmula de co-área esta establecida en un contexto general de mapas entre variedades Riemannianas. Dicha fórmula es muy útil pues nos permite desglosar una integral en integrales iteradas (Fubini Geométrico). Para nuestro caso queremos usar la fórmula para computar $\int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f d\mathbb{P}(\mathcal{H}_{(d)})$.

La referencias básicas para este capítulo son [5] y [22].

3.2. Teorema de Shub-Smale

La prueba del *Teorema 4* se divide en varios pasos.

Consideremos la variedad solución $V \subset \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{R}^{n+1})$. Como subvariedad Riemanniana del producto, V tiene una forma de volúmen asociada. Sea $\pi_1 : V \rightarrow \mathbb{P}(\mathcal{H}_{(d)})$ la proyección en la primer coordenada. Observar que casi todo punto de $\mathbb{P}(\mathcal{H}_{(d)})$ es regular por el teorema de Sard, y por lo tanto $\pi_1^{-1}(f)$ es una subvariedad diferenciable en V , que por ser V compacta y de igual dimensión que $\mathbb{P}(\mathcal{H}_{(d)})$ se tiene que $\#\pi_1^{-1}(f)$ es finito.

Observar que Σ' tiene medida nula en V (por ser una subvariedad algebraica de V), Además, $\Sigma \subset \mathbb{P}(\mathcal{H}_{(d)})$ también tiene medida nula. Veamos eso:

Consideremos la proyección $\pi_1 : V \rightarrow \mathbb{P}(\mathcal{H}_{(d)})$ en su primer coordenada. Como π_1 es un mapa C^∞ (en realidad tiene la misma clase de diferenciabilidad que la que tiene la variedad diferenciable V) se tiene por el *Teorema de Sard*, que el conjunto de valores regulares tiene medida total. Ahora como los valores regulares son exactamente el conjunto $\Sigma \subset \mathbb{P}(\mathcal{H}_{(d)})$, se concluye que Σ tiene medida cero.² Por lo tanto, al integrar en V o en $\mathbb{P}(\mathcal{H}_{(d)})$ podemos despreciarlos.³

En lo que sigue notaremos por $V_\zeta := \pi_2^{-1}(\zeta) \subset V$ para todo $\zeta \in \mathbb{P}(\mathbb{C}^{n+1})$.

Proposición 6. Se tiene la siguiente igualdad

$$\int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f d\mathbb{P}(\mathcal{H}_{(d)}) = \text{Vol}(\mathbb{P}(\mathbb{R}^{n+1})) \cdot \int_{V_{e_0}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} dV_{e_0} \quad (3.3)$$

donde en el miembro derecho de la igualdad en (3.3) estamos eligiendo un representante $f \in \mathbb{P}(\mathcal{H}_{(d)})$ de norma uno, y donde $Tf(e_0) : T_{e_0} \rightarrow \mathbb{R}^n$ esta dado por $Tf(e_0) = D_{e_0}f|_{T_{e_0}}$.

Antes de demostrar la proposición enunciaremos dos lemas que su demostración se incluirán en la próxima sección.

²En realidad, cuando usamos el Teorema de Sard, probamos que Σ tiene medida cero, para la medida inducida por las parametrizaciones (donde los conjuntos de medida nula están bien definidos). Solo resta ver que los conjuntos de medida cero son conjuntos de medida nula para la medida inducida por la métrica Riemanniana de $\mathbb{P}(\mathcal{H}_{(d)})$. Una manera de hacerlo es identificar $\mathcal{H}_{(d)}$ con $\mathbb{R}^{\dim \mathcal{H}_{(d)}}$ e inducir una métrica en el segundo. Luego basta ver que dicha métrica genera una forma de volúmen en $\mathbb{R}^{\dim \mathcal{H}_{(d)}}$ que es absolutamente continua respecto a la medida de Lebesgue, y eso es fácilmente verificable (ver sección *Nuevo Enfoque*).

³En el apéndice incluiremos una demostración auto-contenida del hecho que Σ tiene medida nula en $\mathbb{P}(\mathcal{H}_{(d)})$.

Recordar que Σ' es el conjunto de puntos críticos de $\pi_1 : V \rightarrow \mathbb{P}(\mathcal{H}_{(d)})$. Además vimos en la página 24 que si $(f, \zeta) \in V - \Sigma'$ entonces $D_\zeta f$ es sobreyectivo. Entonces considerando la función evaluación $\epsilon : \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{R}^{n+1}) \rightarrow \mathbb{R}^n$ y observando que la derivada de ϵ respecto a la segunda fibra $\partial_2 \epsilon$ resulta

$$\partial_2 \epsilon(f, \zeta) = D_\zeta f,$$

entonces podemos usar el teorema de la función implícita y encontrar para cada $(f_0, \zeta_0) \in V - \Sigma'$ una función $G = G_{(f_0, \zeta_0)} : U(f_0) \rightarrow \mathbb{P}(\mathbb{R}^{n+1})$ con $U(f_0) \subset \mathbb{P}(\mathcal{H}_{(d)})$ entorno de f_0 tal que:

$$\epsilon(f, G(f)) = 0 \quad \forall f \in U(f_0), \quad \text{y} \quad G(f_0) = \zeta_0.$$

Es fácil chequear que

$$D_\zeta f|_{T_\zeta} \left(D_f G(\dot{f}) \right) = \dot{f}(\zeta) \quad \forall \dot{f} \in T_f \mathbb{P}(\mathcal{H}_{(d)}). \quad (3.4)$$

Lema 4. Para todo par $(f, \zeta) \in V - \Sigma'$, $D_f G$ es sobreyectivo, y en ese caso vale

$$|\det(D_{(f, \zeta)} \pi_1)| \cdot \frac{1}{NJ(\pi_2(f, \zeta))} = \det(D_f G D_f G^*)^{-1/2}.$$

Lema 5. En general, si M, N, V son variedades Riemannianas con $V \subset M \times N$, y $\sigma_1, \sigma_2, \sigma_3$ son isometrías de V, N y M respectivamente, entonces:

$$(i) \text{ si } \pi_2 \sigma_1 = \sigma_2 \pi_2 \text{ entonces } NJ(\pi_2(\sigma_2(x))) = NJ(\pi_2(x)),$$

$$(ii) \text{ si } \pi_1 \sigma_1 = \sigma_3 \pi_1 \text{ entonces } |\det(D_{\sigma(x)} \pi_1)| = |\det(D_x \pi_1)|,$$

siendo π_1, π_2 las proyecciones de $V \subset M \times N$ en su primera y segunda fibra, respectivamente.

Observar que si O es una transformación ortogonal de \mathbb{R}^{n+1} en \mathbb{R}^{n+1} ($O \in \mathcal{O}(n+1)$) entonces ésta induce una transformación $\hat{O} : V \rightarrow V$ definida por

$$\hat{O}(f, \zeta) := (f \circ O^{-1}, O(\zeta)).$$

Dicha transformación deja invariante V ya que $f \circ O^{-1}(O(\zeta)) = f(\zeta)$. Además es una transformación biyectiva.

Podemos pensar a la transformación \hat{O} como una acción libre del grupo ortogonal $\mathcal{O}(n+1)$ sobre la variedad solución, y por lo tanto podemos decir que V es *ortogonalmente invariante* (es decir, es invariante bajo la acción del grupo ortogonal). Esta propiedad es crucial en la prueba de la Proposición 6.

Dem. de Proposición 6. Usando la fórmula de co-área (3.2) para $X = V, Y = \mathbb{P}(\mathcal{H}_{(d)})$, $F = \pi_1$ y $\varphi = |\det D\pi_1|$ resulta

$$\int_V |\det D\pi_1| dV = \int_{f \in \mathbb{P}(\mathcal{H}_{(d)})} \int_{\pi_1^{-1}(f)} |\det D\pi_1| \frac{1}{NJ(\pi_1)} d\pi_1^{-1}(f) d\mathbb{P}(\mathcal{H}_{(d)}).$$

Como estamos integrando en $V - \Sigma'$ se tiene que $D_{(f, \zeta)} \pi_1$ es sobreyectivo (por definición), y como V tiene igual dimensión que $\mathbb{P}(\mathcal{H}_{(d)})$ se tiene que $\text{Ker } D_{(f, \zeta)} \pi_1 = \{0\}$. Entonces resulta que $\forall (f, \zeta) \in V - \Sigma'$

$$NJ(\pi_1(f, \zeta)) = |\det D_{(f, \zeta)} \pi_1|$$

y por lo tanto

$$\int_V |\det D\pi_1| dV = \int_{f \in \mathbb{P}(\mathcal{H}_{(d)})} \int_{\pi_1^{-1}(f)} 1 d\pi_1^{-1}(f) d\mathbb{P}(\mathcal{H}_{(d)}), \quad (3.5)$$

donde

$$N_0^f = \int_{\pi_1^{-1}(f)} 1 d\pi_1^{-1}(f).$$

Usando nuevamente la fórmula de co-área pero para π_2 se tiene

$$\int_V |\det D\pi_1| dV = \int_{\zeta \in \mathbb{P}(\mathbb{R}^{n+1})} \int_{(f,\zeta) \in V_\zeta} |\det D_{(f,\zeta)}\pi_1| \frac{1}{NJ(\pi_2(f,\zeta))} dV_\zeta d\mathbb{P}(\mathbb{R}^{n+1}).$$

Sea $O : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{n+1}$ una transformación ortogonal, tal que $O(\zeta) = e_0$. Dicha transformación induce otra $\hat{O} : V_\zeta \rightarrow V_{e_0}$ definida de manera natural como $\hat{O}(f,\zeta) = (f \circ O^{-1}, e_0)$. Es fácil chequear que ahora \hat{O} es una biyección isométrica de V_ζ en V_{e_0} (ver página 13). Entonces aplicando el cambio de variable \hat{O} en V_ζ y usando el *Lema 5* con $V = V_\zeta$, $M = \mathbb{P}(\mathcal{H}_{(d)})$, $N = \mathbb{P}(\mathbb{R}^{n+1})$, $\sigma_1 = \hat{O}$, σ_2 y σ_3 las inducidas también por O en cada fibra, se tiene:

$$\int_{(f,\zeta) \in V_\zeta} |\det D_{(f,\zeta)}\pi_1| \frac{1}{NJ(\pi_2(f,\zeta))} dV_\zeta = \int_{(g,e_0) \in V_{e_0}} |\det D_{(g,e_0)}\pi_1| \frac{1}{NJ(\pi_2(g,e_0))} dV_{e_0} \quad (3.6)$$

Entonces se tiene

$$\int_V |\det D\pi_1| dV = \text{Vol}(\mathbb{P}(\mathbb{R}^{n+1})) \cdot \int_{V_{e_0}} |\det D\pi_1| \frac{1}{NJ(\pi_2)} dV_{e_0}$$

y usando el *Lema 4*

$$\int_V |\det D\pi_1| dV = \text{Vol}(\mathbb{P}(\mathbb{R}^{n+1})) \cdot \int_{V_{e_0}} \det(D_f G D_f G^*)^{-1/2} dV_{e_0}. \quad (3.7)$$

Resumiendo, de (3.5) y de (3.7) resulta

$$\int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f d\mathbb{P}(\mathcal{H}_{(d)}) = \text{Vol}(\mathbb{P}(\mathbb{R}^{n+1})) \cdot \int_{V_{e_0}} \det(D_f G D_f G^*)^{-1/2} dV_{e_0}. \quad (3.8)$$

Por lo tanto solo falta chequear que

$$\det(D_f G D_f G^*)^{-1} = \det(Tf(e_0) Tf(e_0)^*).$$

Tomemos $f \in \mathbb{P}(\mathcal{H}_{(d)})$ de norma uno como representante. Observar que $D_f G : T_f \rightarrow T_{e_0} \mathbb{R}^{n+1} \simeq \mathbb{R}^n$, $D_{e_0} f|_{T_{e_0}} : T_{e_0} \mathbb{R}^{n+1} \simeq \mathbb{R} \rightarrow \mathbb{R}^n$. Dichas identificaciones las podemos hacer pues

$$T_{e_0} = \{(0, u_1, \dots, u_n) \in \mathbb{R}^{n+1} : u_i \in \mathbb{R}, i = 1, \dots, n\} \simeq \mathbb{R}^n$$

Consideremos el subespacio lineal \hat{K}_{e_0} de $\mathcal{H}_{(d)}$ formado por los $f \in \mathcal{H}_{(d)}$ tal que su polinomio coordinado $f_i(x) = \sum a_j^{(i)} x^j$ es tal que $a_j^{(i)} = 0$ cuando $j \neq (d_i, 0, \dots, 0)$. Entonces se tiene que $\hat{K}_{e_0} \perp \hat{V}_{e_0} = \{f \in \mathcal{H}_{(d)} : f(e_0) = 0\}$ y $\mathcal{H}_{(d)} = \hat{K}_{e_0} \oplus \hat{V}_{e_0}$.

Como $\epsilon(h, G(h)) = 0$ para todo h en un entorno de f , se tiene

$$D_{e_0}f \circ D_fG(\dot{f}) = -\dot{f}(e_0), \quad \forall \dot{f} \in T_f.$$

Observando que $D_fG(\dot{f}) = -(D_{e_0}f|_{T_{e_0}})^{-1} \dot{f}(e_0)$, entonces $D_fG(\dot{f}) = 0$ para todo $\dot{f} \in \hat{V}_{e_0}$ y por lo tanto podemos identificar D_fG con $D_fG|_{\hat{K}_{e_0}}$. Si tomamos $\dot{f} \in \hat{K}_{e_0}$ entonces podemos escribir

$$\dot{f} = \underbrace{(\dot{a}_{(d_1, 0, \dots, 0)}^{(1)}, \dots, \dot{a}_{(d_n, 0, \dots, 0)}^{(n)}, 0, \dots, 0)}_{N = \dim \mathcal{H}_{(d)}}$$

tomando la base de los monomios y reordenando, entonces claramente podemos identificar \hat{K}_{e_0} con \mathbb{R}^n . Además

$$\dot{f}(e_0) = (\dot{a}_{(d_1, 0, \dots, 0)}^{(1)}, \dots, \dot{a}_{(d_n, 0, \dots, 0)}^{(n)}).$$

Entonces podemos identificar $\dot{f} \in \hat{K}_{e_0}$ con $\dot{f}(e_0) \in \mathbb{R}^n$. Por lo tanto podemos encontrar una base de \hat{K}_{e_0} tal que $D_{e_0}f \circ D_fG = Id$, y como $D_fG : T_f \rightarrow T_{e_0}\mathbb{R}^{n+1}$ se tiene que

$$D_{e_0}f|_{T_{e_0}} \circ D_fG = Id$$

de donde podemos identificar $D_{e_0}f|_{T_{e_0}}$ con D_fG^{-1} y viceversa concluyendo lo que queríamos. \square

Observación

Podemos obtener un resultado más general si nos restringimos a un subconjunto U de V , que tenga una propiedad de invarianza ortogonal. Dicha propiedad es la necesaria para que se verifique (3.6). Decimos que un subconjunto U de V es *ortogonalmente invariante* si $O(U_\zeta) = O(U)_{O(\zeta)}$ donde $U_\zeta := U \cap V_\zeta$. Es fácil ver que si U es ortogonalmente invariante podemos modificar la prueba anterior sin cambios significativos y obtener:

Proposición 7. *Si $U \subset V$ es ortogonalmente invariante, entonces*

$$\int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f(U) d\mathbb{P}(\mathcal{H}_{(d)}) = \text{Vol}(\mathbb{P}(\mathbb{R}^{n+1})) \cdot \int_{U \cap V_{e_0}} (\det(Tf(e_0)Tf(e_0)^*))^{1/2} dV_{e_0}$$

Demostración del Teorema 4. Definimos

$$A_{(d)} = \frac{1}{\text{Vol}(\mathbb{P}(\mathcal{H}_{(d)}))} \cdot \int_{\mathbb{P}(\mathcal{H}_{(d)})} N_0^f d\mathbb{P}(\mathcal{H}_{(d)}).$$

Usando la *Proposición 6* obtenemos

$$A_{(d)} = \frac{\text{Vol}(\mathbb{P}(\mathbb{R}^{n+1}))}{\text{Vol}(\mathbb{P}(\mathcal{H}_{(d)}))} \cdot \int_{V_{e_0}} (\det(Tf(e_0)Tf(e_0)^*))^{1/2} dV_{e_0}. \quad (3.9)$$

Sea S_{e_0} la esfera unitaria en \hat{V}_{e_0} . Como V_{e_0} lo podemos obtener cocientando S_{e_0} por el mapa antípoda (es decir identificamos x con $-x$) y además el mapa cociente $Q : S_{e_0} \rightarrow V_{e_0}$ es una

isometría (entre variedades Riemannianas) que cubre dos veces V_{e_0} (es decir $\#Q^{-1}(\bar{x}) = 2, \forall \bar{x} \in V_{e_0}$) entonces

$$\int_{V_{e_0}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} dV_{e_0} = \frac{1}{2} \underbrace{\int_{S_{e_0}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} dS_{e_0}}_{(I)}. \quad (3.10)$$

Sea $L_{(d)} \subset \hat{V}_{e_0}$ el subespacio lineal definido por los polinomios $f = (f_1, \dots, f_n) \in \hat{V}_{e_0}$ tal que $f_i(x) = x_0^{d_i-1} \cdot \sum_{k=1}^n a_k^{(i)} x_k$ para $i = 1, \dots, n$. Notemos por $\hat{\pi} : \hat{V}_{e_0} \rightarrow L_{(d)}$ a la proyección ortogonal (ver Figura 3.2).

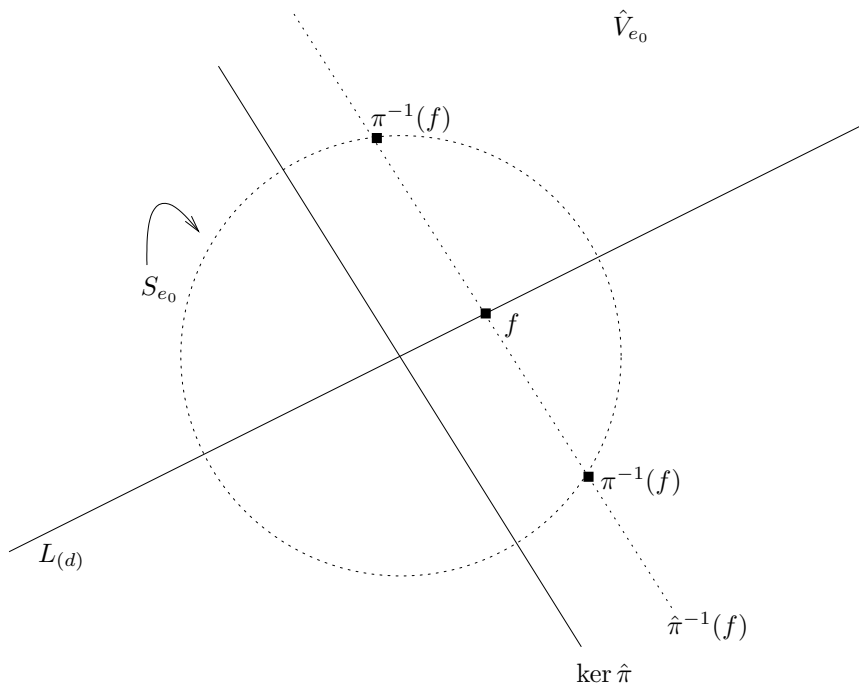


Figura 3.2: Subespacio $L_{(d)}$ en \hat{V}_{e_0} , y proyecciones $\hat{\pi} : \hat{V}_{e_0} \rightarrow L_{(d)}$ y $\pi = \hat{\pi}|_{S_{e_0}}$.

Observar que si $f \in \text{Ker } \hat{\pi}$ entonces f tiene que ser de la forma

$$f_i(x) = x_0^{d_i-2} \cdot \sum_{\|j\|=2} a_{(j_1, \dots, j_n)}^{(i)} x_1^{j_1} \cdots x_n^{j_n} + x_0^{d_i-3} \cdot \sum_{\|j\|=3} a_{(j_1, \dots, j_n)}^{(i)} x_1^{j_1} \cdots x_n^{j_n} + \dots,$$

y entonces $D_{e_0}g \equiv 0$ para todo $g \in \text{Ker } \hat{\pi}$. Por lo tanto para todo $f \in \hat{V}_{e_0}$

$$Tf(e_0) = T[\hat{\pi}(f) + (f - \hat{\pi}(f))](e_0) = T(\hat{\pi}(f))(e_0)$$

es decir,

$$Tf(e_0) = T(\hat{\pi}(f))(e_0). \quad (3.11)$$

Si notamos ahora por $\pi : S_{e_0} \rightarrow L_{(d)}$ la restricción de $\hat{\pi}$ a S_{e_0} o sea $\pi = \hat{\pi}|_{S_{e_0}}$ (ver Figura 3.2),

usando la fórmula de co-área particular (3.1) se tiene

$$\begin{aligned}
(I) &= \int_{S_{e_0}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} dS_{e_0} \\
&= \int_{\substack{f \in L_{(d)} \\ \|f\| \leq 1}} \int_{\pi^{-1}(f)} \overbrace{(\det(Tf(e_0) Tf(e_0)^*))^{1/2}}^{\text{cte en } \pi^{-1}(f)} \frac{1}{(1 - \|f\|^2)^{1/2}} d\pi^{-1}(f) dL_{(d)} \\
&= \int_{\substack{f \in L_{(d)} \\ \|f\| \leq 1}} \text{Vol}(S^l_{\sqrt{1-\|f\|^2}}) (\det(Tf(e_0) Tf(e_0)^*))^{1/2} \frac{1}{(1 - \|f\|^2)^{1/2}} dL_{(d)}
\end{aligned}$$

donde $l = \dim \pi^{-1}(f)$.

$l = \dim S_{e_0} - \dim L_{(d)}$ donde $\dim L_{(d)} = n^2$ y $\dim S_{e_0} = \dim \hat{V}_{e_0} - 1 = \dim \mathcal{H}_{(d)} - n - 1$. Es decir $l = N - n^2 - n - 1$, donde notamos por $N = \dim \mathcal{H}_{(d)}$.

Además como $\text{Vol}(S_r^k) = r^k \text{Vol}(S^k)$ se tiene

$$(I) = \text{Vol}(S^{N-n^2-n-1}) \cdot \underbrace{\int_{\substack{f \in L_{(d)} \\ \|f\| \leq 1}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} (1 - \|f\|^2)^{(l-1)/2} dL_{(d)}}_{(II)}.$$

Sea el mapa $A : L_{(d)} \rightarrow \mathcal{M}_n^{\mathbb{R}}$ (donde $\mathcal{M}_n^{\mathbb{R}}$ es el espacio de las matrices $n \times n$ con coeficientes reales) definido por

$$A(f) = \Delta(d_i^{-1/2}) \cdot [Tf(e_0)]$$

donde notamos por $\Delta(d_i^{-1/2})$ la matriz diagonal $D(d_1^{-1/2}, \dots, d_n^{-1/2})$ y $[\cdot]$ indica la matriz asociada en la base $\{e_1, \dots, e_n\}$ de T_{e_0} . Consideremos la norma de Frobenius $\|\cdot\|_F$ en $\mathcal{M}_n^{\mathbb{R}}$. Es fácil ver que para todo $f \in L_{(d)}$ se tiene que

$$[D_{e_0} f|_{T_{e_0}}] = \left(\left(a_{(d_i-1, e_k)}^{(i)} \right)_{\substack{i=1, \dots, n \\ k=1, \dots, n}} \right)$$

donde notamos por e_k para $k = 1, \dots, n$ las coordenadas de la base canónica en \mathbb{R}^n . Entonces

$$\|Tf(e_0)\|_F = \sum_{i=1}^n \left(\sum_{k=1}^n (a_{(d_i-1, e_k)}^{(i)})^2 \right).$$

Por otro lado

$$\begin{aligned}
\|f\|^2 &= \sum_{i=1}^n \left(\sum_{k=1}^n \frac{(a_{(d_i-1, e_k)}^{(i)})^2}{(d_i-1, e_k)} \right) \\
&= \sum_{i=1}^n \frac{1}{d_i} \left(\sum_{k=1}^n (a_{(d_i-1, e_k)}^{(i)})^2 \right),
\end{aligned}$$

entonces,

$$\begin{aligned} \|f\|^2 &= \sum_{i=1}^n \left(\sum_{k=1}^n \left[\frac{a_{(d_i-1, e_k)}^{(i)}}{\sqrt{d_i}} \right]^2 \right) \\ &= \|\Delta(d_i^{-1/2}) \cdot Tf(e_0)\|_f = \|A(f)\|_F. \end{aligned}$$

Con esto concluimos que el mapa $A : L_{(d)} \rightarrow \mathcal{M}_n^{\mathbb{R}}$ es una isometría lineal. Haciendo el cambio de variable lineal dado por A resulta

$$\begin{aligned} (II) &:= \int_{\substack{f \in L_{(d)} \\ \|f\| \leq 1}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} (1 - \|f\|^2)^{(l-1)/2} dL_{(d)} \\ &= \int_{\substack{M \in \mathcal{M}_n^{\mathbb{R}} \\ \|M\| \leq 1}} \left(\det \left(\Delta(d_i^{1/2}) M \left(\Delta(d_i^{1/2}) M \right)^* \right) \right)^{1/2} (1 - \|M\|^2)^{(l-1)/2} d\mathcal{M}_n^{\mathbb{R}} \\ &= \int_{\substack{M \in \mathcal{M}_n^{\mathbb{R}} \\ \|M\| \leq 1}} \det(\Delta(d_i))^{1/2} \det(M M^*)^{1/2} (1 - \|M\|^2)^{(l-1)/2} d\mathcal{M}_n^{\mathbb{R}}. \end{aligned}$$

Haciendo el cambio a coordenadas polares y observando que

$$\det(M M^*) = \|M\|^{2n} \cdot \det \left(\frac{M}{\|M\|} \left(\frac{M}{\|M\|} \right)^* \right)$$

se tiene

$$(II) = \mathcal{D}^{1/2} \int_0^1 \rho^{n^2-1} \rho^n (1 - \rho^2)^{(N-n^2-n-2)/2} \cdot \int_{\substack{M \in \mathcal{M}_n^{\mathbb{R}} \\ \|M\|=1}} \det(M M^*)^{1/2} d\mathcal{M}_n^{\mathbb{R}}.$$

Mediante un cálculo elemental se prueba que

$$\int_0^1 \rho^{n^2-1} \rho^n (1 - \rho^2)^{(N-n^2-n-2)/2} d\rho = \frac{1}{2} \frac{\Gamma \left(\frac{N-n^2-n}{2} \right) \cdot \Gamma \left(\frac{n^2+n}{2} \right)}{\Gamma \left(\frac{N}{2} \right)}$$

y llamando

$$h(n) = \int_{\substack{M \in \mathcal{M}_n^{\mathbb{R}} \\ \|M\|=1}} \det(M M^*)^{1/2} d\mathcal{M}_n^{\mathbb{R}}.$$

tenemos que:

$$\begin{aligned} (I) &= \mathcal{D}^{1/2} \cdot \text{Vol}(S^{N-n^2-n-1}) \cdot \frac{1}{2} \frac{\Gamma \left(\frac{N-n^2-n}{2} \right) \cdot \Gamma \left(\frac{n^2+n}{2} \right)}{\Gamma \left(\frac{N}{2} \right)} \cdot \Gamma \left(\frac{n^2+n}{2} \right) \cdot h(n) \\ &= \mathcal{D}^{1/2} \cdot \frac{\pi^{\frac{N-n^2-n}{2}}}{\Gamma \left(\frac{N}{2} \right)} \cdot h_1(n) = \mathcal{D}^{1/2} \cdot \frac{\pi^{\frac{N}{2}}}{\Gamma \left(\frac{N}{2} \right)} \cdot h_2(n) \end{aligned}$$

o sea

$$\int_{V_{e_0}} (\det(Tf(e_0) Tf(e_0)^*))^{1/2} dV_{e_0} = \frac{1}{2} \cdot \mathcal{D}^{1/2} \cdot \frac{\pi^{\frac{N}{2}}}{\Gamma \left(\frac{N}{2} \right)} \cdot h_2(n) \quad (3.12)$$

Resumiendo de (3.9) y de (3.12) obtenemos

$$\begin{aligned} A_{(d)} &= \frac{1}{2} \cdot \frac{\text{Vol}(\mathbb{P}(\mathbb{R}^{n+1}))}{\text{Vol}(\mathbb{P}(\mathcal{H}_{(d)}))} \mathcal{D}^{1/2} \cdot \frac{\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)} \cdot h_2(n) \\ &= \frac{\text{Vol}(\mathbb{P}(\mathbb{R}^{n+1}))}{\text{Vol}(S^N)} \mathcal{D}^{1/2} \cdot \frac{\pi^{\frac{N}{2}}}{\Gamma\left(\frac{N}{2}\right)} \cdot h_2(n) \\ &= \mathcal{D}^{1/2} \cdot H(n). \end{aligned}$$

En vez de calcular $H(n)$ que es una función que solo depende de n , lo computamos en un caso particular en el que conozcamos de antemano $A_{(d)}$ y \mathcal{D} . Basta considerar el caso más conocido de sistemas lineales en el cual consideramos $(d) = (d_1, \dots, d_n)$ que resulta fácilmente que $A_{(d)} = 1 = \mathcal{D}^{1/2}$, y entonces $H(n) = 1$ obteniéndose el teorema. \square

3.2.1. Complementos Técnicos

En esta sección demostraremos el *Lema 4* y el *Lema 5*.

Dem. de Lema 4. Sean H_1, H_2 espacios vectoriales de dimensión finita con producto interno, y $A : H_1 \rightarrow H_2$ una transformación lineal, donde el gráfico (lineal) lo notamos por $\Gamma(A)$, es decir, $\Gamma(A) = \{(x, Ax); x \in H_1\}$ (ver Figura 3.3). Sean $\pi_i : \Gamma(A) \rightarrow H_i$ $i = 1, 2$ las proyecciones

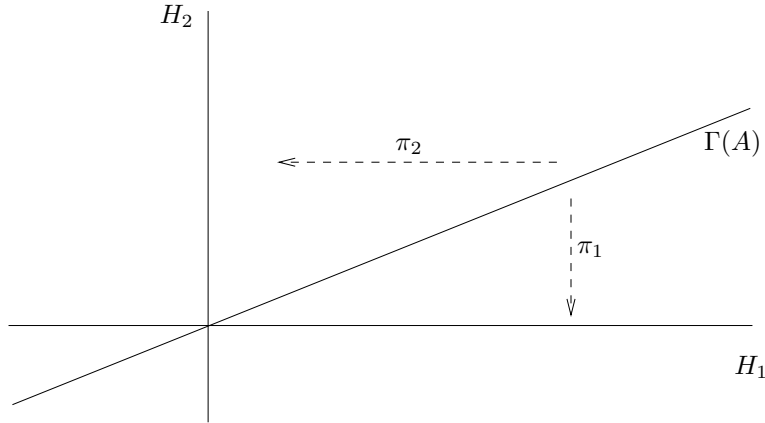


Figura 3.3: Gráfico $\Gamma(A)$ y proyecciones π_1, π_2 sobre H_1, H_2 respectivamente.

respectivas, donde estamos considerando $\Gamma(A)$ como subespacio lineal de $H_1 \times H_2$ con el producto interno heredado. Fijemos bases ortonormales en $\Gamma(A)$, H_1 y en H_2 (observar que $\dim \Gamma(A) = \dim H_1$).

Considerando $H_1 = T_f \mathbb{P}(\mathcal{H}_{(d)})$, $H_2 = T_x \mathbb{P}(\mathbb{R}^{n+1})$ y $A = D_f G$ el *Lema 4* sigue del siguiente lema. \square

Lema 6. Si A es sobreyectivo entonces

$$|\det \pi_1| \cdot \frac{1}{|\det(\pi_2|_{\text{Ker} \pi_2^\perp})|} = \det(A A^*)^{-1/2}.$$

Demostración. $\pi_1^{-1} : H_1 \rightarrow \Gamma(A)$ esta dado por $\pi_1^{-1}(x) = (x, Ax)$, es decir

$$\pi_1^{-1} = \begin{pmatrix} I \\ A \end{pmatrix}.$$

Sea $O : H_1 \rightarrow \Gamma(A)$ una transformación lineal ortogonal tal que:

$$O \left(\left(\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} \right)^{1/2} \right) = \begin{pmatrix} I \\ A \end{pmatrix}.$$

Tal transformación siempre existe pues:

$$\left\| \begin{pmatrix} I \\ A \end{pmatrix} x \right\|^2 = \left\langle \begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} x, x \right\rangle = \left\langle \left(\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} \right)^{1/2} x, \left(\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} \right)^{1/2} x \right\rangle$$

donde usamos que todo operador simétrico definido positivo en dimensión finita tiene una raíz cuadrada (basta diagonalizar y observar que los valores propios son positivos). Entonces

$$|\det \pi_1^{-1}| = \det \left[\left(\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} \right)^{1/2} \right] = \det \left[\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} \right]^{1/2}$$

Ahora

$$\begin{aligned} \left\langle \begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} x, y \right\rangle_{H_1} &= \langle (x, Ax), (y, Ay) \rangle_{\Gamma(A)} \\ &= \langle x, y \rangle_{H_1} + \langle Ax, Ay \rangle_{H_2} \\ &= \langle x, y \rangle_{H_1} + \langle A^* Ax, y \rangle_{H_2} \\ &= \langle (I + A^* A)x, y \rangle_{H_1}, \end{aligned}$$

entonces

$$\begin{pmatrix} I \\ A \end{pmatrix} \begin{pmatrix} I \\ A \end{pmatrix} = I + A^* A \Rightarrow |\det \pi_1| = \frac{1}{|\det(I + A^* A)|}. \quad (3.13)$$

Computemos ahora

$$|\det(\pi_2|_{\text{Ker } \pi_2^\perp})|.$$

Observar que $\text{Ker } \pi_2 = \{(x, 0) : x \in \text{Ker } A\}$. Sea

$$W := \{(A^*(AA^*)^{-1}y, y) : y \in H_2\}$$

donde $(AA^*)^{-1}$ tiene sentido pues como A es sobreyectivo, A^* es inyectivo y como $\text{Im } A^* = \text{Ker } A^\perp$ resulta que AA^* es invertible. Además $W \perp \text{Ker } \pi_2$ y $W \subset \Gamma(A)$ pues $A(A^*(AA^*)^{-1}y) = y$. Tenemos que $\dim W = \dim H_2 = \dim \text{Im } A$ entonces como $\dim \text{Ker } \pi_2 = \dim \text{Ker } A$ se tiene que

$$\text{Ker } \pi_2 \oplus W = \Gamma(A)$$

y por lo tanto $(\text{Ker } \pi_2)^\perp = W$. Entonces usando la fórmula (3.13) pero para $\pi_2|_W : W \rightarrow H_2$ se tiene

$$\begin{aligned} \frac{1}{|\det \pi_2|_W} &= \det \left[I + (A^*(AA^*)^{-1})^* (A^*(AA^*)^{-1}) \right] \\ &= \det(I + (AA^*)^{-1})^{1/2}. \end{aligned}$$

Ahora usando que $\det(I + AA^*) = \det(I + A^*A)$ (por tener AA^* y A^*A los mismos valores propios) se tiene:

$$\begin{aligned} |\det \pi_1| \cdot \frac{1}{|\det \pi_2| |W|} &= \det(I + (AA^*)^{-1})^{1/2} \cdot \det(I + (A^*A)^{-1})^{1/2} \\ &= \det(AA^*((AA^*)^{-1} + I))^{1/2} \cdot \det(I + (A^*A)^{-1})^{1/2} \\ &= \det(AA^*)^{-1/2}. \end{aligned}$$

□

Dem. de Lema 5. (i) Tenemos que $\pi_2 \circ \sigma_1 = \sigma_2 \circ \pi_2$. Entonces si $v \in \text{Ker } D_x \pi_2$ entonces $D_x \sigma_1(v) \in \text{Ker } D_{\sigma_1(x)} \pi_2$ pues:

$$\begin{aligned} D_{\sigma_1(x)} \pi_2(D_x \sigma_1(v)) &= D_x(\pi_2 \circ \sigma_1)(v) \\ &= D_x(\sigma_2 \circ \pi_2)(v) \\ &= D_{\pi_2(x)} \sigma_2 \circ D_x \pi_2(v) = 0. \end{aligned}$$

(Ver Figura 3.4)

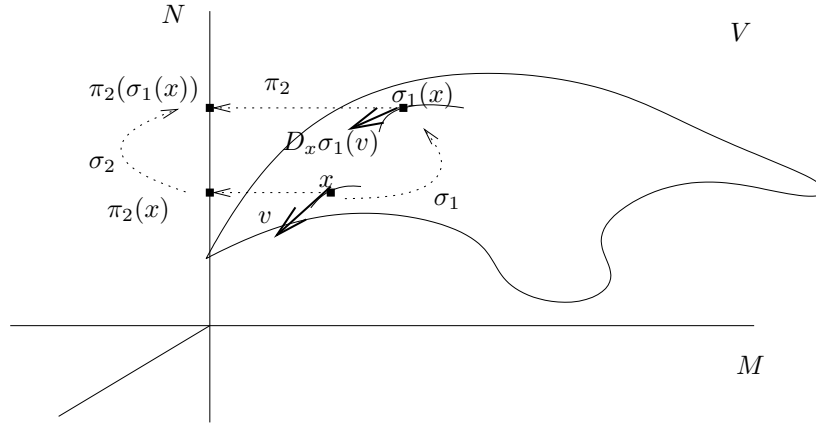


Figura 3.4: $\pi_2 \circ \sigma_1 = \sigma_2 \circ \pi_2$, $v \in \text{Ker } D_x \pi_2$.

Análogamente para la inversa de σ_1 , entonces obtenemos

$$D_x \sigma_1(\text{Ker } D_x \pi_2) = \text{Ker } D_{\sigma_1(x)} \pi_2.$$

Además como σ_1 es una isometría se tiene que

$$D_x \sigma_1(H_x) = H_{\sigma_1(x)}.$$

Entonces :

$$\begin{aligned} D_x \pi_2|_{H_x} &= D_x(\sigma_2^{-1} \circ \pi_2 \circ \sigma_1)|_{H_x} \\ &= D_{\pi_2(x)} \sigma_2^{-1} \circ D_{\sigma_1(x)} \pi_2|_{H_{\sigma_1(x)}} \circ D_x \sigma_1|_{H_x} \end{aligned}$$

y entonces

$$\begin{aligned}
 NJ(\pi_2(\sigma_1(x))) &= |\det D_{\sigma_1(x)}|_{H_{\sigma_1(x)}}| \\
 &= \left| \det \left(D_{\pi_2(x)} \sigma_2^{-1} \circ D_{\sigma_1(x)} \pi_2 |_{H_{\sigma_1(x)}} \circ D_x \sigma_1 |_{H_x} \right) \right| \\
 &= NJ(\pi_2(x))
 \end{aligned}$$

por ser $D\sigma_2^{-1}$ y $D\sigma_1$ isometrías.

(ii) Tenemos que $\pi_1 \circ \sigma_1 = \sigma_3 \circ \pi_1$.

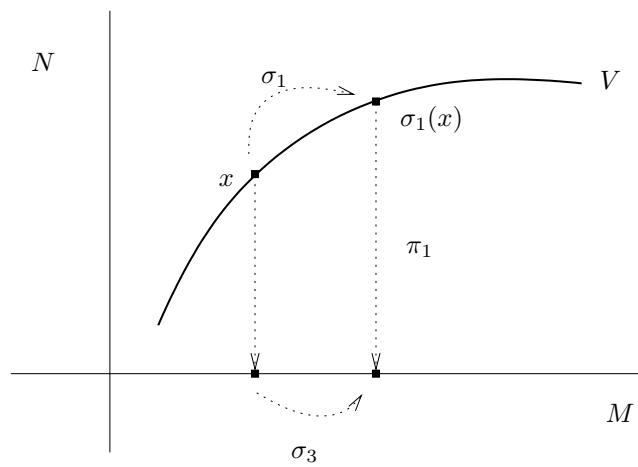


Figura 3.5: $\pi_1 \circ \sigma_1 = \sigma_3 \circ \pi_1$.

Entonces $\pi_1 = \sigma_3^{-1} \circ \pi_1 \circ \sigma_1$, y

$$\begin{aligned}
 |\det D_x \pi_1| &= |\det(D_{\pi_1(x)} \sigma_3^{-1} \circ D_{\sigma_1(x)} \pi_1 \circ D_x \sigma_1)| \\
 &= |\det D_{\sigma_1(x)} \pi_1|
 \end{aligned}$$

nuevamente por ser $D\sigma_3^{-1}$ y $D\sigma_1$ isometrías.

□

Capítulo 4

Nuevo Enfoque

4.1. Introducción

En este capítulo, vamos a desarrollar el segundo enfoque para estudiar el problema del número de soluciones de un sistema polinomial, y estableceremos la conexión entre los dos puntos de vista. Además de ayudar a la comprensión del problema facilitando el pasaje de un punto de vista al otro, el estudio de la conexión de los dos enfoques nos permitirá dar una nueva y sencilla demostración de la invarianza del producto interno de Weyl. También nos ayudará a mostrar que en el caso real existen otros productos internos invariantes en el espacio de polinomios (resultado, por cierto interesante, ya que en el caso complejo el producto interno de Weyl es único a menos de constantes multiplicativas.).

Básicamente en el capítulo anterior se construyó una medida en $\mathbb{P}(\mathcal{H}_d)$ invariante bajo la acción del grupo ortogonal $\mathcal{O}(n+1)$. Nuestro objetivo ahora es darle un enfoque probabilístico al problema donde ahora los coeficientes son tomados al azar de forma tal que su distribución sea invariante bajo la acción del grupo ortogonal $\mathcal{O}(n+1)$. Basta considerar una distribución de los coeficientes (o equivalentemente, del sistema polinomial en \mathcal{H}_d) que al restringirlos a la esfera unidad de \mathcal{H}_d (con la norma (2.10)) sea invariante bajo la acción del grupo ortogonal.

Es fácil ver que basta con encontrar tal distribución en los coeficientes para el caso de \mathcal{H}_d (el espacio de polinomios homogéneos en $n+1$ variables de grado d). Para eso consideremos la identificación natural de \mathcal{H}_d con $\mathbb{R}^{\dim \mathcal{H}_d}$:

-dado $f : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$,

$$f(x_0, \dots, x_n) = \sum_{\|j\|=d} a_j x_0^{j_0} \cdots x_n^{j_n}$$

consideramos la identificación:

$$f \rightsquigarrow ((a_j))_{\|j\|=d} \in \mathbb{R}^{\dim \mathcal{H}_d}.$$

De esta manera existe un producto interno $\langle \cdot, \cdot \rangle_{\mathcal{H}_d}$ en $\mathbb{R}^{\dim \mathcal{H}_d}$ dado por

$$\langle f, g \rangle_{\mathcal{H}_d} = \sum_{\|j\|=d} a_j b_j \binom{d}{j}^{-1}$$

donde $f = ((a_j))_{\|j\|=d}$ y $g = ((b_j))_{\|j\|=d}$. Con tal identificación, el grupo ortogonal $\mathcal{O}(n+1)$ actúa sobre $\mathbb{R}^{\dim \mathcal{H}_d}$ de la siguiente manera:

-dado $O \in \mathcal{O}(n+1)$ se tiene

$$O \cdot ((a_j))_{\|j\|=d} = ((b_j))_{\|j\|=d}$$

donde

$$\sum_{\|j\|=d} b_j x^j = f \circ O(x),$$

identificando $f \rightsquigarrow ((a_j))_{\|j\|=d}$.

Sabemos que si η es un vector aleatorio en \mathbb{R}^N con distribución Gaussiana estándar entonces $\eta/\|\eta\|$ tiene distribución en S^{N-1} invariante bajo el grupo ortogonal $\mathcal{O}(N)$ (considerando la norma euclídeana). Para eso basta ver que la medida de probabilidad de η viene dada por

$$\frac{1}{(\sqrt{2\pi})^N} \cdot e^{-\|x\|^2/2} dx_1 \dots dx_N.$$

De esta manera, resulta natural considerar en $\mathbb{R}^{\dim \mathcal{H}_d}$ con la norma $\|\cdot\|_{\mathcal{H}_d}$ la medida de probabilidad

$$\frac{1}{(\sqrt{2\pi})^{\dim \mathcal{H}_d}} \cdot e^{-\frac{\|f\|_{\mathcal{H}_d}^2}{2}} \prod_{\|j\|=d} \binom{d}{j}^{-1/2} da_j \quad (4.1)$$

donde estamos identificando $f(x) = \sum_{\|j\|=d} a_j x^j \in \mathcal{H}_d$ con el vector $((a_j))_{\|j\|=d}$ en $\mathbb{R}^{\dim \mathcal{H}_d}$. Para convencerse, basta considerar η con distribución Gaussiana estándar en $\mathbb{R}^{\dim \mathcal{H}_d}$ con la norma Euclídeana y luego transformarlo mediante el operador lineal $T : \mathbb{R}^{\dim \mathcal{H}_d} \rightarrow \mathbb{R}^{\dim \mathcal{H}_d}$ definido por

$$T((a_j))_{\|j\|=d} := \left(\binom{d}{j}^{1/2} a_j \right)_{\|j\|=d}.$$

El operador T es una isometría, considerando la norma Euclídeana en el dominio y la norma $\|\cdot\|_{\mathcal{H}_d}$ en el codominio. Es fácil ver que $T(\eta)$ verifica la propiedad deseada¹.

Notar que (4.1) se puede escribir como:

$$\prod_{\|j\|=d} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{a_j^2}{2\binom{d}{j}}} \cdot \binom{d}{j}^{-1/2} da_j$$

y por lo tanto, resulta que la distribución encontrada en los coeficientes del sistema, es la de considerar que los coeficientes a_j sean independientes entre sí con distribución Gaussiana centrada y con varianza $\binom{d}{j}$ para $\|j\|=d$.

En general, identificando $\mathcal{H}_{(d)}$ con $\mathbb{R}^{\dim \mathcal{H}_{(d)}}$ de manera análoga a la anterior, se tiene que basta considerar la medida de probabilidad dada por

$$\frac{1}{(\sqrt{2\pi})^{\dim \mathcal{H}_{(d)}}} \cdot e^{-\|f\|_{\mathcal{H}_{(d)}}^2/2} \prod_{i=1, \dots, n} \prod_{\|j\|=d_i} \binom{d_i}{j}^{-1/2} da_j^{(i)}$$

donde $f = (f_1, \dots, f_n) \in \mathcal{H}_{(d)}$,

$$f_i(x) = \sum_{\|j\|=d_i} a_j^{(i)} x^j \quad (i = 1, \dots, n).$$

¹Observar que si g es la densidad de η entonces $|\det T|^{-1}g \circ T$ es la densidad de $T(\eta)$.

Observación

Existe un isomorfismo lineal $\Phi : \mathcal{H}_{(d)} \rightarrow \mathcal{P}_{(d)}$ definido de manera natural como $\Phi(f, \dots, f_n) = (\Phi_1(f_1), \dots, \Phi_n(f_n))$ donde $\Phi_i := \Lambda^{-1} : \mathcal{H}_{d_i} \rightarrow \mathcal{P}_{d_i}$ definido en (2.4) para el caso $d = d_i$. Mediante este isomorfismo inducimos un producto interno en $\mathcal{P}_{(d)}$ invariante por el grupo ortogonal $\mathcal{O}(n)$. De esta manera el “nuevo enfoque” para el caso $\mathcal{P}_{(d)}$ resulta que la medida de probabilidad en los coeficientes esta dada por

$$\frac{1}{(\sqrt{2\pi})^{\dim \mathcal{P}_{(d)}}} \cdot e^{-\|f\|_{\mathcal{P}_{(d)}}^2/2} \prod_{i=1, \dots, n} \prod_{\|j\| \leq d_i} \binom{d_i}{j}^{-1/2} da_j^{(i)}$$

donde $f = (f_1, \dots, f_n) \in \mathcal{H}_{(d)}$,

$$f_i(x_1, \dots, x_n) = \sum_{\|j\| \leq d_i} a_j^{(i)} x_1^{j_1} \cdots x_n^{j_n} \quad (i = 1, \dots, n)$$

y donde ahora j es el multi-índice $j := (j_1, \dots, j_n)$, y

$$\binom{d_i}{j} = \frac{d_i!}{j_1! \cdots j_n! (d_i - \|j\|)!}.$$

Este nuevo enfoque nos permite estudiar el problema del número de soluciones de sistemas polinomiales en $\mathcal{P}_{(d)}$ desde otro punto de vista.

4.2. Formalización del Nuevo Enfoque y Extensiones

Sea \mathcal{H}_d el espacio de polinomios homogéneos definidos en \mathbb{R}^{n+1} . Notaremos por $\tilde{\mathcal{H}}_d := \{f|_{S^n} : f \in \mathcal{H}_d\}$. Dado $f \in \tilde{\mathcal{H}}_d$, notamos $f(t) = \sum_{\|j\|=d} f_j t^j$. Sea el producto interno de Weyl en $\tilde{\mathcal{H}}_d$, es decir

$$\langle f, g \rangle_W = \sum_{\|j\|=d} f_j g_j \binom{d}{j}^{-1}.$$

Sea la función $K : S^n \times S^n \rightarrow \mathbb{R}$, dada por $K(s, t) := \langle s, t \rangle^d = (s_0 t_0 + \cdots + s_n t_n)^d$. Observar que K cumple la siguiente propiedad:

$$\langle K(\cdot, t), f \rangle_W = \sum_{\|j\|=d} \binom{d}{j} t^j f_j \binom{d}{j}^{-1} = f(t). \quad (4.2)$$

Esta es la propiedad de *núcleo reproductor*.

Sea $\mathcal{F}(S^n, \mathbb{R})$ el espacio de funciones de S^n a valores reales, y $L_W^2(S^n, \mathbb{R})$ al subespacio (de Hilbert) de $\mathcal{F}(S^n, \mathbb{R})$ que resulta de completar $\tilde{\mathcal{H}}_d$ con el producto interno de $\langle \cdot, \cdot \rangle_W$.

Observar que la clausura respecto a $\langle \cdot, \cdot \rangle_W$ del subespacio lineal generado por $\{K(\cdot, t) : t \in S^n\}$ es exactamente $\tilde{\mathcal{H}}_d$. En caso contrario existiría $f \in \tilde{\mathcal{H}}_d$ no nulo, tal que f es ortogonal a todo elemento de $\{K(\cdot, t) : t \in S^n\}$, y por lo tanto usando la propiedad reproductora (4.2) resultaría $f(t) = 0$ para todo $t \in S^n$.

En el otro sentido, sea $X : \Omega \times S^n \rightarrow \mathbb{R}$ un proceso estocástico Gaussiano centrado² en cierto espacio de probabilidad (Ω, \mathcal{A}, P) , donde la covarianza es dada por

$$E(X(s)X(t)) := K(s, t).$$

Sea $L_X^2(\Omega) \subset L^2(\Omega, \mathcal{A}, P)$ la clausura del subespacio lineal generado por $\{X(t) : t \in S^n\}$ con el producto interno natural en $L^2(\Omega, \mathcal{A}, P)$.

Sea $\Theta : L(\{K(\cdot, t) : t \in S^n\}) \rightarrow L_X^2(\Omega)$ la aplicación lineal dada por $\Theta(K(\cdot, t)) = X(t)$, donde $L(A)$ indica el subespacio lineal generado por A . Observando que

$$E(X(t)X(s)) = K(s, t) = \langle K(\cdot, t), K(\cdot, s) \rangle_W,$$

obtenemos, que podemos extender Θ a la W -clausura de $L(\{K(\cdot, t) : t \in S^n\})$, que no es otra cosa que $\tilde{\mathcal{H}}_d$. De esta manera construimos un isomorfismo isométrico

$$\Theta : \tilde{\mathcal{H}}_d \rightarrow \Theta(\tilde{\mathcal{H}}_d) \subset L_X^2(\Omega),$$

donde es fácil chequear que $\Theta(\tilde{\mathcal{H}}_d) = L_X^2(\Omega)$ como consecuencia de que los espacios en cuestión tienen dimensión finita.

Este isomorfismo muestra la conexión de los dos enfoques. Resta probar que el proceso Gaussiano X tiene la forma deseada. Para eso sea la familia $\{\varphi_j : \|j\| = d\} \subset \tilde{\mathcal{H}}_d$ donde

$$\varphi_j(t) = \binom{d}{j}^{1/2} t^j.$$

Es fácil chequear que dicha familia forma una base ortonormal de $(\tilde{\mathcal{H}}_d, \langle \cdot, \cdot \rangle_W)$. Entonces la familia $\{\xi_j : \|j\| = d\}$ donde $\xi_j = \Theta(\varphi_j)$ es una base ortonormal de $L_X^2(\Omega)$. Es decir, $\{\xi_j : \|j\| = d\}$ es una familia de variables aleatorias Gaussianas estándar independientes. Además, se tiene

$$X(t) = \sum_{\|j\|=d} E(\xi_j X(t)) \xi_j.$$

Pero usando nuevamente la isometría Θ , obtenemos

$$E(\xi_j X(t)) = \langle \varphi_j, K(\cdot, t) \rangle_W = \varphi_j(t)$$

y por consiguiente

$$X(t) = \sum_{\|j\|=d} \binom{d}{j}^{1/2} \xi_j t^j,$$

como era de esperar.

En la construcción del isomorfismo isométrico Θ , el dominio S^n no jugó ningún papel en especial, y lo mismo pasa para el espacio de polinomios homogéneos. De igual manera, podemos construir el isomorfismo para el caso de polinomios definidos en \mathbb{R}^{n+1} . Esto nos permitirá extender los resultados partiendo en el otro sentido. Es decir, partimos de un proceso polinomial Gaussiano indexado en \mathbb{R}^{n+1} , con cierta ley invariante bajo el grupo ortogonal de \mathbb{R}^{n+1} (es decir, $K(Us, Ut) =$

²En el apéndice, agregamos una sección sobre *Procesos Estocásticos*.

$K(s, t)$ para toda transformación ortogonal de \mathbb{R}^{n+1}) y llegamos a un producto interno en el espacio de polinomios, invariante bajo la acción del grupo ortogonal.³

Concretamente, sea $X : \mathbb{R}^m \rightarrow \mathbb{R}$ un proceso polinomial Gaussiano centrado no degenerado. Sea $K : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ la covarianza, es decir, $K(s, t) = E(X(s)X(t))$ para todo $s, t \in \mathbb{R}^m$. Por ser X polinomial, resulta que $K(s, t)$ es un polinomio. De esta manera podemos definir un producto interno $\langle \cdot, \cdot \rangle_K$ en el espacio \mathcal{P}_d de polinomios en m variables de grado d , siendo d el grado del proceso polinomial X . Basta considerar

$$\langle K(\cdot, t), K(\cdot, s) \rangle_K := K(s, t).$$

Luego lo extendemos a la K -clausura $\overline{L(\{K(\cdot, t) : t \in \mathbb{R}^m\})}^{\langle \cdot, \cdot \rangle_K}$. Ahora como, $L(\{K(\cdot, t) : t \in \mathbb{R}^m\})$ es un subespacio lineal del espacio de dimensión finita \mathcal{P}_d , entonces usando que todas las normas son equivalentes, obtenemos:

$$\overline{L(\{K(\cdot, t) : t \in \mathbb{R}^m\})}^{\langle \cdot, \cdot \rangle_K} = \overline{L(\{K(\cdot, t) : t \in \mathbb{R}^m\})}^{\langle \cdot, \cdot \rangle_W} = \mathcal{P}_d,$$

siendo $\langle \cdot, \cdot \rangle_W$ el producto interno de Weyl en \mathcal{P}_d .

Veamos como la propiedad del núcleo reproductor nos permite dar una nueva prueba de la invarianza bajo la acción del grupo ortogonal (o unitario en el caso complejo) del producto interno de Weyl, y de su unicidad en el caso complejo.

Teorema 6. *La producto interno de Weyl en $\tilde{\mathcal{H}}_d$, es invariante por transformaciones ortogonales de \mathbb{R}^{n+1} .*

Demostración. Queremos probar que para toda isometría U de \mathbb{R}^{n+1} , se tiene $\langle Uf, Ug \rangle_W = \langle f, g \rangle_W$ para todo $f, g \in \tilde{\mathcal{H}}_d$ siendo $Uf(t) = f(Ut)$. Sea como antes, $K(s, t) = \langle s, t \rangle^d$, para todo $s, t \in S^n$. Observar que $UK(\cdot, t) = K(\cdot, U^T t)$. Entonces usando la propiedad reproductora (4.2) de K obtenemos:

$$\begin{aligned} \langle UK(\cdot, s), UK(\cdot, t) \rangle_W &= \langle K(\cdot, U^T s), K(\cdot, U^T t) \rangle_W = \langle U^T s, U^T t \rangle \\ &= \langle s, t \rangle = \langle K(\cdot, s), K(\cdot, t) \rangle_W, \end{aligned}$$

es decir, $\langle UK(\cdot, s), UK(\cdot, t) \rangle_W = \langle K(\cdot, s), K(\cdot, t) \rangle_W$. Luego $\langle \cdot, \cdot \rangle_W$ sigue siendo invariante en la W -clausura de $L(\{K(\cdot, t) : t \in S^n\})$. Y como $\overline{L(\{K(\cdot, t) : t \in S^n\})}^W = \tilde{\mathcal{H}}_d$, se obtiene la invarianza en $\tilde{\mathcal{H}}_d$. \square

Observar que la función $\Gamma_W : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ dada por $\Gamma_W(s, t) = \langle K(\cdot, t), K(\cdot, t) \rangle_W$ (que es una función polinomial de grado $2d$, homogénea en cada variable, e invariante por transformaciones ortogonales) determina completamente el producto interno de Weyl $\langle \cdot, \cdot \rangle_W$. Recíprocamente, sea $\Gamma : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ polinomial de grado $2d$, tal que, en cada variable es un polinomio homogéneo de grado d , invariante bajo transformaciones ortogonales ($\Gamma(U_s, U_t) = \Gamma(s, t)$). Se puede probar que

$$\Gamma(s, t) = Q(\langle s, t \rangle, \|s\|^2, \|t\|^2), \quad (4.3)$$

³Pedimos que el proceso Gaussiano sea polinomial para que $K(\cdot, t)$ sea un polinomio.

donde Q es un polinomio en tres variables reales y $Q(u, v, w) = Q(u, w, v)$ para cualquier elección de $u, v, w \in \mathbb{R}^n$ (ver Spivak [24]). De esta manera, en el caso que Γ sea definida positiva, podemos definir en $\mathcal{H}_d(\mathbb{R}^n)$ un producto interno $\langle \cdot, \cdot \rangle_\Gamma$ inducido por Γ , $\mathcal{O}(n)$ -invariante. Basta encontrar t_1, \dots, t_N elementos en \mathbb{R}^n (siendo $N = \dim \mathcal{H}_d(\mathbb{R}^n)$) tales que $\{K(\cdot, t_1), \dots, K(\cdot, t_n)\}$ sea una base de $\mathcal{H}_d(\mathbb{R}^n)$, y luego definir:

$$\langle K(\cdot, t_i), K(\cdot, t_j) \rangle_\Gamma := \Gamma(t_i, t_j), \quad (i, j = 1, \dots, N) \quad (4.4)$$

para luego extenderlo a todo $\mathcal{H}_d(\mathbb{R}^n)$. En particular, si Γ es una función de covarianza de algún proceso (polinomial), obtendremos un producto interno. En [17] se puede encontrar una caracterización completa de los polinomios Q que definen una covarianza en (4.3).

4.2.1. Ejemplo de nuevo producto interno $\mathcal{O}(n)$ -invariante

Mostraremos en el caso $\mathcal{H}_2(\mathbb{R}^2)$, un nuevo producto interno invariante bajo la acción de $\mathcal{O}(2)$. Un polinomio genérico en $\mathcal{H}_2(\mathbb{R}^2)$ lo notamos $f(x, y) = f_1x^2 + 2f_2xy + f_3y^2$, con $x, y \in \mathbb{R}^2$.

Es fácil chequear que los polinomios $E_1 := K(\cdot, (1, 0))$, $E_2 := K(\cdot, (1, 1))$, $E_3 := K(\cdot, (0, 1))$, forman una base de $\mathcal{H}_2(\mathbb{R}^2)$, y que la matriz de productos internos viene dada por:

$$((\langle E_i, E_j \rangle_\Gamma))_{\{i, j=1, 2, 3\}} = \begin{pmatrix} 2 & 3 & 1 \\ 3 & 8 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

Además, todo elemento de $\mathcal{H}_2(\mathbb{R}^2)$ se escribe⁴

$$f = (f_2 - f_1)E_1 + f_2E_2 + (f_3 - f_2)E_3.$$

Sea $\Gamma : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ dada por $\Gamma(s, t) = \langle s, t \rangle^2 + \|s\|^2\|t\|^2$. Es fácil chequear que Γ es definida positiva. Definiendo $\langle \cdot, \cdot \rangle_\Gamma$ como en (4.4) y extendiendo, resulta

$$\langle f, g \rangle_\Gamma := 2 \cdot (f_1g_1 + f_2g_2 + f_3g_3) + f_1g_3 + f_3g_1. \quad (4.5)$$

Es un ejercicio sencillo observar que (4.5) define un producto interno $\mathcal{O}(2)$ -invariante. Verifiquemos la invarianza:

Usando la identidad de polarización, basta probar que preserva normas. Sea la matriz

$$U = \begin{pmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{pmatrix}.$$

Entonces,

$$\begin{aligned} Uf(x, y) &= (f_1 \cos^2 \theta + 2f_2 \cos \theta \sin \theta + f_3 \sin^2 \theta) x^2 \\ &\quad + 2(f_1 \cos \theta + f_2(\sin^2 \theta - \cos^2 \theta) - f_3 \cos \theta \sin \theta) xy \\ &\quad + (f_1 \sin^2 \theta - 2f_2 \cos \theta \sin \theta + f_3 \cos^2 \theta) y^2 \end{aligned}$$

⁴Resulta $E_1(x, y) = x^2$, $E_2(x, y) = x^2 + 2xy + y^2$, $E_3(x, y) = y^2$.

de donde

$$\begin{aligned}
\|Uf\|_{\Gamma}^2/2 &= f_1^2 (\cos^4 \theta + 2 \cos^2 \theta \sin^2 \theta + \sin^4 \theta) + \\
&+ f_2^2 (4 \cos^2 \theta \sin^2 \theta + (\sin^2 \theta - \cos^2 \theta)^2) \\
&+ f_3^2 (\sin^4 \theta + 2 \cos^2 \theta \sin^2 \theta + \cos^4 \theta) \\
&+ f_1 f_2 (2 \cos^2 \theta \cos \theta \sin \theta + 2 \cos \theta \sin \theta (\sin^2 \theta - \cos^2 \theta) - 2 \sin^2 \theta \cos \theta \sin \theta) \\
&+ f_1 f_3 (2 \cos^2 \theta \sin^2 \theta + \cos^4 \theta + \sin^4 \theta) \\
&+ f_2 f_3 (2 \cos \theta \sin \theta \sin^2 \theta - 2(\sin^2 \theta - \cos^2 \theta) \cos \theta \sin \theta - 2 \cos \theta \sin \theta \cos^2 \theta) \\
&= f_1^2 + f_2^2 + f_3^2 + f_1 f_3 = \|f\|_{\Gamma}^2/2.
\end{aligned}$$

4.2.2. Unicidad del producto interno de Weyl en el caso complejo

Teorema 7. *El producto interno de Weyl en $\mathcal{H}_d(\mathbb{C}^n)$ es el único producto interno invariante bajo la acción del grupo unitario $\mathcal{U}(\mathbb{C}^n)$, a menos de constantes multiplicativas.*

Demostración. Sea $\langle \cdot, \cdot \rangle_N$ un producto interno en $\mathcal{H}_d(\mathbb{C}^n)$ invariante y $\Gamma : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ dada por $\Gamma(z, w) = \langle K(\cdot, w), K(\cdot, z) \rangle_N$, donde como antes $K(z, w) := \langle z, w \rangle^d = (z_1 \bar{w}_n + \dots + z_n \bar{w}_1)^d$. Entonces Γ es una función polinomial homogénea en la primer variable y polinomial conjugada homogénea en la segunda. Además, Γ es $\mathcal{U}(\mathbb{C}^n)$ -invariante. Basta probar que bajo estas hipótesis $\Gamma(z, w) = \alpha \langle z, w \rangle^d$ para algún α .

Fijado $0 \neq w \in \mathbb{C}^n$ sea $\varphi^{(w)} : \mathbb{C}^n \rightarrow \mathbb{C}$ la función polinomial dada por

$$\varphi^{(w)}(z) := \Gamma(z, w) - \frac{\Gamma(w, w)}{\|w\|^{2d}} \langle z, w \rangle^d.$$

Afirmación 1. *Para todo $U \in \mathcal{U}(\mathbb{C}^n)$ se tiene $\varphi^{(w)}(Uz) = \varphi^{(U^*w)}(z)$, siendo U^* la transformación adjunta.*

La afirmación resulta de

$$\begin{aligned}
\varphi^{(w)}(Uz) &= \Gamma(Uz, w) - \frac{\Gamma(w, w)}{\|w\|^{2d}} \langle Uz, w \rangle^d \\
&= \Gamma(z, U^*w) - \frac{\Gamma(w, w)}{\|w\|^{2d}} \langle z, U^*w \rangle^d \\
&= \varphi^{(U^*w)}(z),
\end{aligned}$$

donde usamos que U^* también es unitario (y por lo tanto $UU^*w = w$).

Sea el subgrupo $\mathcal{U}_w(\mathbb{C}^n)$ de $\mathcal{U}(\mathbb{C}^n)$ que deja invariante a w , y T_w el subespacio ortogonal a w . Entonces

$$\varphi^{(w)}(Uz) = \varphi^{(w)}(z) \quad \forall U \in \mathcal{U}_w(\mathbb{C}^n).$$

Dado $\lambda \in \mathbb{C}$, sea $f_\lambda := \varphi^{(w)}|_{\lambda w + T_w}$ la restricción de $\varphi^{(w)}$ al subespacio afín $\lambda w + T_w$. Identificando $\lambda w + T_w$ con \mathbb{C}^{n-1} , es fácil chequear que $f_\lambda : \mathbb{C}^{n-1} \rightarrow \mathbb{C}$ es una función polinomial invariante bajo la acción del grupo unitario $\mathcal{U}(\mathbb{C}^{n-1})$, es decir $f_\lambda(Uz) = f_\lambda(z)$ para todo $U \in \mathcal{U}(\mathbb{C}^{n-1})$: -cuando identificamos el subespacio afín con \mathbb{C}^{n-1} , toda transformación unitaria de este espacio esta identificada con una transformación unitaria en T_w , y toda transformación unitaria en T_w se extiende a una transformación en $\mathcal{U}_w(\mathbb{C}^n)$.

Resumiendo, $f_\lambda : \mathbb{C}^{n-1} \rightarrow \mathbb{C}$ es un polinomio en $n - 1$ variables complejas invariante bajo la acción del grupo unitario. Resulta entonces que f_λ debe ser constante, pues al restringir f_λ a cualquier subespacio uno-dimensional (complejo), la función (que resulta analítica) permanece constante en la circunferencia unidad de la recta. Concluimos entonces que para todo $\lambda \in \mathbb{C}$, $f_\lambda = f_\lambda(0) = \varphi^{(w)}(\lambda w) = 0$ de donde resulta $\varphi^{(w)}$ es el polinomio nulo. \square

4.3. Herramienta principal

En esta sección mostraremos la herramienta fundamental que permite atacar el problema desde el punto de vista de los campos aleatorios. Esta herramienta, es como ya comentamos en la introducción la *Fórmula de Rice*.

Daremos hipótesis suficientes para nuestro propósito, pero en el *Apéndice* estableceremos las hipótesis generales de dicha fórmula.

Teorema 8. *Sea $Z : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ un sistema polinomial aleatorio, con distribución Gaussiana. Entonces para todo $u \in \mathbb{R}^m$ se tiene:*

$$E(N_u^Z(V)) = \int_V E(|\det(Z'(t))|/Z(t) = u) \cdot p_{Z(t)}(u) dt, \quad (4.6)$$

siendo $V \subset \mathbb{R}^m$ de Borel.

Demostración. Usando argumentos estándares de extensión basta probar el resultado cuando $V = I$ un paralelepípedo en \mathbb{R}^m . Usando la fórmula del área se tiene que para todo $\omega \in \Omega$

$$\int_{\mathbb{R}^m} g(u) N_u^Z(I) du = \int_I g(Z(t)) |\det(Z'(t))| dt \quad (4.7)$$

siendo $g : \mathbb{R}^m \rightarrow \mathbb{R}$ medible, positiva y acotada (ver *Apéndice, Teorema 17*). Luego tomando esperanza en ambos miembros de (4.7), y aplicando el teorema de Fubini-Tonelli, obtenemos

$$\int_{\mathbb{R}^m} g(u) E(N_u^Z(I)) du = \int_I E(g(Z(t)) |\det(Z'(t))|) dt,$$

(observar que podemos usar Fubini-Tonelli por ser los integrandos funciones medibles positivas). Fijado t y condicionando respecto a $Z(t) = u$ resulta

$$E(g(Z(t)) |\det(Z'(t))|) = \int_{\mathbb{R}^m} g(u) E(|\det(Z'(t))|/Z(t) = u) p_{Z(t)}(u) du,$$

y entonces aplicando nuevamente Fubini-Tonelli se tiene

$$\int_{\mathbb{R}^m} g(u) E(N_u^Z(I)) du = \int_{\mathbb{R}^m} g(u) \left(\int_I E(|\det(Z'(t))|/Z(t) = u) p_{Z(t)}(u) dt \right) du.$$

Por lo tanto tenemos que (4.6) vale para casi todo $u \in \mathbb{R}^m$. Resta probar la continuidad de ambos miembros de (4.6), para concluir la prueba.

Fijemos $u_0 \in \mathbb{R}^m$ y $\omega \in \Omega$. Usando el *Lema 9* y *Lema 11* demostrados en el *Apéndice* podemos

asumir que u_0 es un valor regular y que sus pre-imágenes están en el interior de I . Por el teorema de la Función Inversa, existe $\delta(u_0)$ (que depende de ω) y entornos abiertos U_1, \dots, U_N (siendo $N = N_{u_0}^Z(I)$ que es finito por ser I compacto) tales que $Z^{-1}(B_{\delta(u_0)}(u_0)) = \cup_1^N U_i$, y es uno a uno en cada U_i de $1, \dots, N$. Resulta entonces que para todo $u \in B_{\delta(u_0)}(u_0)$ la cantidad de pre-imágenes es la misma, es decir $N_u^Z(I) = N_{u_0}^Z(I)$. Entonces para casi todo $\omega \in \Omega$, se tiene

$$\lim_{u \rightarrow u_0} N_u^Z(I) = N_{u_0}^Z(I).$$

La continuidad de $E(N_u^Z(I))$ respecto a u resulta de aplicar el teorema de Lebesgue de convergencia dominada, el cual puede aplicarse pues $N_u^Z(I)$ esta uniformemente acotado (salvo en un conjunto de probabilidad cero) por el número de Bézout del sistema polinomial en cuestión. Veamos la continuidad del segundo miembro. Haciendo una regresión Gaussiana de $Z'(t)$ respecto a $Z(t)$ obtenemos

$$Z'_j(t) = Y_j(t) + \alpha_j(t) Z(t)$$

donde $Z'_j(t)$ ($j = 1, \dots, m$) son las columnas de la matriz $Z'(t)$, $Y_j(t)$ para $j = 1, \dots, m$ son vectores Gaussianos independientes de $Z(t)$, y $\alpha_j(t)$ es una matriz continua en la variable t . La anterior construcción es posible usando el hecho de que $Z(t)$ no degenera en I ($\text{var}(Z(t)) \succ 0$). Entonces obtuvimos

$$E(|\det(Z'(t))|/Z(t) = u) = E(|\det[(Y_j(t) + \alpha_j(t)u)_{j=1\dots m}]|).$$

Como las variables en cuestión son Gaussianas y continuas respecto a las variables $u \in \mathbb{R}^m$ y $t \in I$ se tiene que el determinante esta uniformemente acotado por una constante, luego aplicando el teorema de convergencia dominada, se tiene

$$E(|\det[(Y_j(t) + \alpha_j(t)u)_{j=1\dots m}]|) \rightarrow E(|\det[(Y_j(t) + \alpha_j(t)u_0)_{j=1\dots m}]|)$$

cuando $u \rightarrow u_0$. Además la constante uniforme se puede tomar independiente de t y u , luego aplicando nuevamente convergencia dominada se obtiene el resultado. \square

Observaciones:

- En realidad no es ésta la versión que utilizaremos, pues $Z_i(t)$ será la normalización de $f_i(t)$ por su varianza. Pero todos los argumentos utilizados en esta prueba funcionan igual para este caso.
- Como la fórmula tiene un carácter local es posible extenderla al caso en que Z este definido en una variedad diferenciable encajada en algún espacio Euclídeo. En ese caso hay que remplazar la medida de Lebesgue por la medida geométrica de la variedad.
- La demostración anterior se basa en la demostración general de la fórmula de Rice. En este caso, ésta es más sencilla pues usamos fuertemente la hipótesis de que el número de pre-imágenes (a menos de un conjunto de probabilidad nula) esta uniformemente acotado (*Teorema de Bézout*). Esa diferencia, lleva a la simplificación de la prueba. En el caso general, se necesita poder controlar la cercanía entre las pre-imágenes y poder de esa manera aplicar algún teorema límite. La demostración del caso general se puede encontrar en [4].

Veamos como esta fórmula se aplica al estudio del número de soluciones de un sistema de ecuaciones aleatorias, que en principio no tiene por que ser polinomial.

Sea $f = (f_1, \dots, f_m)$ un sistema de ecuaciones donde cada $f_i(t)$ es independiente de las demás, con distribución Gaussiana no degenerada.

Definamos los nuevos procesos

$$Z_i(t) = \frac{f_i(t)}{\sqrt{\text{Var}(f_i(t))}} \quad (i = 1, \dots, m).$$

Es claro que el número de soluciones no se ve afectado, es decir, $N^f = N^Z$ donde $Z = (Z_1, \dots, Z_m)$. Como consecuencia de la normalización

$$E(Z_i^2(t)) = 1$$

y por lo tanto derivando respecto a las coordenadas t_l con $l = 1, \dots, m$ se obtiene que

$$E(Z_i(t) \frac{\partial}{\partial t_l} Z_i(t)) = 0$$

lo cual implica (por la gaussianidad de las variables) que $Z_i(t)$ y $\nabla Z_i(t)$ son independientes, y por lo tanto $Z(t)$ y $Z'(t)$ también lo son.

Además en el caso centrado se tiene que

$$p_{Z(t)}(0) = \frac{1}{(\sqrt{2\pi})^m}.$$

Entonces obtenemos la siguiente expresión para la cantidad de raíces:

$$E(N^Z) = \frac{1}{(\sqrt{2\pi})^m} \int_{\mathbb{R}^m} E(|\det(Z'(t))|) dt, \quad (4.8)$$

y por lo tanto el problema se reduce a computar

$$E(|\det(Z'(t))|).$$

4.4. Nuevo Teorema de Shub-Smale y Extensiones

En este nuevo enfoque el teorema de Shub-Smale tiene la siguiente forma:

Teorema 9 (Shub-Smale 1992). *Sea*

$$f_i(t) := \sum_{\|j\| \leq d_i} a_j^{(i)} t^j \quad (i = 1, \dots, m).$$

Supongamos que las variables aleatorias a valores reales $a_j^{(i)}$ son independientes, Gaussianas, centradas, y con varianzas:

$$\text{Var}(a_j^{(i)}) = \binom{d_i}{j_1 \dots j_m} := \frac{d_i!}{j_1! \dots j_m! (d_i - \|j\|)}.$$

Entonces

$$E(N^f) = \sqrt{\mathcal{D}}$$

donde $\mathcal{D} = d_1 \dots d_m$ es el número de Bézout del sistema polinomial f .

Daremos una demostración sencilla de este resultado para mostrar la utilidad de la fórmula de Rice, y más adelante se extenderá este resultado.

Demostración. Sea $\hat{f} = (\hat{f}_1, \dots, \hat{f}_m) : \mathbb{R}^{m+1} \rightarrow \mathbb{R}^m$ donde $\hat{f}_i : \mathbb{R}^{m+1} \rightarrow \mathbb{R}$ son las homogenizaciones de los polinomios f_i (ver página 15). Por lo tanto tenemos

$$N^f = \frac{1}{2} N^{\hat{f}}(S^m).$$

Es fácil ver que la covarianza de \hat{f}_i es:

$$E(\hat{f}_i(s)\hat{f}_i(t)) = \sum_{\|j\|=d_i} \binom{d_i}{j} s^j t^j = \langle s, t \rangle^{d_i}.$$

Definamos

$$Z_i(t) := \frac{\hat{f}_i(t)}{\|t\|} \quad (i = 1, \dots, m),$$

y $g := (g_1, \dots, g_m) : S^m \rightarrow \mathbb{R}^m$ donde

$$g_i := \hat{f}_i \Big|_{S^m}$$

Entonces, como $Z_i(t) = g_i(t)$ en S^m (es decir, Z_i es una extensión de g_i a todo \mathbb{R}^m) para todo $i = 1, \dots, m$, se concluye de la sección anterior que $g(t)$ y $g'(t)$ son independientes. Por lo tanto aplicando la fórmula (4.8) solo resta computar

$$E(|\det(g'(t))|) \tag{4.9}$$

en alguna base ortonormal de $T_t(S^m)$. Por la invarianza por isometrías basta con calcular (4.9) en $t = e_{m+1}$, siendo $\{e_1, \dots, e_{m+1}\}$ la base canónica de \mathbb{R}^{m+1} . Para este caso, es fácil calcular $g'(e_{m+1})$ y observar que en la base $\{e_1, \dots, e_m\}$ la matriz es de la forma $((d_i^{1/2} \xi_{ij}))_{i,j=1\dots m}$ donde ξ_{ij} son normales estándar independientes. Entonces

$$E(|\det(g'(t))|) = (d_1 \cdots d_m)^{1/2} \cdot E(|\det G|)$$

con $G = ((\xi_{ij}))_{i,j=1\dots m}$. Es un cálculo sencillo ver que

$$E(|\det G|) = \frac{1}{\sqrt{2\pi}} 2^{(m+1)/2} \cdot \Gamma\left(\frac{m+1}{2}\right),$$

si interpretamos el módulo del determinante como el volumen del paralelepípedo formado por las columnas de G .⁵

Entonces

$$\begin{aligned} E(N^f) &= \frac{1}{2} E(N^{\hat{f}}(S^m)) \\ &= \frac{1}{2} \int_{S^m} E(|\det G|) \frac{1}{(2\pi)^{m/2}} dS^m \\ &= \frac{1}{2} \cdot \text{Vol}(S^m) \cdot \frac{1}{(2\pi)^{m/2}} \cdot \sqrt{d_1 \cdots d_m} \cdot \frac{1}{\sqrt{2\pi}} 2^{(m+1)/2} \cdot \Gamma\left(\frac{m+1}{2}\right) \\ &= \frac{2^{(m-1)/2}}{(2\pi)^{(m+1)/2}} \cdot \sqrt{d_1 \cdots d_m} \cdot \Gamma\left(\frac{m+1}{2}\right) \cdot \text{Vol}(S^m) \\ &= \sqrt{d_1 \cdots d_m}. \end{aligned}$$

⁵En la sección *Smooth Analysis* calcularemos dicho determinante.

□

Más generalmente, sea $X : (X_1, \dots, X_m) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ un campo polinomial Gaussiano centrado, donde los procesos estocásticos a valores reales

$$X_1(t), \dots, X_m(t)$$

son independientes, y donde cada uno tiene una ley invariante por isometrías de \mathbb{R}^m . Esto significa que la función de covarianza

$$r^{X_i}(s, t) := E(X_i(s) X_i(t)) \quad (i = 1, \dots, m)$$

satisface

$$r^{X_i}(Us, Ut) = r^{X_i}(s, t)$$

para cualquier isometría $U : \mathbb{R}^m \rightarrow \mathbb{R}^m$.

Bajo esta propiedad de invarianza, es fácil ver que

$$r^{X_i}(s, t) = Q^{(i)}(\langle s, t \rangle, \|s\|^2, \|t\|^2) \quad (4.10)$$

donde $Q^{(i)}$ es una función de tres variables reales a valores. Aquí $\langle \cdot, \cdot \rangle$ indica el producto escalar usual en \mathbb{R}^m y $\|\cdot\|$ indica la norma Euclidea de \mathbb{R}^m .

Como ya mencionamos en la sección anterior, se puede probar que $Q^{(i)}$ es un polinomio (ver Spivak [24]) y además satisface la siguiente condición de simetría

$$Q^{(i)}(u, v, w) = Q^{(i)}(u, w, v) \quad \forall u, v, w \in \mathbb{R}. \quad (4.11)$$

Es interesante saber que clase de polinomios $Q^{(i)}$ son funciones de covarianza de algún proceso. Kostlan en [17] describe esta clase.

Para esta sección y la siguiente:

$Q_u^{(i)}$, $Q_v^{(i)}$, $Q_{uv}^{(i)}$, ... denotan las derivadas parciales de $Q^{(i)}$,

$$q_i(x) := \frac{Q_u^{(i)}}{Q^{(i)}}, \quad (4.12a)$$

$$r_i(x) := \frac{Q^{(i)} \left(Q_{uu}^{(i)} + 2Q_{uv}^{(i)} + 2Q_{uv}^{(i)} + 4Q_{vv}^{(i)} \right) - \left(Q_u^{(i)} + Q_v^{(i)} + Q_w^{(i)} \right)^2}{(Q^{(i)})^2}, \quad (4.12b)$$

$$h_i(x) := 1 + x \frac{r_i(x)}{q_i(x)} \quad (4.12c)$$

donde el miembro derecho de las igualdades (4.12a) y (4.12b) están evaluados en la terna (x, x, x) . Con estas notaciones podemos enunciar el siguiente teorema probado en [3]:

Teorema 10 (Azaïs-Wschebor 2005). *Para todo Boreliano $V \subset \mathbb{R}^m$ se tiene:*

$$E(N^X(V)) = \frac{1}{\sqrt{2} \pi^{\frac{m+1}{2}}} \Gamma\left(\frac{m}{2}\right) \int_V \left[\prod_{i=1}^m q_i(\|t\|^2) \right]^{1/2} \cdot E_h(\|t\|^2) dt \quad (4.13)$$

donde

$$E_h(x) = E \left(\left[\sum_{i=1}^m h_i(x) \xi_i^2 \right]^{1/2} \right)$$

y ξ_1, \dots, ξ_m denotan variables aleatorias independientes con distribución normal estándar.

Observación

Veamos como en algunos ejemplos la fórmula (4.13) se simplifica.

- Si $h_i(x)$ no depende de i , notando por $h(x)$ al valor común se tiene

$$E_h(x) = \sqrt{h(x)} E(\|\eta_m\|)$$

donde η_m es un vector Gaussiano estándar en \mathbb{R}^m . Es fácil chequear que

$$E(\|\eta_m\|) = \sqrt{2} \frac{\Gamma((m+1)/2)}{\Gamma(m/2)},$$

- (ver *Lema 8* más adelante). Si estamos bajo la hipótesis $Q^{(i)}(u, v, w) = Q^{d_i}(u)$ para cierto polinomio Q , entonces se tiene que

$$\begin{aligned} q_i(x) &= d_i \frac{Q'(u)}{Q(u)} = d_i q(x) \\ r_i(x) &= d_i \left[\frac{Q''(u)}{Q(u)} - \frac{Q'^2(u)}{Q(u)^2} \right] \\ h_i(x) &= h(x) = 1 - x \frac{Q'^2(x) - Q(x)Q''(x)}{Q(x)Q'(x)}. \end{aligned}$$

Usando el cambio a polares se obtiene que el número de soluciones en \mathbb{R}^m se expresa

$$\begin{aligned} E(N^X) &= \frac{1}{\sqrt{2}} \frac{1}{\pi^{\frac{m+1}{2}}} \Gamma\left(\frac{m}{2}\right) \sigma_{m-1} E(\|\eta_m\|) \sqrt{d_1 \cdots d_m} \int_0^{+\infty} \rho^{m-1} q(\rho^2)^{\frac{m}{2}} \sqrt{h(\rho^2)} d\rho \\ &= \sqrt{\frac{2}{\pi}} E(\|\eta_m\|) \sqrt{d_1 \cdots d_m} \int_0^{+\infty} \rho^{m-1} q(\rho^2)^{\frac{m}{2}} \sqrt{h(\rho^2)} d\rho, \end{aligned}$$

donde σ_{m-1} es la medida geométrica de la esfera unitaria S^{m-1} en \mathbb{R}^m , que resulta $\sigma_{m-1} = 2\pi^{m/2}/\Gamma(m/2)$.

Es fácil ver que en el caso Shub-Smale

$$Q^{(i)}(u, v, w) = (1+u)^{d_i} \quad (i = 1, \dots, m)$$

y

$$h(x) = q(x) = \frac{1}{1+x}.$$

Pasemos a demostrar el *Teorema 10*.

Demostración. Consideremos la normalización

$$Z_i(t) := \frac{X_i(t)}{(Q^{(i)}(\|t\|^2, \|t\|^2, \|t\|^2))^{1/2}} \quad (i = 1, \dots, m)$$

donde $Z := (Z_1, \dots, Z_m)^T$. De esta manera obtenemos (como ya vimos en (4.8))

$$E(N^Z(V)) = \frac{1}{(\sqrt{2\pi})^m} \int_V E(|\det(Z'(t))|) dt. \quad (4.14)$$

Mediante un cálculo se obtiene

$$\begin{aligned} E\left(\frac{\partial Z_i}{\partial t_\alpha}(t) \cdot \frac{\partial Z_i}{\partial t_\beta}(t)\right) &= \frac{\partial^2}{\partial s_\alpha \partial t_\beta} r^{Z_i}(s, t)|_{s=t} \\ &= r_i(\|t\|^2) t_\alpha t_\beta + q_i(\|t\|^2) \delta_{\alpha\beta}, \end{aligned}$$

entonces

$$\text{Var}(\nabla Z_i(t)) = q_i(\|t\|^2) I_m + r_i(\|t\|^2) t \cdot t^T,$$

donde I_l indica la matriz identidad de $l \times l$.

Sea U_t una isometría de \mathbb{R}^m tal que $U_t(t/\|t\|) = e_1$. Sea $\{t/\|t\|, v_2, \dots, v_m\}$ base ortonormal de \mathbb{R}^m , y A la matriz $t \cdot t^T$. Observar que $U_t A(w) = \langle t, w \rangle U_t(t)$, entonces

$$\begin{aligned} U_t A\left(\frac{t}{\|t\|}\right) &= e_1 \\ U_t A(v_i) &= 0. \end{aligned}$$

Por lo tanto

$$\text{Var}(U_t \nabla Z_i(t)) = \text{Diag}\left[r_i(\|t\|^2) \cdot \|t\|^2 + q_i(\|t\|^2), q_i(\|t\|^2), \dots, q_i(\|t\|^2)\right]$$

(Observar que $\text{Var}(U_t \nabla Z_i(t)) = U_t \text{Var}(\nabla Z_i(t)) U_t^T$). Entonces

$$\text{Var}\left(\frac{U_t \nabla Z_i(t)}{\sqrt{q_i(\|t\|^2)}}\right) = \text{Diag}(h_i(\|t\|^2), 1, \dots, 1). \quad (4.15)$$

Sean los vectores columnas

$$T_i = \frac{U_t \nabla Z_i(t)}{\sqrt{q_i(\|t\|^2)}} \quad (i = 1, \dots, m)$$

y

$$T := \begin{bmatrix} T_1 & \cdots & T_m \end{bmatrix}$$

la matriz formada por las columnas T_i para $i = 1, \dots, m$. Observar que

$$|\det(Z'(t))| = |\det T| \cdot \prod_{i=1}^m (q_i(\|t\|^2))^{1/2}. \quad (4.16)$$

Solo nos resta computar $|\det T|$. Para eso, sean W_1, \dots, W_m los vectores fila de T , es decir

$$T = \begin{bmatrix} W_1 \\ \vdots \\ W_m \end{bmatrix}.$$

Veamos como se distribuyen los vectores W_1, \dots, W_m . Para eso escribamos

$$T_j = \begin{pmatrix} T_j^{(1)} \\ \vdots \\ T_j^{(m)} \end{pmatrix}.$$

De (4.15) sabemos que

$$\begin{aligned} \text{Var}(T_j^{(i)}) &= 1 \quad \text{for } i = 2, \dots, m; j = 1, \dots, m \\ \text{Var}(T_j^{(1)}) &= h_j(\|t\|^2) \quad \text{for } j = 1, \dots, m. \end{aligned}$$

Se tiene que

$$W_i = (T_1^{(i)}, \dots, T_m^{(i)}) \quad (i = 1, \dots, m).$$

Por la independencia de los vectores T_j para $j = 1, \dots, m$ las coordenadas de W_i son independientes para $i = 1, \dots, m$. Se concluye de (4.15) que para $i = 2, \dots, m$ los vectores W_i tienen distribución Gaussiana estándar en \mathbb{R}^m . Además el vector W_1 tiene distribución Gaussiana centrada pero con matriz de varianza

$$\text{Var}(W_1) = \text{Diag}(h_1, \dots, h_m) \quad (4.17)$$

donde en todas las funciones del miembro derecho de (4.17) están evaluadas en $\|t\|^2$. Para computar $|\det T|$ basta calcular el volumen del paralelepípedo $\mathcal{P}_{W_1, \dots, W_m}$ engendrado por los vectores $\{W_1, \dots, W_m\}$, donde $\mathcal{P}_{W_1, \dots, W_m}$ es el conjunto de vectores de \mathbb{R}^m que pueden ser expresados por

$$\sum_{i=1}^m a_i W_i$$

con $0 \leq a_i \leq 1$ ($i = 1, \dots, m$).

Entonces se tiene

$$|\det T| = \text{vol}^{(m-1)}(W_1, \dots, W_{m-1}) \cdot d(W_m, L_{m-1})$$

donde $\text{vol}^{(l)}$ indica el volumen l -dimensional en \mathbb{R}^m , L_l es el subespacio lineal generado por $\{W_1, \dots, W_l\}$, y $d(W_l, L_{l-1})$ indica la distancia Euclídea de W_l al subespacio L_{l-1} (ver Figura 4.1).

Procediendo de la misma manera se obtiene:

$$|\det T| = \|W_1\| \cdot \prod_{i=2}^m d(W_i, L_{i-1}). \quad (4.18)$$

Ahora hay que tomar esperanza en ambos miembros de (4.18). La idea es usar la independencia de los W_i , y usar fuertemente la invarianza por isometrías de la distribución Gaussiana.

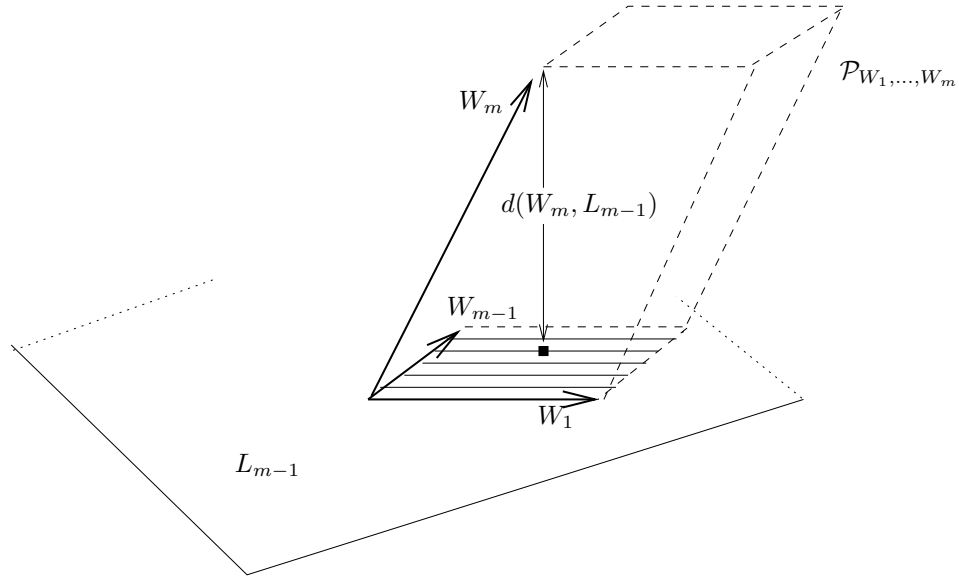


Figura 4.1: Paralelepípedo $\mathcal{P}_{W_1, \dots, W_m}$.

Sea $\sigma(W_1, \dots, W_l)$ la σ -álgebra engendrada por $\{W_1, \dots, W_l\}$ para $l = 1, \dots, m$. Entonces

$$\begin{aligned} E(|\det T|) &= E\left(\|W_1\| \cdot \prod_{i=2}^m d(W_i, L_{i-1})\right) \\ &= E\left[E\left(\|W_1\| \cdot \prod_{i=2}^m d(W_i, L_{i-1}) \mid \sigma(W_1, \dots, W_{m-1})\right)\right] \\ &= E\left[\|W_1\| \cdot \prod_{i=2}^{m-1} d(W_i, L_{i-1}) E\left(d(W_m, L_{m-1}) \mid \sigma(W_1, \dots, W_{m-1})\right)\right]. \end{aligned}$$

Para calcular

$$E\left(d(W_m, L_{m-1}) \mid \sigma(W_1, \dots, W_{m-1})\right)$$

consideremos el subespacio lineal L_{m-1} formado por los vectores W_1, \dots, W_{m-1} . Sean $\{v_1, \dots, v_m\}$ una base ortonormal (aleatoria) de \mathbb{R}^m tal que $\{v_1, \dots, v_{m-1}\}$ sea una base ortonormal de L_{m-1} . Podemos construir dicha base de forma tal que cada elemento sea medible respecto a $\sigma(W_1, \dots, W_{m-1})$. Eso se puede hacer mediante el proceso de Gramm-Schmidt.

Escribimos $W_m = \sum_{i=1}^m a_i v_i$ donde

$$a_m = \langle W_m, v_m \rangle$$

(ver Figura 4.2).

Observar que $\eta_1 := \langle W_m, v_m \rangle$ tiene distribución normal estándar debido a la independencia entre W_m y v_m . Entonces se tiene

$$d(W_m, L_{m-1}) = |\eta_1|.$$

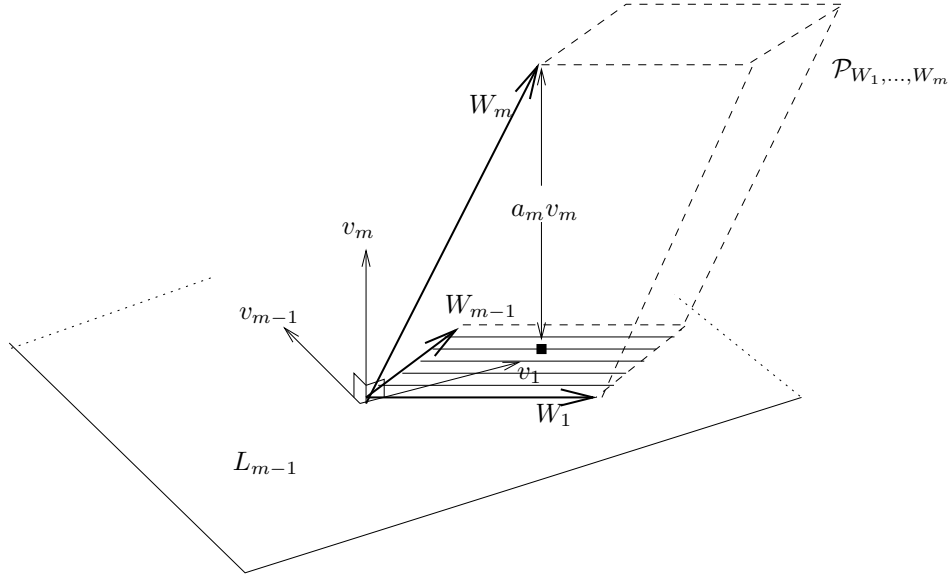


Figura 4.2: $\mathcal{P}_{W_1, \dots, W_m}$ representado en la base v_1, \dots, v_m .

Por lo tanto concluimos

$$E(|\det T|) = E(|\eta_1|) \cdot E\left[\|W_1\| \cdot \prod_{i=2}^m d(W_i, L_{i-1})\right]$$

donde $\eta_1 \sim \mathcal{N}(0, 1)$.

Procediendo de igual manera y condicionando sucesivamente obtenemos

$$E(|\det T|) = E(\|W_1\|) \cdot \prod_{l=1}^{m-1} E(\|\eta_l\|_l) \quad (4.19)$$

donde $\eta_l \sim \mathcal{N}(0, I_l)$ en \mathbb{R}^l , y $\|\cdot\|_l$ indica la norma Euclidea l -dimensional.

Definiendo por $K_l = E(\|\eta_l\|_l)$ y usando (4.14), (4.16) y (4.19) obtenemos

$$E(N^X(V)) = \frac{\prod_{l=1}^{m-1} K_l}{(\sqrt{2\pi})^m} \int_V \left[\prod_{i=1}^m q_i(\|t\|^2) \right]^{1/2} \cdot E_h(\|t\|^2) dt$$

Sabiendo que

$$K_l = \sqrt{2} \frac{\Gamma(\frac{l+1}{2})}{\Gamma(\frac{l}{2})}$$

se tiene mediante cálculos elementales que

$$\prod_{l=1}^n K_l = \frac{1}{\sqrt{2\pi}} 2^{\frac{n+1}{2}} \Gamma\left(\frac{n+1}{2}\right).$$

Entonces

$$\frac{\prod_{l=1}^{m-1} K_l}{(\sqrt{2\pi})^m} = \frac{\Gamma(\frac{m}{2})}{\sqrt{2\pi}^{\frac{m+1}{2}}}$$

obteniéndose

$$E(N^X(V)) = \frac{1}{\sqrt{2} \pi^{\frac{m+1}{2}}} \Gamma\left(\frac{m}{2}\right) \int_V \left[\prod_{i=1}^m q_i(\|t\|^2) \right]^{1/2} \cdot E_h(\|t\|^2) dt.$$

□

Capítulo 5

Smooth Analysis

5.1. Introducción

Consideremos

$$P_i(t) = 0 \quad (i = 1, \dots, m)$$

un sistema de ecuaciones polinomiales determinístico, donde el grado de $P_i(t)$ es d_i , y sea $X : (X_1, \dots, X_m) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ un campo polinomial Gaussiano centrado, donde los procesos

$$X_1(\cdot), \dots, X_m(\cdot)$$

a valores reales sean independientes, y donde cada uno tiene una ley invariante por isometrías de \mathbb{R}^m , como en la sección anterior. La idea de este capítulo es estudiar que sucede con el número de soluciones del nuevo sistema

$$P_i(t) + X_i(t) = 0 \quad (i = 1, \dots, m).$$

Más precisamente, se expondrá el resultado obtenido en [1].¹

Para nuestros fines, primero supondremos que el polinomio de covarianza es de la forma

$$Q^{(i)}(u, v, w) = Q_0(u, vw)$$

donde Q_0 es un polinomio de dos variables reales con coeficientes no-negativos. Para dichos polinomios, es fácil ver que son funciones de covarianza de algún proceso. Esto resulta de que:

- los mapas $(s, t) \rightarrow \langle s, t \rangle$ y $(s, t) \rightarrow \|s\|^2 \|t\|^2$ son funciones de covarianza²,
- cualquier combinación lineal de coeficientes no-negativos de funciones de covarianza es una función de covarianza,
- el producto de funciones de covarianzas es una función de covarianza.

¹Es recomendable recordar la sección *Objetivo* en la *Introducción*.

²Para el primer mapa basta tomar $\xi \sim \mathcal{N}(0, I_m)$ y tomar $X(t) = \langle \xi, t \rangle$, para el segundo basta tomar $Y(t) = \zeta \|t\|^2$ donde $\zeta \sim \mathcal{N}(0, 1)$.

En realidad, para lo que nos interesa, nos restringiremos al caso en que $Q^{(i)}$ sólo dependa de la variable u , es decir la función covarianza solo depende de $\langle s, t \rangle$. Es decir

$$Q^{(i)}(u) = \sum_{k=0}^{d_i} c_k^{(i)} u^k. \quad (5.1)$$

En este caso, la condición necesaria y suficiente para que $Q^{(i)}$ sea una covarianza es que $c_k^{(i)} \geq 0$ para todo $k = 0, \dots, d_i$.

Por ejemplo, el caso Shub-Smale (1.2) corresponde a la elección

$$c_k^{(i)} = \binom{d_i}{k} \quad (k = 0, 1, \dots, d_i).$$

Para eso basta observar que fijado i

$$\begin{aligned} r^{X_i}(s, t) &= E(X_i(s)X_i(t)) \\ &= \sum_{\|j\| \leq d_i} E\left((a_j^{(i)})^2\right) s^j t^j \\ &= \sum_{\|j\| \leq d_i} \binom{d_i}{j} s^j t^j \\ &= (1 + \langle s, t \rangle)^{d_i}, \end{aligned}$$

entonces

$$Q^{(i)}(u) = (1 + u)^{d_i} = \sum_{k=0}^{d_i} \binom{d_i}{k} u^k.$$

Se puede construir un proceso muy fácilmente, que tenga una covarianza de la forma

$$Q(u) = \sum_{k=0}^d c_k u^k.$$

con los coeficientes $c_k \geq 0$ dados para $k = 0, \dots, d$. Para eso tomemos un proceso X de la forma

$$X(t) = \sum_{\|j\| \leq d_i} a_j t^j$$

donde las variables aleatorias a_j son Gaussianas, centradas, e independientes. Notando por $\sigma_j := \text{Var}(a_j)$, $\eta_i := s_i t_i$ para $i = 1, \dots, m$ y $\eta := \eta_1 \cdots \eta_m (= s \cdot t)$, se tiene

$$\begin{aligned} r^X(s, t) &= E(X(s)X(t)) \\ &= \sum_{\|j\| \leq d_i} \sigma_j \eta^j \end{aligned} \quad (5.2)$$

Por otro lado

$$\begin{aligned} Q(\langle s, t \rangle) &= \sum_{k=0}^d c_k (\eta_1 + \cdots + \eta_m)^k \\ &= \sum_{k=0}^d c_k \left[\sum_{\|l\|=k} \binom{k}{l} \eta^l \right] \end{aligned} \quad (5.3)$$

igualando (5.2) con (5.3) se tiene

$$\sum_{\|j\| \leq d_i} \sigma_j \eta^j = \sum_{k=0}^d c_k \left[\sum_{\|l\|=k} \binom{k}{l} \eta^l \right].$$

Entonces,

$$\sigma_j = c_{\|j\|} \binom{\|j\|}{j} = c_{\|j\|} \frac{\|j\|!}{j_1! \cdots j_m!} \quad 0 \leq \|j\| \leq d. \quad (5.4)$$

Hemos concluido que si X_i es un proceso polinomial de la forma

$$X_i(t) = \sum_{\|j\| \leq d_i} a_j^{(i)} t^j \quad \text{para } i = 1, \dots, m$$

con $a_j^{(i)}$ variables aleatorias Gaussianas, centradas, independiente y con varianza

$$\text{Var}(a_j^{(i)}) = c_{\|j\|}^{(i)} \frac{\|j\|!}{j!} \quad (i = 1, \dots, m; \|j\| \leq d_i)$$

entonces su función de covarianza esta dada por (5.1).

5.2. Teorema de Perturbación Aleatoria

Antes de enunciar el resultado obtenido en [1] es necesario establecer algunas hipótesis y agregar notaciones adicionales.

Notaremos por $P : \mathbb{R}^m \rightarrow \mathbb{R}^m$ al campo vectorial

$$P(t) := \begin{pmatrix} P_1(t) \\ \vdots \\ P_m(t) \end{pmatrix}.$$

Vamos a asumir que para cada t fijo las distribuciones uni-dimensionales de $X_i(t)$ no degeneren para $i = 1, \dots, m$. Esto significa que la función de covarianza $Q^{(i)}$ no se anule para $u \geq 0$. También asumiremos que los polinomios $Q^{(i)}$ tienen grado *efectivo* d_i , esto significa que

$$c_{d_i}^{(i)} > 0 \quad (i = 1, \dots, m).$$

El resultado enunciado en el *Teorema 11* más abajo, es un resultado de carácter asintótico en el tamaño del sistema, es decir, es un resultado que interesa cuando m es grande. Es por eso que necesitaremos hipótesis que controlen el comportamiento del sistema (tanto aleatorio como determinístico) independientemente de m . Se puede probar mediante un cálculo elemental que para cada polinomio $Q^{(i)}$, cuando $u \rightarrow +\infty$:

$$q_i(u) \sim \frac{d_i}{1+u} \quad (5.5a)$$

$$h_i(u) \sim \frac{c_{d_i-1}^{(i)}}{d_i c_{d_i}^{(i)}} \cdot \frac{1}{1+u}. \quad (5.5b)$$

Sin embargo las relaciones (5.5a) y (5.5b) pueden variar según m , pues el sistema polinomial P y los polinomios $Q^{(i)}$ (para $i = 1, \dots, m$) pueden ir cambiando a medida que m varia (las funciones q_i, r_i, h_i para $i = 1, \dots, m$ dependen de m , pero para no hacer pesada la notación las omitimos). Es por esta razón que necesitamos algunas relaciones que controlen el sistema independientemente de m .

H_1) h_i es independiente de i ($i = 1, \dots, m$), pero se permite que pueda variar con m . Notaremos $h = h_i$.

H_2) Existen constantes positivas D_i, E_i ($i = 1, \dots, m$) y \underline{q} tales que:

$$\underline{q} \leq (1+u)q_i(u) \leq D_i, \quad \text{y} \quad q_i(u) + \frac{E_i}{(1+u)^2} \geq \frac{D_i}{1+u}. \quad (5.6)$$

para todo $u \geq 0$. Además

$$\max_{1 \leq i \leq m} D_i, \quad \max_{1 \leq i \leq m} E_i$$

permanecen acotadas por constantes \bar{D}, \bar{E} respectivamente. Las constantes \underline{q}, \bar{D} y \bar{E} son independientes de m . También existen constantes \underline{h}, \bar{h} tales que

$$\underline{h} \leq (1+u)h(u) \leq \bar{h} \quad (5.7)$$

para todo $u \geq 0$. Estas constantes son también independientes de m .

Este primer conjunto de hipótesis concierne solamente a la parte aleatoria del sistema. Necesitamos ahora un segundo conjunto de hipótesis en el cual se establezca como debe ser la relación de la “señal” P con el “ruido” X . Para estas hipótesis necesitamos un poco de notación adicional.

Sea P un polinomio en m variables reales y coeficientes reales de grado d , y Q un polinomio de una variable con coeficientes no negativos de grado efectivo d , $Q(u) = \sum_{k=0}^d c_k u^k$. Asumiremos que el polinomio Q no se anula. Definimos para $r > 0$

$$L(P, Q, r) := \inf_{\|t\| \geq r} \frac{P(t)^2}{Q(\|t\|^2)},$$

y

$$H(P, Q) := \sup_{t \in \mathbb{R}^m} \left\{ (1 + \|t\|) \cdot \left\| \nabla \left(\frac{P}{\sqrt{Q(\|t\|^2)}} \right) (t) \right\| \right\}$$

$$K(P, Q) := \sup_{t \in \mathbb{R}^m \setminus \{0\}} \left\{ (1 + \|t\|^2) \cdot \left| \frac{\partial}{\partial \rho} \left(\frac{P}{\sqrt{Q(\|t\|^2)}} \right) (t) \right| \right\}$$

donde $\frac{\partial}{\partial \rho}$ denota la derivada en la dirección de definida por $\frac{t}{\|t\|}$, en el punto $t \neq 0$. Es fácil ver mediante un cálculo elemental que

$$H(P, Q) < \infty, \quad K(P, Q) < \infty.$$

para cualquier par de polinomio P y Q que verifican las condiciones establecidas.

Con esta notación podemos enunciar las restantes hipótesis

$H_3)$

$$A_m := \frac{1}{m} \cdot \sum_{i=1}^m \frac{H^2(P_i, Q^{(i)})}{i} = o(1) \quad \text{cuando } m \rightarrow +\infty \quad (5.8a)$$

$$B_m := \frac{1}{m} \cdot \sum_{i=1}^m \frac{K^2(P_i, Q^{(i)})}{i} = o(1) \quad \text{cuando } m \rightarrow +\infty. \quad (5.8b)$$

$H_4)$ Existen constantes positivas r_0, l tales que, si $r \geq r_0$:

$$L(P_i, Q^{(i)}, r) \geq l \quad \text{para todo } i = 1, \dots, m.$$

Ahora estamos en condiciones de enunciar el resultado:

Teorema 11. *Bajo las hipótesis $H_1), \dots, H_4)$, se tiene que existen constantes positivas C, θ , con $0 < \theta < 1$ tal que:*

$$E(N^{P+X}) \leq C \theta^m E(N^X). \quad (5.9)$$

Observaciones generales sobre el enunciado

Observar que el problema que estamos estudiando es exactamente el mismo si intercambiamos de lugar los polinomios del sistema

$$P_i(t) + X_i(t) = 0 \quad (i = 1, \dots, m).$$

Pero si observamos las condiciones (5.8a) y (5.8b) en la hipótesis $H_3)$, es evidente que dependen del orden en que se encuentre el sistema. Para solucionar esto, pediremos que exista un orden $i = 1, \dots, m$ del sistema en el cual las hipótesis (5.8a) y (5.8b) son satisfechas.

Observación sobre la naturaleza de las hipótesis

Antes de comenzar con la demostración del *Teorema 11* daremos un significado a las hipótesis establecidas.

- La hipótesis $H_1)$ es de carácter puramente técnico, en el sentido de que quizás se pueda extender el resultado al caso dependiente de i .
- La hipótesis $H_2)$ expresa cierto control sobre la perturbación cuando trabajamos con resultados asintóticos.
- La hipótesis $H_3)$ puede interpretarse como un control del cociente entre la señal y el ruido. Más precisamente da una cota superior sobre el gradiente de este cociente. La hipótesis $H_4)$ da una cota en el otro sentido.
- Veamos porque la necesidad de imponer hipótesis sobre la relación entre la señal y el ruido. Para eso, basta observar lo siguiente: consideremos el sistema

$$P_i(t) + \sigma X_i(t) = 0 \quad (i = 1, \dots, m) \quad (5.10)$$

donde σ es un parámetro real positivo. Si hacemos que $\sigma \downarrow 0$ entonces el valor esperado del número de raíces del sistema (5.10) tiende al número de raíces del sistema determinístico $P_i(t) = 0$, ($i = 1, \dots, m$). Pero es claro que en principio el sistema puede degenerar y no es esperable un resultado del tipo del *Teorema 11*. En este caso la relación de la señal sobre el ruido tiende a infinito. Ahora, si $\sigma \rightarrow +\infty$, tampoco es esperable un resultado del tipo del *Teorema 11* pues el valor esperado del número de raíces tiende a $E(N^X)$. En este caso la relación señal sobre el ruido tiende a cero.

Demostración. La demostración sigue las líneas de la demostración del *Teorema 10* en la página 55. Sea

$$Z_j(t) := \frac{X_j(t) + P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \quad (j = 1, \dots, m)$$

y notemos por Z al sistema

$$Z := \begin{pmatrix} Z_1 \\ \vdots \\ Z_m \end{pmatrix}.$$

Claramente $N^{X+P} = N^Z$. Entonces igual que en el caso centrado obtenemos que $Z_j(t)$ y $\nabla Z_j(t)$ son independientes y entonces tenemos que la fórmula de Rice toma la siguiente forma:

$$E(N^Z(V)) = \int_V E(|\det(Z'(t))|) \cdot p_{Z(t)}(0) dt, \quad (5.11)$$

donde

$$p_{Z(t)}(0) = \frac{1}{(2\pi)^{m/2}} \cdot \exp \left\{ -\frac{1}{2} \left(\frac{P_1(t)^2}{Q^{(1)}(\|t\|^2)} + \dots + \frac{P_m(t)^2}{Q^{(m)}(\|t\|^2)} \right) \right\}.$$

Nuestro objetivo ahora es calcular $E(|\det(Z'(t))|)$. Para eso procedemos igual que en el caso centrado. Entonces se obtiene

$$\text{Cov} \left(\frac{\partial Z_i}{\partial t_\alpha}(t), \frac{\partial Z_i}{\partial t_\beta}(t) \right) = r_i(\|t\|^2) t_\alpha t_\beta + q_i(\|t\|^2) \delta_{\alpha\beta}$$

para $i, j, \alpha, \beta = 1, \dots, m$, donde las funciones q_i, r_i están definidas en (4.12a) y (4.12b).

Aplicamos una transformación ortogonal al espacio \mathbb{R}^m , (a la cual llamaremos U_t) de manera tal que nos exprese las coordenadas en una nueva base con primer elemento $t/\|t\|$ (para más detalles ver página 55), obtenemos

$$\text{Var} \left(\frac{U_t \nabla Z_j(t)}{\sqrt{q_j(\|t\|^2)}} \right) = \text{Diag} (h(\|t\|^2), 1, \dots, 1) \quad (5.12)$$

donde notamos $\nabla Z_j(t)$ como el vector columna,

$$\nabla Z_j(t) := \begin{pmatrix} \frac{\partial Z_j}{\partial t_1}(t) \\ \vdots \\ \frac{\partial Z_j}{\partial t_m}(t) \end{pmatrix}.$$

Entonces

$$\frac{U_t \nabla Z_j(t)}{\sqrt{q_j(\|t\|^2)}} = \zeta_j + \alpha_j \quad (j = 1, \dots, m)$$

donde ζ_j es un vector con distribución Gaussiana, centrado en \mathbb{R}^m que tiene matriz de covarianza dada por (5.12), y α_j es el vector determinístico

$$\alpha_j := \frac{U_t \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right)}{\sqrt{q_j(\|t\|^2)}} = \begin{pmatrix} \alpha_j^{(1)} \\ \vdots \\ \alpha_j^{(m)} \end{pmatrix} \quad (j = 1, \dots, m). \quad (5.13)$$

Observar que ζ_1, \dots, ζ_m son independientes.

Sea T la matriz aleatoria $m \times m$ que tiene como columnas a $\zeta_j + \alpha_j$, ($j = 1, \dots, m$). Entonces tenemos

$$|\det(Z'(t))| = |\det(T)| \cdot \prod_{i=1}^m (q_i(\|t\|^2))^{1/2}$$

por lo tanto

$$E(|\det(Z'(t))|) = \prod_{i=1}^m (q_i(\|t\|^2))^{1/2} \cdot E(|\det(T)|). \quad (5.14)$$

Nuestro objetivo es computar el lado derecho de (5.14).

Sean η_1, \dots, η_m las columnas de T , es decir

$$\eta_j = \zeta_j + \alpha_j \quad (j = 1, \dots, m)$$

con

$$\zeta_j = \begin{pmatrix} \zeta_j^{(1)} \\ \vdots \\ \zeta_j^{(m)} \end{pmatrix},$$

donde las variables $\zeta_j^{(i)}$ son Gaussianas, centradas, independientes y

$$\text{Var}(\zeta_j^{(i)}) = 1 \quad \text{para } i = 2, \dots, m; j = 1, \dots, m$$

$$\text{Var}(\zeta_j^{(1)}) = h(\|t\|^2) \quad \text{para } j = 1, \dots, m.$$

En otras palabras,

$$\zeta_j = \begin{pmatrix} \sqrt{h(\|t\|^2)} \psi_j^{(1)} \\ \psi_j^{(2)} \\ \vdots \\ \psi_j^{(m)} \end{pmatrix}$$

siendo $\psi_j = \begin{pmatrix} \psi_j^{(1)} \\ \vdots \\ \psi_j^{(m)} \end{pmatrix} \sim \mathcal{N}(0, I_m)$.

Entonces

$$\begin{aligned} |\det(T)| &= |\det(\zeta_1 + \alpha_1, \dots, \zeta_m + \alpha_m)| \\ &= \sqrt{h(\|t\|^2)} \cdot |\det(\psi_1 + \tilde{\alpha}_1, \dots, \psi_m + \tilde{\alpha}_m)| \end{aligned}$$

donde

$$\tilde{\alpha}_j := \begin{pmatrix} \alpha_j^{(1)} / \sqrt{h(\|t\|^2)} \\ \alpha_j^{(2)} \\ \vdots \\ \alpha_j^{(m)} \end{pmatrix} \quad \text{para } j = 1 \dots m, \quad (5.15)$$

siendo $\alpha_j^{(k)}$ la coordenada k -ésima del vector α_j dado en (5.13).

Nuevamente interpretamos $|\det(T)|$ como el volúmen del paralelepípedo engendrado por η_1, \dots, η_m . Usando la invarianza por isometrías del espacio de la distribución Gaussiana obtenemos la siguiente cota

$$E(|\det T|) \leq \sqrt{h(\|t\|^2)} \cdot \prod_{j=1}^m E(\|\xi_j + c_j(t)\|_j) \quad (5.16)$$

donde como antes $\|\cdot\|_j$ denota la norma Euclidea en \mathbb{R}^j ($\|\cdot\| = \|\cdot\|_m$), ξ_j es un vector con distribución normal estándar en \mathbb{R}^j y $c_j(t)$ es un vector determinístico en \mathbb{R}^j teniendo norma

$$\|\tilde{\alpha}_j\|, \quad j = 1, \dots, m$$

(ver *Afirmación 1* en *Complementos Técnicos* en la página 76).

Notamos por

$$\gamma_j(c) := E(\|\xi_j + c\|_j)$$

donde $c \in \mathbb{R}^j$ es un vector determinístico. Entonces

$$\gamma_j(c) \leq \gamma_j(0) \left(1 + \frac{1}{2j} \|c\|_j^2\right). \quad (5.17)$$

(ver *Lema 8* en *Complementos Técnicos*). Usando la desigualdad elemental

$$1 + x \leq e^x \quad \forall x \in \mathbb{R},$$

reemplazando en (5.11), y usando (5.14) y (5.17) se tiene:

$$\begin{aligned} E(N^Z) &\leq \frac{1}{(2\pi)^{m/2}} L_m \cdot \int_{\mathbb{R}^m} \left\{ \sqrt{h(\|t\|^2)} \cdot \right. \\ &\quad \left. \left(\prod_{i=1}^m q_i(\|t\|^2) \right)^{1/2} \cdot \exp \left[-\frac{1}{2} \sum_{i=1}^m \frac{P_i(t)^2}{Q^{(i)}(\|t\|^2)} + \frac{1}{2} \sum_{j=1}^m \|c_j(t)\|_j^2 \frac{1}{j} \right] \right\} dt \end{aligned} \quad (5.18)$$

donde

$$L_m := \prod_{i=1}^m E(\|\xi_i\|_i) = \frac{1}{\sqrt{2\pi}} 2^{(m+1)/2} \Gamma\left(\frac{m+1}{2}\right).$$

Lo que resta, es tratar de acotar el integrando en (5.18)

Observar que en realidad $\tilde{\alpha}_j$ difiere de $\frac{1}{\sqrt{q_j(\|t\|^2)}} U_t \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right)$ en su primer coordenada donde

$$\alpha_j^{(1)} = \frac{1}{\sqrt{q_j(\|t\|^2)}} \langle u_{1\bullet}, \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right) \rangle \quad (5.19)$$

$$= \frac{1}{\sqrt{q_j(\|t\|^2)}} \frac{\partial}{\partial u_{1\bullet}} \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right) \quad (5.20)$$

siendo $u_{1\bullet}$ la primer fila de la matriz U_t que no es otra cosa que $t/\|t\|$ (ver página 55). Entonces usando las hipótesis H_2) y H_3) se tiene

$$|\tilde{\alpha}_j^{(1)}| = \frac{1}{\sqrt{h(\|t\|^2) q_j(\|t\|^2)}} \cdot \left| \frac{\partial}{\partial \rho} \frac{P_j(\|t\|^2)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right| \leq \frac{1}{\sqrt{h\underline{q}}} K(P_j, Q^{(j)})$$

Como

$$\|\tilde{\alpha}_j\|^2 = \left(\frac{\alpha_j^{(1)}}{\sqrt{h(\|t\|^2)}} \right)^2 + (\alpha_j^{(2)})^2 + \dots + (\alpha_j^{(m)})^2,$$

podemos acotar $\|\tilde{\alpha}_j\|^2$ por

$$\|\tilde{\alpha}_j\|^2 \leq \left(\frac{\alpha_j^{(1)}}{\sqrt{h(\|t\|^2)}} \right)^2 + \|\alpha_j\|^2, \quad (5.21)$$

donde nuevamente por H_2) y H_3)

$$\|\alpha_j\| = \frac{1}{\sqrt{q_j(\|t\|^2)}} \left\| \nabla \left(\frac{P_j(t)}{\sqrt{Q^{(j)}(\|t\|^2)}} \right) \right\| \leq \frac{1}{\sqrt{\underline{q}}} H(P_j, Q^{(j)}).$$

Entonces obtenemos

$$\|\tilde{\alpha}_j\|^2 \leq \frac{1}{h\underline{q}} K^2(P_j, Q^{(j)}) + \frac{1}{\underline{q}} H^2(P_j, Q^{(j)})$$

lo cual implica

$$\sum_{j=1}^m \|c_j\|_j^2 \cdot \frac{1}{j} \leq \frac{1}{\underline{q}} m A_m + \frac{1}{h\underline{q}} m B_m. \quad (5.22)$$

Si reemplazamos (5.22) en (5.18) obtenemos la siguiente cota

$$E(N^Z) \leq s_m H_m$$

donde

$$s_m := \left(\frac{\bar{h}}{\underline{h}} \right)^{1/2} \cdot e^{\frac{1}{2} \left(\frac{1}{\underline{q}} m A_m + \frac{1}{\bar{h}\underline{q}} m B_m \right)} = e^{o(m)} \quad (\text{as } m \rightarrow +\infty)$$

y

$$H_m := \frac{\Gamma\left(\frac{m}{2}\right)}{\sqrt{2} \pi^{(m+1)/2}} \cdot \int_{\mathbb{R}^m} \left(\prod_{i=1}^m q_i(\|t\|^2) \right)^{1/2} \cdot \sqrt{h(\|t\|^2)} E(\|\xi_m\|) e^{-\frac{1}{2} \sum_{i=1}^m \frac{P_i(t)^2}{Q^{(i)}(\|t\|^2)}} dt \quad (5.23)$$

Observar que la fórmula (5.23) y la fórmula para el caso centrado (4.13) solo difieren en la exponencial, la cual para valores grandes de $\|t\|$ nos va a ayudar a controlar la cantidad de soluciones del nuevo sistema.

Para estudiar el comportamiento de la exponencial en H_m es necesario partir el dominio de integración. Para eso escribimos

$$H_m = H_m^{(1)}(r) + H_m^{(2)}(r)$$

donde $H_m^{(1)}(r)$ corresponde a integrar en el conjunto $\|t\| \leq r$ y $H_m^{(2)}(r)$ corresponde a integrar en $\|t\| > r$.

Primero consideremos r suficientemente grande para que la hipótesis H_4) sea satisfecha (basta $r \geq r_0$). Entonces obtenemos la siguiente cota para $H_m^{(2)}(r)$,

$$H_m^{(2)}(r) \leq s_m \cdot e^{-lm/2} E(N^X). \quad (5.24)$$

Ahora ocupemosnos de $H_m^{(1)}(r)$. Primero acotamos groseramente la exponencial en el integrando de (5.23) por 1. Luego usando la hipótesis H_2) se tiene

$$H_m^{(1)}(r) \leq \frac{1}{\sqrt{2} \pi^{(m+1)/2}} \Gamma\left(\frac{m}{2}\right) E(\|\xi_m\|) \cdot \int_{\{\|t\| \leq r\}} \frac{(\prod_{i=1}^m D_i)^{1/2}}{(1 + \|t\|^2)^{m/2}} \cdot \frac{\bar{h}^{-1/2}}{(1 + \|t\|^2)^{1/2}} dt$$

y pasando a polares

$$H_m^{(1)}(r) \leq \frac{1}{\sqrt{2} \pi^{(m+1)/2}} \Gamma\left(\frac{m}{2}\right) \bar{h}^{-1/2} E(\|\xi_m\|) \left(\prod_{i=1}^m D_i^{1/2} \right) \sigma_{m-1} \int_0^r \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho \quad (5.25)$$

donde σ_{m-1} es la medida de área $(m-1)$ -dimensional de la esfera S^{m-1} . Un cálculo sencillo muestra que

$$\int_0^r \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho \leq \frac{\pi}{2} \left(\frac{r^2}{1 + r^2} \right)^{\frac{m-1}{2}}.$$

Entonces

$$H_m^{(1)}(r) \leq \alpha_m \bar{h}^{-1/2} \pi \left(\frac{r^2}{1 + r^2} \right)^{\frac{m-1}{2}}. \quad (5.26)$$

donde

$$\alpha_m = \frac{1}{\pi^{(m+1)/2}} \Gamma\left(\frac{m}{2}\right) E(\|\xi_m\|) \left(\prod_{i=1}^m D_i^{1/2} \right) \sigma_{m-1}$$

Recordando la fórmula (4.13) y usando H_2) tenemos

$$E(N^X) \geq \alpha_m \underline{h}^{1/2} \int_0^{+\infty} \left[\prod_{i=1}^m \left(\frac{D_i}{1 + \|t\|^2} - \frac{E_i}{(1 + \|t\|^2)^2} \right)^{1/2} \right] \frac{1}{(1 + \|t\|^2)^{1/2}} dt$$

y cambiando a coordenadas polares se tiene

$$\begin{aligned} E(N^X) &= \alpha_m \underline{h}^{1/2} \int_0^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} \prod_{i=1}^m \left(1 - \frac{F_i}{1 + \rho^2} \right)^{1/2} d\rho \\ &\geq \alpha_m \underline{h}^{1/2} \int_a^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} \prod_{i=1}^m \left(1 - \frac{F_i}{1 + \rho^2} \right)^{1/2} d\rho \\ &\geq \alpha_m \underline{h}^{1/2} \left(1 - \frac{\max_{1 \leq i \leq m} F_i}{1 + a^2} \right)^{1/2} \int_a^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho \end{aligned}$$

donde notamos por $F_i = E_i/D_i$, ($i = 1, \dots, m$). Por H_1) se tiene

$$F_i \leq \frac{\max_{1 \leq k \leq m} E_k}{\underline{q}} \quad (i = 1, \dots, m).$$

y por tanto $\overline{F_m} := \max_{1 \leq i \leq m} F_i$ permanece acotado. Entonces

$$E(N^X) \geq \alpha_m \overline{h}^{1/2} \left(1 - \frac{\overline{F_m}}{1 + a^2} \right)^{1/2} \int_a^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho.$$

Elegimos a suficientemente grande para que

$$\lambda^2 = \frac{r^2}{1 + r^2} < 1 - \frac{\overline{F_m}}{1 + a^2} = \nu^2.$$

Entonces, usando la equivalencia

$$\int_a^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho \sim \sqrt{\frac{\pi}{2m}} \quad \text{cuando } m \rightarrow +\infty.$$

para $a > 0$ fijo, se tiene

$$\frac{H_m^{(1)}(r)}{E(N^X)} \leq \frac{\overline{h}^{1/2}}{\underline{h}^{1/2}} \frac{\pi}{2} \left(\frac{\lambda}{\nu} \right)^m \frac{1}{\int_a^{+\infty} \frac{\rho^{m-1}}{(1 + \rho^2)^{(m+1)/2}} d\rho} \leq C_1 \cdot \lambda_1^m \quad (5.27)$$

tomando $\lambda/\nu < \lambda_1 < 1$, y siendo C_1 una constante positiva. Entonces obtenemos

$$H_m^1(r) \leq C_1 \lambda_1^m E(N^X) \quad (5.28)$$

Resumiendo, hemos obtenido de (5.24) y de (5.28)

$$E(N^{P+X}) \leq s_m \left[C_1 \lambda_1^m + e^{-lm/2} \right] E(N^X) \leq C \theta^m E(N^X)$$

para constantes positivas C, θ , $0 < \theta < 1$. □

5.3. Ejemplos

Ahora nos dedicaremos a mostrar algunos ejemplos, clasificándolos en distintos tipos de ruidos que se pueden considerar. Debido a la gran variedad de señales que se pueden considerar solo nos remitiremos a mostrar como afectan distintos ruidos a una misma señal.

Aunque la notación lo omita, es importante recordar que los polinomios $Q^{(i)}$ y P_i para $i = 1, \dots, m$ también son polinomios que dependen de m .

5.3.1. Shub-Smale

Como ya vimos, el caso Shub-Smale corresponde a

$$q_i(u) = \frac{d_i}{1+u}, \quad h_i(u) = h(u) = \frac{1}{1+u}.$$

Es fácil chequear que las hipótesis $H_1)$ y $H_2)$ son verificadas si asumimos que los grados d_i permanecen uniformemente acotados por una constante \bar{d} .

Las hipótesis $H_3)$ y $H_4)$ son relacionadas con la señal P . Vemos dos ejemplos de naturaleza totalmente distinta.

(i) Sea

$$P_i(t) = \|t\|^{d_i} - r^{d_i},$$

donde d_i es par (simplemente para que sea un polinomio) y $r > 0$. Asumimos que r permanece uniformemente acotado a medida que varia m por una constante R . Se tiene mediante un calculo elemental

$$\begin{aligned} \frac{\partial}{\partial \rho} \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) &= \frac{d_i \|t\|^{d_i-1} + d_i r^{d_i} \|t\|}{(1 + \|t\|^2)^{\frac{d_i}{2}+1}} \leq \frac{d_i(1 + r^{d_i})}{(1 + \|t\|^2)^{3/2}} \\ \nabla \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) &= \frac{d_i \|t\|^{d_i-2} + d_i r^{d_i}}{(1 + \|t\|^2)^{\frac{d_i}{2}+1}} \cdot t \end{aligned}$$

lo cual implica

$$\left\| \nabla \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) (t) \right\| \leq \frac{d_i(1 + r^{d_i})}{(1 + \|t\|^2)^{3/2}}.$$

Veamos que $H_3)$ se verifica. Recordando las notaciones de $H(P, Q)$ y $K(P, Q)$ en la página 63 se tiene

$$H(P_i, Q^{(i)}) \leq \sup_{t \in \mathbb{R}^m} \left\{ (1 + \|t\|) \cdot \frac{d_i(1 + r^{d_i})}{(1 + \|t\|^2)^{3/2}} \right\}$$

donde

$$\frac{1 + \|t\|}{(1 + \|t\|^2)^{3/2}} \leq 1 \quad \forall t \in \mathbb{R}^m.$$

Entonces

$$H(P_i, Q^{(i)}) \leq d_i(1 + r_i^{d_i}).$$

Lo mismo ocurre con $K(P_i, Q^{(i)})$. Entonces

$$\begin{aligned} A_m &\leq \frac{1}{m} \sum_{i=1}^m \frac{d_i^2 (1 + r_i^d)^2}{i} \\ &\leq \bar{d}^2 (1 + R^{\bar{d}})^2 \underbrace{\frac{1}{m} \sum_{i=1}^m \frac{1}{i}}_{o(1)} \end{aligned}$$

donde $\frac{1}{m} \sum_{i=1}^m \frac{1}{i} = o(1)$. Análogo para B_m .

Para chequear la hipótesis H_4), basta estudiar el comportamiento de la función

$$f_i(\rho) = \frac{(\rho^{d_i} - r^{d_i})^2}{(1 + \rho^2)^{d_i}}$$

y observar que $\lim_{|\rho| \rightarrow +\infty} f_i(\rho) = 1$ uniforme en $i = 1, \dots, m$, donde además $f_i(\rho) = 0$ si y solo si $\rho = r$, siendo creciente en el intervalo $[r, +\infty)$. Basta tomar $r_0 = 2R$ en H_4) y l suficientemente chico.

Lo interesante de este caso es que el sistema $P_i(t) = 0$ ($i = 1, \dots, m$) tiene infinitas raíces (la esfera centrada en el origen y de radio r es solución). Sin embargo el número esperado de soluciones del sistema perturbado es geoméricamente menor que el esperado en el caso Shub-Smale, cuando m es grande.

- (ii) Veamos otro caso de naturaleza distinta al anterior. Sea T un polinomio de grado d en una variable real con d raíces reales distintas. Definamos el sistema

$$P_i(t_1, \dots, t_m) = T(t_i) \quad (i = 1, \dots, m).$$

Es fácil chequear que este caso también cae en las hipótesis. Entonces se tiene que existen constantes positivas C, θ con $0 < \theta < 1$ tal que

$$E(N^{P+X}) \leq C \theta^m d^{m/2}$$

donde $d^{m/2}$ es el valor esperado del número de soluciones en el caso Shub-Smale. Sin embargo $N^P = d^m$. Nuevamente se observa claramente el efecto de la perturbación.

5.3.2. $Q^{(i)} = Q$, con raíces solamente reales

Vamos a asumir que $Q^{(i)}$ no dependen de i , $Q^{(i)} = Q$, donde Q es un polinomio con todas sus raíces reales y diferentes. Observar que por hipótesis el polinomio $Q(u)$ no se anula para $u \geq 0$, entonces todas sus raíces deben ser estrictamente negativas. Llamaremos,

$$-\alpha_1, \dots, -\alpha_d$$

las raíces de $Q(u)$ (es decir $Q(u) = (u + \alpha_1) \cdots (u + \alpha_d)$) donde

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_d.$$

Observar que se puede asumir que $\alpha_1 \geq 1$. Ésto se debe a que si hacemos una homotecia del espacio \mathbb{R}^m con factor $\sqrt{\alpha_1}$ se tiene que el número de raíces del sistema se mantiene inalterado, y el nuevo polinomio Q tiene ahora $\alpha_1 \geq 1$.

Vamos a asumir igual que antes que el grado d esta uniformemente acotado al variar m , y vamos a pedir que las raíces también estén acotadas,

$$\alpha_k \leq \bar{\alpha} \quad (k = 1, \dots, d)$$

por $\bar{\alpha}$ independiente de m .

No es difícil chequear que

$$q_i(u) = q(u) = \sum_{k=1}^d \frac{1}{u + \alpha_k} \quad , \quad h_i(u) = h(u) = \frac{1}{q(u)} \sum_{k=1}^d \frac{\alpha_k}{(u + \alpha_k)^2}.$$

Se puede verificar (5.6), eligiendo $D_i = d$, $E_i = d \cdot \max_{1 \leq k \leq d} (\alpha_k - 1)$. Similarmente se prueba (5.7).

Tomando la señal del primer caso en el ejemplo anterior, es decir

$$P_i(t) = \|t\|^{d_i} - r^{d_i}$$

en las mismas condiciones que antes, se tiene

$$\left| \frac{\partial}{\partial \rho} \left(\frac{P_i}{\sqrt{Q^{(i)}}} \right) \right| \leq d_i (\bar{\alpha} + r^{d_i}) \frac{1}{(1 + \|t\|^2)^{3/2}}.$$

Se sigue que $K(P_i, Q^{(i)})$ esta uniformemente acotado. Análogo para $H(P_i, Q^{(i)})$. Con respecto a H_4) es fácil ver que

$$L(P_i, Q^{(i)}, r) \geq \left(\frac{1}{1 + \bar{\alpha}} \right)^{\bar{d}}$$

para $i = 1, \dots, m$ y cualquier $r \geq 1$. Entonces podemos aplicar el *Theorem 11* para este caso.

5.3.3. $Q^{(i)} = Q^{l_i}$ con $i = 1, \dots, m$.

Consideramos ahora el caso en que $Q^{(i)}$ es de la forma

$$Q^{(i)}(u, v, w) = [Q(u)]^{l_i} \quad (i = 1, \dots, m)$$

donde $Q(u) = \sum_{k=0}^{\nu} b_k u^k$ es un polinomio de grado ν , y donde los coeficientes son positivos ($b_k > 0$ para $k = 0, \dots, \nu$).

Como hemos venido advirtiendo, aquí el polinomio Q puede depender de m , tanto como los exponentes l_1, \dots, l_m . Notar que $d_i = \nu \cdot l_i$ ($i = 1, \dots, m$).

Observar que el caso Shub-Smale corresponde a la elección $Q(u) = 1 + u$, $l_i = d_i$ ($i = 1, \dots, m$).

Tenemos que

$$q_i(u) = l_i \frac{Q'(u)}{Q(u)},$$

$$h_i(u) = h(u) = 1 - u \frac{Q'^2(u) - Q(u)Q''(u)}{Q(u)Q'(u)}$$

donde obviamente H_1) se satisface.

Veamos que ocurre con H_2). Primero veamos que

$$D_i \leq (1+u)q_i(u)$$

con la elección $D_i = d_i$ para $i = 1, \dots, m$. Como $Q(u) > 0$ para todo $u \geq 0$, basta ver que

$$0 \leq d_i Q(u) - (1+u)l_i Q'(u). \quad (5.29)$$

Es fácil ver que podemos escribir el miembro izquierdo de (5.29) como

$$\sum_{k=0}^{\nu-1} a_k u^k$$

donde

$$a_k = l_i [(\nu - k)b_k - (k + 1)b_{k+1}] \quad (k = 0, \dots, \nu - 1).$$

Entonces basta asumir

$$b_{k+1} \leq \frac{\nu - k}{k + 1} b_k \quad (k = 0, 1, \dots, \nu - 1)$$

Veamos que si suponemos

$$l_1, \dots, l_m$$

están uniformemente acotados por constantes independientes de m , y existen constantes positivas $\bar{\nu}$, \underline{b} , \bar{b} tales que

$$\underline{b} \leq b_0, b_1, \dots, b_\nu \leq \bar{b}, \quad \nu \leq \bar{\nu}$$

entonces se verifica H_2). Veamos que con esas condiciones

$$D_i - (1+u)q_i(u) \leq \frac{E_i}{1+u}.$$

Para eso basta ver que

$$(1+u) \sum_{k=0}^{\nu-1} a_k u^k \leq E_i Q(u), \quad (5.30)$$

donde el miembro izquierdo de (5.30) es el polinomio

$$\sum_{k=0}^{\nu-1} (a_k + a_{k+1}u^k + a_{\nu-1}u^\nu).$$

Basta elegir E_i ($i = 1, \dots, m$) tal que

$$\frac{a_k + a_{k+1}}{b_k} \leq E_i \quad \frac{a_{\nu-1}}{b_\nu} \leq E_i \quad (i = 1, \dots, m).$$

Pero eso es sencillo pues para $k = 0, \dots, \nu - 1$

$$\begin{aligned} a_k &= l_i [(\nu - k)b_k - (k + 1)b_{k+1}] \\ &\leq l_i [\bar{\nu}b - \underline{b}]. \end{aligned}$$

El resto de las condiciones de H_2) se prueban fácilmente . Con respecto a las hipótesis H_3) y H_4) para las señales (i) y (ii) del primer ejemplo se prueban de manera análoga a los ejemplos anteriores.

Observar que si tomamos en el caso $\nu = 2$ el polinomio

$$Q(u) = 1 + 2a u + b u^2$$

con las condiciones $0 < a \leq 1, \sqrt{b} > a \geq b > 0$, se tiene que caemos en las hipótesis de este ejemplo, sin embargo este polinomio no tiene raíces reales, y por tanto esta nueva clase de ejemplos es diferente de la anterior.

5.4. Complementos Técnicos

Para empezar necesitamos un pequeño lema elemental:

Lema 7. Sea $\xi \sim \mathcal{N}(0, I_k)$ y α, β en \mathbb{R}^k vectores, tales que $\|\alpha\| \leq \|\beta\|$, donde $\|\cdot\|$ indica la norma Euclideana en \mathbb{R}^k . Entonces

$$E(\|\xi + \alpha\|) \leq E(\|\xi + \beta\|).$$

Demostración. Usando la invarianza por isometrías de la distribución Gaussiana, basta probar que

$$E\left(\underbrace{[(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_k^2]^{1/2}}_{\varphi(a)}\right) \leq E\left(\underbrace{[(\xi_1 + b)^2 + \xi_2^2 + \dots + \xi_k^2]^{1/2}}_{\varphi(b)}\right)$$

donde

$$\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix}$$

y $a = \|\alpha\|, b = \|\beta\|$.

Es evidente que basta demostrar el lema para el caso $k = 1$, es decir que

$$\varphi(a) \leq \varphi(b) \quad \text{donde} \quad \varphi(x) := E(|\xi_1 + x|).$$

Para eso basta mostrar que

$$\int_{-\infty}^{+\infty} |x| e^{-(x-a)^2/2} dx \leq \int_{-\infty}^{+\infty} |x| e^{-(x-b)^2/2} dx.$$

Sea $A = (b + a)/2$, partiendo la integral en las semirrectas definidas por A se tiene

$$\begin{aligned} \varphi(b) - \varphi(a) &= \\ &= \int_{-\infty}^{+\infty} |x| e^{-(x-b)^2/2} dx - \int_{-\infty}^{+\infty} |x| e^{-(x-a)^2/2} dx \\ &= \int_{-\infty}^A |x| \left(e^{-(x-b)^2/2} - e^{-(x-a)^2/2} \right) dx + \int_A^{+\infty} |x| \left(e^{-(x-b)^2/2} - e^{-(x-a)^2/2} \right) dx \end{aligned}$$

Haciendo el cambio de variable $y = 2A - x$ en la primer integral se tiene

$$\begin{aligned} \varphi(b) - \varphi(a) &= \\ &= \int_A^{+\infty} |y - 2A| \left(e^{-(y-a)^2/2} - e^{-(y-b)^2/2} \right) dy + \int_A^{+\infty} |x| \left(e^{-(x-b)^2/2} - e^{-(x-a)^2/2} \right) dx \\ &= \int_A^{+\infty} \underbrace{(|x - 2A| - |x|)}_{\geq 0} \underbrace{\left(e^{-(x-a)^2/2} - e^{-(x-b)^2/2} \right)}_{\geq 0} dx \geq 0 \end{aligned}$$

donde el integrando es mayor que cero. \square

Afirmación 2. Sean ψ_1, \dots, ψ_m vectores aleatorios en \mathbb{R}^m independientes con distribución común Gaussiana estándar, y μ_1, \dots, μ_m vectores en \mathbb{R}^m . Definimos

$$\eta_j := \psi_j + \mu_j \quad \text{para } j = 1, \dots, m.$$

Entonces

$$E(|\det(\eta_1, \dots, \eta_m)|) \leq \prod_{j=1}^m E(\|\xi_j + c_j\|_j) \quad (5.31)$$

donde $c_j \in \mathbb{R}^j$, con la propiedad $\|c_j\|_j = \|\mu_j\|$ (ver notación en página 67), y ξ_j es un vector con distribución normal estándar en \mathbb{R}^j para $j = 1, \dots, m$.

Demostración. Al igual que en el caso centrado consideremos el subespacio lineal $L(\eta_1, \dots, \eta_{m-1})$ formado por los vectores $\eta_1, \dots, \eta_{m-1}$. Sean $\{v_1, \dots, v_m\}$ una base ortonormal (aleatoria) de \mathbb{R}^m tal que $\{v_1, \dots, v_{m-1}\}$ sea una base ortonormal de $L(\eta_1, \dots, \eta_{m-1})$. Como ya vimos podemos construir dicha base de forma tal que cada elemento sea medible respecto a la σ -álgebra engendrada por $\eta_1, \dots, \eta_{m-1}$ (que notamos por $\sigma(\eta_1, \dots, \eta_{m-1})$).

$$\begin{aligned} E(|\det T|) &= E(\text{vol}(\eta_1, \dots, \eta_m)) \\ &= E\left(d(\eta_m, L(\eta_1, \dots, \eta_{m-1})) \text{vol}^{(m-1)}(\eta_1, \dots, \eta_{m-1})\right) \\ &= E\left(E\left[d(\eta_m, L(\eta_1, \dots, \eta_{m-1})) \text{vol}^{(m-1)}(\eta_1, \dots, \eta_{m-1}) \mid \sigma(\psi_1, \dots, \psi_{m-1})\right]\right) \\ &= E\left(\text{vol}^{(m-1)}(\eta_1, \dots, \eta_{m-1}) E\left[d(\eta_m, L(\eta_1, \dots, \eta_{m-1})) \mid \sigma(\psi_1, \dots, \psi_{m-1})\right]\right) \end{aligned}$$

Escribiendo $\eta_m = \sum_{i=1}^m a_i v_i$ donde

$$a_m = \langle \eta_m, v_m \rangle = \langle \psi_m, v_m \rangle + \langle \mu_m, v_m \rangle.$$

Observar que $\gamma_m := \langle \psi_m, v_m \rangle$ tiene distribución normal estándar debido a la independencia entre ψ_m y v_m . Entonces se tiene

$$d(\eta_m, L(\eta_1, \dots, \eta_{m-1})) = |\gamma_m + \langle \mu_m, v_m \rangle|.$$

Sea $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ Borel medible, dada por

$$\varphi(x) = E(|\gamma_m + x|).$$

Es fácil ver que

$$E(|\gamma_m + \langle \mu_m, v_m \rangle| | \sigma(\psi_1, \dots, \psi_{m-1})) = \varphi(\langle \mu_m, v_m \rangle).$$

Por lo tanto, usando el *Lema 5.4* se tiene

$$E(|\gamma_m + \langle \mu_m, v_m \rangle| | \sigma(\psi_1, \dots, \psi_{m-1})) \leq E(|\gamma_m + \|\mu_m\||),$$

(ya que $\langle \mu_m, v_m \rangle \leq \|\mu_m\|$). Entonces

$$E(|\det(\eta_1, \dots, \eta_m)|) \leq E\left(\text{vol}^{(m-1)}(\eta_1, \dots, \eta_{m-2})\right) \cdot E(|\xi_1 + c_1|)$$

donde $\xi_1 \sim \mathcal{N}(0, 1)$ y $c_1 = \|\mu_m\|$.

Análogamente sea $\{v_1, \dots, v_{m-2}\}$ base ortonormal de $L(\eta_1, \dots, \eta_{m-2})$ y la completamos con v_{m-1} y v_m a una base de \mathbb{R}^m igual que antes donde cada vector es medible respecto a $\psi_1, \dots, \psi_{m-2}$. De esta manera tenemos

$$E\left(\text{vol}^{(m-2)}(\eta_1, \dots, \eta_{m-1})\right) \leq E\left(\text{vol}^{(m-2)}(\eta_1, \dots, \eta_{m-2})\right) \cdot E(\|\xi_2 + c_2\|_2)$$

donde $\xi_2 \sim \mathcal{N}(0, I_2)$ y c_2 es un vector en \mathbb{R}^2 de norma $\|\mu_{m-1}\|$. Continuando de esta manera obtenemos

$$E(|\det(\eta_1, \dots, \eta_m)|) \leq \prod_{j=1}^m E(\|\xi_j + c_j\|_j) \quad (5.32)$$

donde $\xi_j \sim \mathcal{N}(0, I_j)$, c_j es un vector en \mathbb{R}^j de norma $\|\mu_j\|$ para $j = 1, \dots, m$. \square

Lema 8. Sea $\gamma : \mathbb{R}^j \rightarrow \mathbb{R}$, $j \geq 1$ definido como

$$\gamma(c) = E(\|\xi + c\|)$$

donde ξ es un vector Gaussiano centrado en \mathbb{R}^j , y $c \in \mathbb{R}^j$ ($\|\cdot\|$ es la norma Euclídeana en \mathbb{R}^j). Entonces

$$(i) \quad \gamma(0) = \sqrt{2} \frac{\Gamma(\frac{j+1}{2})}{\Gamma(\frac{j}{2})}.$$

(ii) γ es una función de $\|c\|$ y verifica:

$$\gamma(c) \leq \gamma(0) \left(1 + \frac{1}{2j} \|c\|^2\right). \quad (5.33)$$

Demostración. Por la invarianza por isometrías de \mathbb{R}^j se tiene que

$$E(\|\xi + \alpha\|) = E([\xi_1 + a]^2 + \xi_2^2 + \cdots + \xi_j^2)^{1/2}$$

donde $a = \|\alpha\|$.

Sea

$$\gamma(a) = E([\xi_1 + a]^2 + \xi_2^2 + \cdots + \xi_j^2)^{1/2}.$$

Entonces

$$\gamma(0) = E(\|\xi\|) = \frac{\sigma_{j-1}}{(2\pi)^{j/2}} \int_0^{+\infty} \rho^j e^{-1/2\rho^2} d\rho = \frac{\sigma_{j-1}}{(2\pi)^{j/2}} I_j$$

donde

$$I_j = \int_0^{+\infty} \rho^j e^{-\rho^2/2} d\rho,$$

y σ_{j-1} es la medida geométrica de la esfera $j - 1$ -dimensional.

Haciendo el cambio de variable $u = \rho^2/2$ se tiene

$$I_j = 2^{\frac{j-1}{2}} \int_0^{+\infty} u^{\frac{j-1}{2}} e^{-u} du = 2^{\frac{j-1}{2}} \Gamma\left(\frac{j+1}{2}\right)$$

Entonces como $\sigma_{j-1} = 2\pi^{j/2}/\Gamma(j/2)$ se tiene

$$\gamma(0) = E(\|\xi\|) = \sqrt{2} \frac{\Gamma\left(\frac{j+1}{2}\right)}{\Gamma\left(\frac{j}{2}\right)}$$

Veamos que

$$\gamma'(a) = E\left(\frac{\xi_1 + a}{[\xi_1 + a]^2 + \xi_2^2 + \cdots + \xi_j^2}^{1/2}\right)$$

donde el integrando esta acotado en módulo por 1 y por tanto podemos derivar formalmente bajo el signo de integral.

Es claro que $\gamma'(0) = 0$ como consecuencia de la simetría de ξ_1 .

Veamos que podemos derivar nuevamente $\gamma'(a)$ formalmente bajo el signo de integral obteniendo

$$\gamma''(a) = E\left(\frac{\xi_2^2 + \cdots + \xi_j^2}{[(\xi_1 + a)^2 + \xi_2^2 + \cdots + \xi_j^2]^{3/2}}\right) \quad (5.34)$$

con $j \geq 3$.³ Para eso observemos que

$$E\left(\frac{\xi_2^2 + \cdots + \xi_j^2}{[(\xi_1 + a)^2 + \xi_2^2 + \cdots + \xi_j^2]^{3/2}}\right) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-x_1^2/2} dx_1 \frac{\sigma_{j-2}}{(2\pi)^{(j-1)/2}} \int_0^{+\infty} \frac{\rho^j}{[(x_1 + a)^2 + \rho^2]^{3/2}} e^{-\rho^2/2} d\rho.$$

³Más adelante veremos que también vale para $j = 2$.

Haciendo el cambio de variable ($\rho = |x_1 + a|t$) obtenemos que :

$$\int_0^{+\infty} \frac{\rho^j}{[(x_1 + a)^2 + \rho^2]^{3/2}} e^{-1/2\rho^2} = |x_1 + a|^{j-2} \int_0^{+\infty} \frac{t^j}{(1+t^2)^{3/2}} e^{-(x_1+a)^2 t^2/2} dt$$

y entonces obtenemos que

$$E \left(\frac{\xi_2^2 + \dots + \xi_j^2}{[(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_j^2]^{3/2}} \right) = \frac{\sigma_{j-2}}{(2\pi)^{j/2}} \int_{-\infty}^{+\infty} |x_1 + a|^{j-2} \alpha(x_1, a) e^{-x_1^2/2} dx_1$$

donde

$$\alpha(x, a) = \int_0^{+\infty} \frac{t^j}{(1+t^2)^{3/2}} e^{-(x+a)^2 t^2/2} dt.$$

De lo anterior se observa que podemos derivar bajo el signo de integral para $j \geq 3$ y

$$\gamma''(a) = \frac{\sigma_{j-2}}{(2\pi)^{j/2}} \int_{-\infty}^{+\infty} |x_1 + a|^{j-2} \alpha(x_1, a) e^{-x_1^2/2} dx_1.$$

Estudiemos por separado los casos $j = 1$ y $j = 2$.

Caso $j = 1$:

Para este caso tenemos $\gamma_1(a) = E(|\xi_1 + a|)$ y

$$\gamma_1'(a) = E(\text{sg}(\xi_1 + a)).$$

Entonces

$$\begin{aligned} \gamma_1'(a) &= E(\text{sg}(\xi_1 + a)) \\ &= P(\xi_1 > -a) - P(\xi_1 < -a) \\ &= \frac{1}{\sqrt{2\pi}} \left[\int_{-a}^{+\infty} e^{-x^2/2} dx - \int_{-\infty}^a e^{-x^2/2} dx \right] \end{aligned}$$

y por lo tanto

$$\gamma_1''(a) = \sqrt{\frac{2}{\pi}} e^{-a^2/2}. \quad (5.35)$$

Caso $j = 2$:

Para este caso tenemos $\gamma_2(a) = E([\xi_1 + a]^2 + \xi_2^2)^{1/2}$ y

$$\gamma_2'(a) = E \left(\frac{\xi_1 + a}{[(\xi_1 + a)^2 + \xi_2^2]^{1/2}} \right)$$

Veamos que

$$\gamma_2''(a) = E \left(\frac{\xi_2^2}{[(\xi_1 + a)^2 + \xi_2^2]^{3/2}} \right),$$

y por lo tanto la igualdad (5.34) vale también para $j = 2$.

Tenemos que

$$E \left(\frac{\xi_2^2}{[(\xi_1 + a)^2 + \xi_2^2]^{3/2}} \right) = \frac{2}{2\pi} \int_0^{+\infty} e^{-\rho^2/2} d\rho \underbrace{\int_{-\infty}^{+\infty} e^{-x_1^2/2} \frac{\rho^2}{[(x_1 + a)^2 + \rho^2]^{3/2}} dx_1}_{F(\rho)}$$

es decir

$$\frac{1}{\pi} \int_0^{+\infty} e^{-\rho^2/2} F(\rho) d\rho \quad (5.36)$$

donde $F(\rho) = \int_{-\infty}^{+\infty} e^{-x_1^2/2} \frac{\rho^2}{[(x_1 + a)^2 + \rho^2]^{3/2}} dx_1$.

Es claro que $F(\rho) \leq \frac{\sqrt{2\pi}}{\rho}$, entonces no hay problema cuando integramos en (5.36) en el conjunto $\rho \geq \gamma > 0$. El problema está cuando ρ es próximo a cero y $x_1 = -a$. Pero veremos que $F(\rho)$ permanece acotada en un entorno de 0. Para eso estudiemos el comportamiento de

$$\int_{-a}^A \frac{\rho^2}{[(x_1 + a)^2 + \rho^2]^{3/2}} dx_1 = \int_0^{\tilde{A}} \frac{\rho^2}{[x^2 + \rho^2]^{3/2}} dx = \frac{1}{\rho} \int_0^{\tilde{A}} \frac{1}{[(x/\rho)^2 + 1]^{3/2}} dx$$

donde A es cierto valor distinto de $-a$ y $\tilde{A} = A + a$.

Entonces haciendo el cambio de variable $u = x/\rho$ obtenemos que

$$\frac{1}{\rho} \int_0^{\tilde{A}} \frac{1}{[(x/\rho)^2 + 1]^{3/2}} dx = \int_0^{\tilde{A}/\rho} \frac{1}{(u^2 + 1)^{3/2}} du < \int_0^{+\infty} \frac{1}{(u^2 + 1)^{3/2}} du < +\infty$$

independiente de ρ . Por lo tanto estamos en condiciones de derivar bajo el signo de integral.

Ahora trataremos de acotar $\gamma''(0)$ y luego veremos que esto nos da una cota para $\gamma''(a)$ para todo $a \in \mathbb{R}$. Primero observemos que

$$\gamma''(0) = \left(\frac{j-1}{j} \right) \cdot E \left(\frac{1}{\|\xi\|} \right).$$

Esto resulta de lo siguiente:

$$\begin{aligned} \gamma''(0) &= E \left(\frac{\xi_2^2 + \dots + \xi_j^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} \right) \\ &= E \left(\frac{\xi_1^2 + \dots + \xi_j^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} - \frac{\xi_1^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} \right) \\ &= E \left(\frac{1}{\|\xi\|} \right) - \underbrace{E \left(\frac{\xi_1^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} \right)}_{(I)}. \end{aligned}$$

Pero por simetría se tiene que

$$(I) = E \left(\frac{\xi_l^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} \right) \quad (l = 1, \dots, j)$$

entonces

$$(I) = \frac{1}{j} E \left(\frac{\xi_1^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} + \dots + \frac{\xi_j^2}{(\xi_1^2 + \dots + \xi_j^2)^{3/2}} \right) = \frac{1}{j} E \left(\frac{1}{\|\xi\|} \right),$$

concluyendo la observación.

Computemos $E \left(\frac{1}{\|\xi\|} \right)$:

$$\begin{aligned} E \left(\frac{1}{\|\xi\|} \right) &= \frac{\sigma_{j-1}}{(2\pi)^{j/2}} \int_0^{+\infty} \rho^{j-2} e^{-\rho^2/2} d\rho \\ &= \frac{\sigma_{j-1}}{(2\pi)^{j/2}} \cdot I_{j-2} \\ &= \left(\frac{2\pi^{j/2}}{\Gamma(j/2)} \right) \frac{1}{(2\pi)^{j/2}} 2^{\frac{j-3}{2}} \Gamma \left(\frac{j-1}{2} \right) \\ &= \frac{1}{\sqrt{2}} \frac{\Gamma \left(\frac{j-1}{2} \right)}{\Gamma \left(\frac{j}{2} \right)}. \end{aligned}$$

Entonces se tiene

$$\gamma''(0) = \left(\frac{j-1}{j} \right) \cdot \frac{1}{\sqrt{2}} \frac{\Gamma \left(\frac{j-1}{2} \right)}{\Gamma \left(\frac{j}{2} \right)}$$

entonces

$$\begin{aligned} \frac{\gamma''(0)}{\gamma(0)} &= \left(\frac{j-1}{j} \right) \cdot \frac{1}{\sqrt{2}} \frac{\Gamma \left(\frac{j-1}{2} \right)}{\Gamma \left(\frac{j}{2} \right)} \frac{\Gamma \left(\frac{j}{2} \right)}{\sqrt{2} \Gamma \left(\frac{j+1}{2} \right)} \\ &= \left(\frac{j-1}{j} \right) \cdot \frac{1}{2} \frac{\Gamma \left(\frac{j-1}{2} \right)}{\Gamma \left(\frac{j+1}{2} \right)} \\ &= \left(\frac{j-1}{j} \right) \cdot \frac{1}{2} \frac{\Gamma \left(\frac{j-1}{2} \right)}{\left(\frac{j-1}{2} \right) \Gamma \left(\frac{j-1}{2} \right)} = \frac{1}{j}. \end{aligned}$$

Hemos probado que

$$\gamma''(0) = \frac{1}{j} \gamma(0) \tag{5.37}$$

Veamos que $|\gamma''(a)| \leq \gamma''(0)$:

$$\gamma''(a) = E \left(\frac{\xi_2^2 + \dots + \xi_j^2}{\left[(\xi_1 + a)^2 + \xi_2^2 + \dots + \xi_j^2 \right]^{3/2}} \right)$$

entonces condicionando en ξ_2, \dots, ξ_m se obtiene por la independencia que basta estudiar el comportamiento de la función (aleatoria)

$$\varphi(a) = E \left(\frac{c}{\left[(\xi_1 + a)^2 + c \right]^{3/2}} \right)$$

donde $c \geq 0$.⁴

Es claro que $\lim_{|a| \rightarrow +\infty} \varphi(a) = 0$ y que $\varphi(a) \geq 0$ para todo $a \in \mathbb{R}$.

Veamos que φ es decreciente cuando $a \geq 0$ y creciente si $a \leq 0$. Para esto basta observar que

$$\varphi'(a) = -3c E \left(\frac{\xi_1 + a}{\left[(\xi_1 + a)^2 + c \right]^{5/2}} \right)$$

donde podemos derivar formalmente bajo el signo. Considerando $a \geq 0$ obtenemos

$$\begin{aligned} \varphi'(a) &= -3c \int_{-\infty}^{+\infty} \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx \\ &= -3c \left\{ \int_{-\infty}^0 \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx + \int_0^{+\infty} \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx \right\} \\ &= -3c \left\{ - \int_0^{+\infty} \frac{x}{(x^2 + c)^{5/2}} \underbrace{e^{-(x+a)^2/2}}_{\leq e^{-(x-a)^2/2}} dx + \int_0^{+\infty} \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx \right\} \\ &\leq -3c \left\{ - \int_0^{+\infty} \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx + \int_0^{+\infty} \frac{x}{(x^2 + c)^{5/2}} e^{-(x-a)^2/2} dx \right\} \\ &= 0 \end{aligned}$$

es decir, la función $\varphi(a)$ es decreciente en la semirrecta $[0, +\infty)$ y análogamente se prueba que es creciente en la semirrecta $(-\infty, 0]$.

De esta manera la función φ tiene un único máximo y se da en 0 probando lo que queríamos.

Entonces usando el desarrollo de Taylor en 0 de γ obtenemos:

$$\gamma(a) \leq \gamma(0) + \frac{1}{2} a^2 \gamma''(0)$$

donde se tiene usando (5.37)

$$\gamma''(0) = \frac{1}{j} \gamma(0) \quad \text{y} \quad \gamma(0) = \sqrt{2} \frac{\Gamma(\frac{j+1}{2})}{\Gamma(\frac{j}{2})}$$

para $j \geq 1$.⁵

□

⁴En realidad depende de la aleatoriedad de c , pero aplicando como antes una regresión Gaussiana podemos suponer que c esta fijo, y trabajar sin problemas.

⁵Observar que para el caso $j = 1$, es inmediato por (5.35).

Capítulo 6

Apéndice

6.1. Teorema Principal de Teoría de la Eliminación

En esta sección del apéndice, queremos mostrar bajo que condiciones una variedad algebraica dada en un producto cartesiano, se proyecta en alguna de las fibras en una nueva variedad algebraica. Más precisamente, en la notación del *Teorema de Bézout* queremos ver si $\Sigma := \pi_1(\Sigma') \subset \mathbb{P}(\mathcal{H}_{(d)})$ es una variedad algebraica, donde $\pi_1 : \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1}) \rightarrow \mathbb{P}(\mathcal{H}_{(d)})$ es la proyección en la primera fibra, y Σ' es el conjunto de pares $(f, x) \in \mathbb{P}(\mathcal{H}_{(d)}) \times \mathbb{P}(\mathbb{C}^{n+1})$ tales que f en x tiene un cero degenerado.

Antes de enunciar y demostrar el teorema, necesitaremos algunos pre-requisitos y resultados previos.

Dado \mathbb{K} un cuerpo, notaremos por $\mathbb{K}[x_1, \dots, x_n]$ al anillo de polinomios en las variables x_1, \dots, x_n , con coeficientes en \mathbb{K} . Decimos que $Z \subset \mathbb{K}^n$ es un *variedad algebraica* si es el conjunto de ceros comunes de una familia finita de polinomios f_1, \dots, f_s en $\mathbb{K}[x_1, \dots, x_n]$ (al cual notaremos por $\mathcal{Z}(f_1, \dots, f_s)$).¹ Observar que si consideramos el ideal $I = (f_1, \dots, f_s)$ generado por dichos polinomios, entonces $\mathcal{Z}(I) = \mathcal{Z}(f_1, \dots, f_s)$.

Vale la pena resaltar que en la definición de variedad algebraica en el caso de que \mathbb{K} sea algebraicamente cerrado, no es una restricción imponer que los polinomios que la definen sean una cantidad finita. Eso resulta del siguiente teorema debido a Hilbert, comúnmente llamado *Teorema de la Base de Hilbert*:

Teorema 12. *Si \mathbb{K} es algebraicamente cerrado, entonces todo ideal de $\mathbb{K}[x_1, \dots, x_n]$ es finitamente generado.*

En otras palabras, toda variedad algebraica definida sobre \mathbb{K}^n con \mathbb{K} un cuerpo algebraicamente cerrado es una intersección finita de hiperplanos (es decir $f(x_1, \dots, x_n) = 0$ con f un polinomio).

Se puede chequear fácilmente, que el conjunto de variedades algebraicas sobre \mathbb{K}^n satisface los axiomas de conjuntos cerrados para una topología. A través de esto, se puede definir una topología en \mathbb{K}^n , donde los abiertos son los complementos de las variedades algebraicas. Dicha topología es conocida como *topología de Zariski*.

El siguiente resultado es un corolario del conocido teorema de Hilbert *Nullstellensatz*, que incluiremos sin demostración:

¹En la página 25 hay una pequeña digresión sobre variedades algebraicas que complementa este apéndice.

Teorema 13. Sea \mathbb{K} un cuerpo algebraicamente cerrado, e I un ideal de $\mathbb{K}[x_1, \dots, x_n]$. Entonces $\mathcal{Z}(I) = \emptyset$ si y solo si $1 \in I$. Equivalentemente, f_1, \dots, f_m no tienen un cero en común, si y solo si, existen $g_1, \dots, g_m \in \mathbb{K}[x_1, \dots, x_n]$ tales que $\sum_1^m g_i f_i = 1$.

Teorema 14. Sea \mathbb{K} un cuerpo algebraicamente cerrado, e I el ideal generado por f_1, \dots, f_s polinomios homogéneos en $\mathbb{K}[x_0, \dots, x_n]$. Entonces $\mathcal{Z}(I) = \{0\}$ si y solo si existe $l \in \mathbb{N}$ tal que $x_i^l \in I$ para todo $i = 0, \dots, n$.

Demostración. La condición suficiente es evidente. Probemos la condición necesaria. Dado $f \in \mathbb{K}[x_0, \dots, x_n]$ homogéneo, notaremos por $\tilde{f} \in \mathbb{K}[x_1, \dots, x_n]$ al polinomio $\tilde{f}(x_1, \dots, x_n) := f(1, x_1, \dots, x_n)$. Y en el sentido contrario, dado $g \in \mathbb{K}[x_1, \dots, x_n]$, notaremos por \hat{g} a su homogeneizado ($\hat{g}(x_0, \dots, x_n) := x_0^l g(x_1/x_0, \dots, x_n/x_0)$) siendo l el grado de g .

Sean $f_1, \dots, f_s \in \mathbb{K}[x_0, \dots, x_n]$ polinomios homogéneos tales que $I = (f_1, \dots, f_s)$. Supongamos que $\mathcal{Z}(I) = \{0\}$. Consideremos $\tilde{f}_1, \dots, \tilde{f}_s$ y el ideal $\tilde{I} = (\tilde{f}_1, \dots, \tilde{f}_s)$. Entonces $\mathcal{Z}(\tilde{I}) = \emptyset$, y usando el Teorema 13 obtenemos que existen

$g_1, \dots, g_s \in \mathbb{K}[x_1, \dots, x_n]$ tales que: $1 = \sum_1^m g_i \tilde{f}_i$. Reemplazando la variable x_i por x_i/x_0 , obtenemos

$$1 = \sum_{i=1}^s g_i \left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) \tilde{f}_i \left(\frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right),$$

de donde resulta

$$\begin{aligned} x_0^{l_1 + \dots + l_s} &= x_0^{l_2 + \dots + l_s} \hat{g}_1(x_0, \dots, x_n) f_1 \left(1, \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right) + \dots \\ &\quad \dots + x_0^{l_1 + \dots + l_{s-1}} \hat{g}_n(x_0, \dots, x_n) f_n \left(1, \frac{x_1}{x_0}, \dots, \frac{x_n}{x_0} \right), \end{aligned}$$

siendo l_i el grado del polinomio g_i para $i = 1, \dots, s$. Entonces tenemos que $x_0^l \in I$ donde l es la suma de los grados de los g_i más los grados de los f_i , para $i = 1, \dots, s$. Análogamente, obtenemos que $x_i^l \in I$. \square

Corolario 1. Sea \mathbb{K} un cuerpo algebraicamente cerrado, e $I = (f_1, \dots, f_s)$ donde f_1, \dots, f_s son polinomios homogéneos en $\mathbb{K}[x_0, \dots, x_n]$. Entonces $\mathcal{Z}(I) = \{0\}$ si y solo si existe $d \in \mathbb{N}$ tal que $(x_0, \dots, x_n)^d \subset I$.²

Demostración. Basta tomar $d = (n+1)l$ donde l está dado en el teorema anterior. \square

Teorema 15 (Teorema Principal de Teoría de la Eliminación). La proyección $\pi_1 : \mathbb{P}(\mathbb{C}^{m+1}) \times \mathbb{P}(\mathbb{C}^{n+1}) \rightarrow \mathbb{P}(\mathbb{C}^{m+1})$ es Zariski cerrada. Es decir, si $\mathcal{Z} \subset \mathbb{P}(\mathbb{C}^{m+1}) \times \mathbb{P}(\mathbb{C}^{n+1})$ es una variedad algebraica proyectiva, entonces $\pi_1(\mathcal{Z})$ también lo es.

Demostración. En realidad probaremos un resultado en principio más débil:

si consideramos la identificación de \mathbb{C}^m en $\mathbb{P}(\mathbb{C}^{m+1})$ dada por $(x_1, \dots, x_m) \rightarrow (1, x_1, \dots, x_m)$, entonces la proyección $\pi_1 : \mathbb{C}^m \times \mathbb{P}(\mathbb{C}^{n+1}) \rightarrow \mathbb{C}^m$ es cerrada con la topología de Zariski. Pero como el resultado es local en la imagen, dicha restricción alcanza para probar el teorema. Observar que con esta identificación, \mathbb{C}^m es abierto en $\mathbb{P}(\mathbb{C}^{m+1})$ con la topología de Zariski.

²Si S y T son ideales sobre un anillo conmutativo, definimos ST como el conjunto $\{\sum_1^n s_j t_j : s_j \in S, t_j \in T\}$.

En estas condiciones, sea Z es un cerrado en $\mathbb{C}^m \times \mathbb{P}(\mathbb{C}^{n+1})$, entonces existen polinomios $f_1, \dots, f_s \in \mathbb{C}[y_1, \dots, y_m, x_0, \dots, x_n]$ homogéneos respecto a las variables x_0, \dots, x_m . Sea d_i el grado de f_i para $i = 1, \dots, s$ respecto a las variables x_0, \dots, x_m que definen Z . Entonces para todo $a \in \mathbb{C}^m$ se tiene $a \notin \pi_1(Z)$ si y solo si $f_i(a, x)$ no tienen ceros comunes en $\mathbb{P}(\mathbb{C}^{n+1})$ para $i = 1, \dots, s$, (i.e. $\mathcal{Z}(f_1(a, \cdot), \dots, f_s(a, \cdot)) = \emptyset$). Ahora usando el corolario del *Teorema 14* se tiene que existe $d \in \mathbb{N}$ tal que $(x_0, \dots, x_n)^d \subset I_a := (f_1(a, \cdot), \dots, f_s(a, \cdot))$. Concluimos entonces que

$$\pi_1(Z)^c = \cup_{d \in \mathbb{N}} A_d, \quad \text{donde} \quad A_d := \{a \in \mathbb{C}^m : (x_0, \dots, x_n)^d \subset I_a\}.$$

Por lo tanto, basta probar que A_d es Zariski abierto en \mathbb{C}^m para todo $d \in \mathbb{N}$.

Sea \mathcal{H}_k el espacio vectorial de polinomios homogéneos de grado k en las variables x_0, \dots, x_n . Sea $d \geq 1$ y consideremos para cada $a \in \mathbb{C}^m$ la aplicación lineal

$$T^d(a) : \mathcal{H}_{d-d_1} \oplus \dots \oplus \mathcal{H}_{d-d_s} \rightarrow \mathcal{H}_d$$

dada por

$$T^d(a)(g_1, \dots, g_s) = \sum_{i=1}^s g_i(x) f_i(a, x).$$

(Estamos asumiendo que $\mathcal{H}_k = 0$ cuando $k < 0$). Fijando las bases de los monomios en los respectivos espacios obtenemos que la matriz $n_d \times m_d$ asociada $[T^d(a)]$ de $T^d(a)$ tiene como coeficientes, polinomios en las variables a_1, \dots, a_m . Concluimos entonces que:

$(x_0, \dots, x_n)^d \subset (f_1(a, x), \dots, f_s(a, x))$ si y solo si $T^d(a)$ es sobreyectivo, y esto sucede si y solo si existe un menor complementario M de $[T^d(a)]$, de tamaño $m_d \times m_d$ tal que $\det M \neq 0$. Concluimos entonces que A_d para d fijo es un abierto en la topología de Zariski. □

6.2. Procesos Estocásticos

Un *proceso 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 (Ω, \mathcal{A}, P) , un p.e. es una función $X : \Omega \times T \rightarrow \mathbb{R}$ 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 $X(\cdot, \omega) : T \rightarrow \mathbb{R}$ tal que $t \mapsto X(t, \omega)$ 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 $X(\cdot, \omega) : T \rightarrow \mathbb{R}$ 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, \dots, t_m de índices el vector $(X(t_1), \dots, X(t_m))$ tiene distribución conjunta Gaussiana en \mathbb{R}^m .

En la siguiente sección consideraremos que el conjunto de parámetros es una variedad diferenciable compacta V de dimensión d en \mathbb{R}^d y el campo (aleatorio) $Z : V \rightarrow \mathbb{R}^d$ es Gaussiano.

6.3. El conjunto Σ tiene medida nula en $\mathbb{P}(\mathcal{H}_{(d)})$

La idea de esta sección es hacer una prueba autocontenida del siguiente hecho: los sistemas polinomiales con ceros degenerados tienen medida cero respecto a la medida inducida en $\mathbb{P}(\mathcal{H}_{(d)})$ por la métrica Riemanniana que proviene del producto interno de Weyl (ver nota al pie en página 31). Como vimos en la sección *Nuevo Enfoque*, basta probar que el conjunto de sistemas polinomiales $f \in \mathcal{H}_{(d)}$ con ceros degenerados tiene medida cero respecto a la medida Gaussiana inducida en los coeficientes. Para probar dicho problema, alcanza con probar el siguiente lema:

Lema 9. *Sea $X : \Omega \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ un proceso Gaussiano con $\text{Var}(X(t)) \succ 0$ para todo $t \in \mathbb{R}^d$, con trayectorias de clase por lo menos C^2 . Sea*

$$\Sigma := \{t \in \mathbb{R}^d : X(t) = 0, \det(D_t X) = 0\}.$$

Si le pedimos al proceso que la función de covarianza sea continua (como es en nuestro caso) entonces se tiene

$$P(\{\Sigma \neq \emptyset\}) = 0.$$

Demostración. Cubriendo \mathbb{R}^d con una cantidad numerable de cubos, es fácil ver que basta probar el lema para el caso $X : \Omega \times [0, 1]^d \rightarrow \mathbb{R}^d$. Sea $\Sigma_0 := \{t \in [0, 1]^d : X(t) = 0, \det(D_t X) = 0\}$.

Sea el evento

$$S_M = \left[\sup_{t \in [0, 1]^d} \left\{ \left| \frac{\partial X_k}{\partial t_j}(t) \right|, \left| \frac{\partial^2 X_k}{\partial t_i \partial t_j}(t) \right|, i, j, k = 1, \dots, d \right\} \geq M \right].$$

Resulta de la regularidad de las trayectorias que $P(S_{+\infty}) = 0$, y por lo tanto, dado $\eta > 0$ existe $M(\eta)$ suficientemente grande, tal que $P(S_{M(\eta)}) < \eta$.

Sean la partición $\{Q_j^{(n)}\}_{j=1, \dots, n^d}$ del cubo $[0, 1]^d$ en cubitos disjuntos de lado $1/n$. Análogamente partimos cada cara de los cubitos $Q_j^{(n)}$ en cubitos $\{D_{j,l}^{(n)}\}_{l=1, \dots, 2dn^{d-1}}$ de dimensión $d-1$ y de lado $1/n^2$ (observar que la cantidad de cubitos por cara es n^{d-1} , y un cubo d -dimensional tiene $2d$ caras).

Observar que el evento

$$\{\Sigma_0 \neq \emptyset\} = \bigcup_{j=1}^{n^d} \{\Sigma_0 \cap Q_j^{(n)} \neq \emptyset\},$$

y

$$\begin{aligned} \{\Sigma_0 \cap Q_j^{(n)} \neq \emptyset\} &= \{\exists \tau_j \in Q_j^{(n)} : X(\tau_j) = 0, \det(D_{\tau_j} X) = 0\} \\ &= \{\exists \tau_j \in Q_j^{(n)} \text{ y } 0 \neq v \in \mathbb{R}^d : X(\tau_j) = 0, D_{\tau_j} X(v) = 0\}. \end{aligned}$$

Fijemos un punto de referencia $\tau_{j,l}^*$ en cada cubito $D_{j,l}^{(n)}$ (que depende de n pero la omitimos para no cargar tanto la notación).

Fijemos $\omega \in \{\exists \tau_j \in Q_j^{(n)} \text{ y } 0 \neq v \in \mathbb{R}^d : X(\tau_j) = 0, D_{\tau_j}X(v) = 0\} \cap (S_{M(\eta)})^c$. Sea el punto $\tilde{\tau}_{j,l}$ que resulta de la intersección de la recta en la dirección de v que pasa por τ_j con la cara del cubo $Q_j^{(n)}$ (donde l indica que la intersección resulta en $D_{j,l}^{(n)}$). Luego para cada $k = 1, \dots, d$ obtenemos mediante el desarrollo de Taylor de segundo orden, centrado en τ_j , que

$$|X_k(\tilde{\tau}_{j,l})| \leq \frac{1}{2}M(\eta) d^2 \frac{1}{n^2}.$$

Entonces comparando con $\tau_{j,l}^*$ resulta

$$|X_k(\tau_{j,l}^*)| \leq \frac{1}{2}M(\eta) d^2 \frac{1}{n^2} + M(\eta)(d-1)\frac{1}{n^2} \leq K_0(d)\frac{1}{n^2}$$

donde $K_0(d)$ es una constante que solamente depende de la dimensión d . Entonces obtuvimos que

$$\{\Sigma_0 \cap Q_j^{(n)} \neq \emptyset\} \subset \bigcup_{l=1}^{2dn^{d-1}} \{\|X(\tau_{j,l}^*)\| \leq K(d)/n^2\}.$$

donde $K(d) = dK_0(d)$. Luego,

$$\begin{aligned} P(\{\Sigma_0 \neq \emptyset\}) &\leq \eta + \sum_{j=1}^{n^d} P(\{\Sigma_0 \cap Q_j^{(n)} \neq \emptyset\}) \\ &\leq \eta + \sum_{j=1}^{n^d} \sum_{l=1}^{2dn^{d-1}} P(\{\|X(\tau_{j,l}^*)\| \leq K(d)/n^2\}) \\ &\leq \eta + n^d 2dn^{d-1} C_{[0,1]^d} \lambda_d(B_d(0, K(d)/n^2)) \\ &\leq \eta + n^d 2dn^{d-1} C_{[0,1]^d} K \frac{1}{n^{2d}} \end{aligned}$$

donde $C_{[0,1]^d}$ es una constante que resulta de acotar la varianza del proceso por debajo en el cubo unidad (precisamente

$$C_{[0,1]^d} = \max_{t \in [0,1]^d} p_{X(t)}(0) = \max_{t \in [0,1]^d} \frac{1}{\sqrt{\det(E(X(t)X(t)^T))}},$$

que por ser no-degenerado resulta $C_{[0,1]^d} < +\infty$, y $\lambda_d(B_d(0, r))$ la medida de Lebesgue d -dimensional de la bola centrada en cero y de radio r . Entonces obtuvimos que

$$P(\{\Sigma_0 \neq \emptyset\}) \leq \eta + \frac{K}{n},$$

donde K es una constante independiente de n . Luego tomando limite en n tenemos que

$$P(\{\Sigma_0 \neq \emptyset\}) \leq \eta$$

y como el η era arbitrario, concluimos la demostración. \square

Observaciones:

- (i) El hecho de que X sea Gaussiano no juega un papel importante en la demostración del *Lema 9*.
(ii) La continuidad de la función de covarianza, solo la usamos para probar que la densidad del proceso $X(t)$, para cada $t \in [0, 1]^d$, esta acotada por arriba en el cubo unidad. Con estas observaciones en cuenta podemos extender casi sin cambios el resultado al siguiente lema:

Lema 10. *Sea $X : U \rightarrow \mathbb{R}^d$ con $U \subset \mathbb{R}^d$ compacto, un campo aleatorio con trayectorias C^2 , y $u \in \mathbb{R}^d$. Supongamos que para todo $t \in U$ y $x \in B_d(0, \delta)$ se tiene $p_{X(t)}(x) \leq C$. Entonces*

$$P(\{\exists t \in U, X(t) = u, \det(D_t X) = 0\}) = 0.$$

La demostración del *Lema 10* se puede encontrar en [11], y la demostración del *Lema 9* se basa en dicha referencia.

6.4. Fórmula de Rice, Fórmula del Área y Complementos

En esta sección enunciaremos la *fórmula de Rice* en su versión general, y agregaremos los resultados necesarios para completar la prueba, para el caso de sistemas polinomiales aleatorios (ver sección *Nuevo Enfoque*) en cuyo caso, resulta bastante más sencilla que en el caso general. La fórmula del área es extraída de [2].

Teorema 16. *Sea $Z : V \rightarrow \mathbb{R}^d$ con $V \subset \mathbb{R}^d$ abierto, un campo aleatorio, y $u \in \mathbb{R}^d$. Asumimos lo siguiente:*

- H0) Z es un campo vectorial Gaussiano,*
H1) c.s. el campo Z tiene trayectorias C^1 ,
H2) $\forall t \in V$, $Z(t)$ tiene distribución no-degenerada ($\text{var}(Z(t)) \succ 0$),
H3) $P\{\exists t \in V, Z(t) = u, \det(Z'(t)) = 0\} = 0$.

Entonces para todo conjunto de Borel $B \subset V$ se tiene

$$E(N_u^Z(B)) = \int_B E(|\det(Z'(t))| / Z(t) = u) \cdot p_{Z(t)}(u) dt. \quad (6.1)$$

Teorema 17 (Fórmula del Área). *Sea $f : U \rightarrow \mathbb{R}^d$ una función de clase C^1 definida sobre un abierto $U \subset \mathbb{R}^d$. Asumimos que el conjunto de valores críticos tiene medida de Lebesgue cero en \mathbb{R}^d . Sea $g : \mathbb{R}^d \rightarrow \mathbb{R}$ una función continua y acotada. Entonces se tiene*

$$\int_{\mathbb{R}^d} g(u) N_u^f(B) du = \int_B |\det(f'(t))| g(f(t)) dt \quad (6.2)$$

para todo $B \subset U$ de Borel, tal que las integrales de ambos miembros existen (pueden ser infinitas).

Demostración. Nos podemos reducir a estudiar el caso en que $B = I$ es un paralelepípedo de \mathbb{R}^d . Si $u \in \mathbb{R}^d$ es un valor regular de f , resulta de la compacidad de I que $N_u^f(I) < \infty$. Si $N_u := N_u^f(I) > 0$, sean $\tau_1, \dots, \tau_{N_u}$ las pre-imágenes por f de u . Tomando δ chico, podemos aislar cada una de las pre-imágenes $\tau_1, \dots, \tau_{N_u}$ con entornos abiertos disjuntos U_1, \dots, U_{N_u} respectivamente, de forma tal que:

- f sea un difeomorfismo entre U_i y la bola $B_d(u, \delta)$
- si $x \notin \cup_{i=1}^{N_u} U_i$ entonces $f(x) \notin B_d(u, \delta)$ donde $B_d(u, \delta)$ es la bola Euclideana en \mathbb{R}^d de centro u y radio δ .

Definamos la siguiente función auxiliar $\delta(u)$:

- si u es un valor regular y $N_u > 0$, entonces $\delta(u)$ es el supremo de los δ 's que verifican las condiciones anteriores.

- si $N_u = 0$, lo definimos como la distancia de u a $f(I)$.

- si u es un valor crítico, entonces $\delta(u) = 0$. No es difícil chequear que $\delta(u)$ es una función continua.

Se puede probar que es una función Lipschitz.

Entonces si $\delta < \delta(u)$

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

entonces se tiene

$$N_u^f(I) = \frac{1}{\lambda_d(B_d(u, \delta))} \int_I |\det(f'(t))| \chi_{\{\|f(t)-u\|<\delta\}} dt.$$

para todo $\delta < \delta(u)$.

Podemos asumir que $g : \mathbb{R}^d \rightarrow \mathbb{R}$ es continua no-negativa (el caso general resulta de separar en parte positiva y negativa). Entonces tenemos

$$g(u)N_u^f(I) = g(u) \frac{1}{\lambda_d(B_d(u, \delta))} \int_I |\det(f'(t))| \chi_{\{\|f(t)-u\|<\delta\}} dt \quad \forall \delta < \delta(u).$$

Queremos integrar respecto a u en todo \mathbb{R}^d la anterior igualdad, pero debemos ser cuidadosos pues la anterior igualdad vale para $\delta < \delta(u)$ y no uniforme en todo u . La forma de solucionar esto es agregar una función auxiliar, que tome el valor cero cuando δ sobrepase el valor $\delta(u)$. Con ese objetivo, sea $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ una función continua creciente arbitraria, tal que tome el valor 0 en $(-\infty, 1/2]$ y el valor 1 en $[1, +\infty)$. Observar que, fijando $\delta_0 > 0$ y tomando cualquier $\delta < \delta_0/2$ entonces vale

$$\int_{\mathbb{R}^d} g(u)N_u^f(I) \alpha\left(\frac{\delta(u)}{\delta_0}\right) du = \int_{\mathbb{R}^d} \alpha\left(\frac{\delta(u)}{\delta_0}\right) g(u) \frac{1}{\lambda_d(B_d(u, \delta))} \int_I |\det(f'(t))| \chi_{\{\|f(t)-u\|<\delta\}} dt du.$$

Luego aplicando Fubini

$$\int_{\mathbb{R}^d} g(u)N_u^f(I) \alpha\left(\frac{\delta(u)}{\delta_0}\right) du = \int_I |\det(f'(t))| \left(\frac{1}{\lambda_d(B_d(0, \delta))} \int_{B_d(f(t), \delta)} g(u) \alpha\left(\frac{\delta(u)}{\delta_0}\right) du \right) dt.$$

Como el miembro izquierdo no depende de δ , podemos pasar al límite en el miembro derecho para $\delta \downarrow 0$ y usar el teorema de convergencia dominada y la continuidad de $g(u)\alpha\left(\frac{\delta(u)}{\delta_0}\right)$ respecto a u . Entonces resulta

$$\int_{\mathbb{R}^d} g(u) N_u^f(I) \alpha\left(\frac{\delta(u)}{\delta_0}\right) du = \int_I |\det(f'(t))| g(f(t)) \alpha\left(\frac{\delta(f(t))}{\delta_0}\right) dt.$$

El resultado sigue de tomar $\delta_0 \downarrow 0$ y usar el teorema de convergencia monótona. Observar que en el miembro izquierdo podemos excluir los valores críticos de la función, y en el miembro derecho solamente hay que excluir los puntos del borde (pues en los puntos críticos el determinante es cero). □

El siguiente lema es utilizado para probar la formula de Rice:

Lema 11. *Sea I compacto en \mathbb{R}^d , $Z : I \rightarrow \mathbb{R}^d$ un campo vectorial aleatorio y $u \in \mathbb{R}^d$ tal que:*

1. *El mapa $t \mapsto Z(t)$ 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

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

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

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

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, \dots, B_m de radio δ_j con $j = 1 \dots m$ respectivamente tales que

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

$$P(N_u^Z(\partial I) \neq 0) \leq P\left(N_u^Z\left(\bigcup_{j=1}^m B_j \cap I\right) \geq 1\right) \tag{6.3}$$

$$\leq P\left(\left\{N_u^Z\left(\bigcup_{j=1}^m B_j \cap I\right) \geq 1\right\} \cap A_\varepsilon^c\right) + \underbrace{P(A_\varepsilon)}_{< \varepsilon} \tag{6.4}$$

$$< P\left(\left\{N_u^Z\left(\bigcup_{j=1}^m B_j \cap I\right) \geq 1\right\} \cap A_\varepsilon^c\right) + \varepsilon. \tag{6.5}$$

Estudiemos el primer miembro después de la desigualdad en (6.5).

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

Si $\omega \in \{N_u^Z(B_j \cap I) \geq 1\}$ 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 $\omega \in \{N_u^Z(B_j \cap I) \geq 1\} \cap A_\varepsilon^c$ se tiene

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

donde la primera desigualdad sigue del teorema de valor medio para campos vectoriales.

Tenemos que

$$\|Z(\hat{t}_j) - u\| \leq 2M_0 d \delta_j. \quad (6.6)$$

Concluimos que: si $\omega \in \{N_u^Z(B_j \cap I) \geq 1\} \cap A_\varepsilon^c$ entonces por (6.6) $\|Z(\hat{t}_j)(\omega) - u\| \leq 2M_0 d \delta_j$.

Juntando lo anterior se tiene

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

Entonces

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

Como $M_0 = M_0(\varepsilon)$ depende solamente 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. \square

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) están conectados. Pero esto no es un problema pues en la hipótesis cuando decimos que el mapa $t \mapsto Z(t)$ 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.

6.5. Fórmula de Co-Área

En esta sección demostraremos la fórmula de co-área, con hipótesis que alcanzan para probar el Teorema de Shub-Smale.

Sean M y N variedades Riemannianas de dimensiones m y n respectivamente, y sean dM y dN los elementos de volúmenes respectivos inducidos por la métrica. Supondremos $m \geq n$. Sea $f : M \rightarrow N$ una submersión.³ Entonces como todo punto en la imagen es un valor regular de f , sigue del teorema de la función implícita que $f^{-1}(q)$ es una sub-variedad diferenciable de M de dimensión $m - n$ (co-dimensión n), la cual hereda la métrica Riemanniana de M y por tanto tiene una medida de volumen inducida, que notaremos dM_q .

Sea ϕ una función definida en M a valores reales. Notaremos por ϕ_q a la restricción a la "fibra" $f^{-1}(q)$, y a $\bar{\phi}(q)$ al promedio en la fibra $f^{-1}(q)$, es decir

$$\bar{\phi}(q) := \int_{f^{-1}(q)} \phi_q dM_q.$$

(Notaremos la derivada normal de f por \mathcal{J}_f).

Teorema 18. *Sean M y N variedades Riemannianas de dimensiones m y n ($m \geq n$) respectivamente. Sea $f : M \rightarrow N$ una submersión, y $\phi : M \rightarrow \mathbb{R}$ una función continua con soporte compacto. Entonces valen las siguientes igualdades*

$$\int_M \mathcal{J}_f \cdot \phi dM = \int_N \bar{\phi} dN \left(= \int_{q \in N} \int_{p \in f^{-1}(q)} \phi_q(p) dM_q(p) dN(q) \right), \quad (6.7)$$

$$y \quad \int_M \phi dM = \int_{q \in N} \int_{p \in f^{-1}(q)} \frac{1}{(\mathcal{J}_f)_q(p)} \phi_q(p) dM_q dN. \quad (6.8)$$

Observaciones:

- Es posible extender por métodos estándares de análisis los resultados (6.7) y (6.8) a funciones $\phi : M \rightarrow \mathbb{R}$ integrables.
- Existe una versión más general de la fórmula de co-área para el caso en que: f sea una función Lipschitz, y, M, N sean conjuntos arbitrarios de espacios euclideos. En ese caso, es necesario utilizar la medida de Hausdorff para integrar en las fibras pues a priori las fibras ($f^{-1}(q)$) no tienen estructura diferenciable. Una buena referencia para estos temas es [13].
- La prueba que daremos, sigue las líneas de la prueba de la fórmula de co-área para el caso de *submersiones isométricas* (es decir, $\mathcal{J}_f(p) = 1$ para todo $p \in M$) extraída de [21].

Demostración. Probaremos (6.7), pues (6.8) es análogo, además eligiendo una ϕ adecuada, se puede transformar (6.7) en (6.8). Por la forma local de las submersiones, dado $p \in M$ y $q = f(p) \in N$ existen cartas locales (φ, U) y (ψ, V) de p y q respectivamente tales que:

- $f(U) = V$,
- $\varphi(U) = B_\varepsilon^n(0) \times B_\delta^{m-n}(0)$,
- $\psi(V) = B_\varepsilon^n(0)$,

³Una función entre variedades Riemannianas se dice una *submersión* si es diferenciable y el diferencial en todo punto es sobreyectivo.

- $\psi \circ f \circ \varphi^{-1}(x_1, \dots, x_n, x_{n+1}, \dots, x_m) = (x_1, \dots, x_n)$,

donde notamos por $B_\alpha^j(0)$ a la bola en \mathbb{R}^j de centro 0 y radio α . Es fácil ver que basta reducirnos al caso local y luego extenderlo a toda la variedad usando una partición de la unidad subordinada al cubrimiento generado por la anterior construcción en cada punto (ver final de la prueba).

Sea $\{\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_m}(p)\}$ la base natural de T_pM inducida por la carta para $p \in U$, y $\{\frac{\partial}{\partial x_1}(q), \dots, \frac{\partial}{\partial x_n}(q)\}$ la base natural de T_qN para $q \in V$. Observar que el núcleo de $D_p f$ esta generado por $\{\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_n}(p)\}$, y $D_p f(\frac{\partial}{\partial x_i}(p)) = \frac{\partial}{\partial x_i}(q)$ para $i = n+1, \dots, m$. Por lo tanto el espacio horizontal H_p esta generado por $\{\frac{\partial}{\partial x_{n+1}}(p), \dots, \frac{\partial}{\partial x_m}(p)\}$.

Sean dM y dN las formas de volúmen inducidos en M y N , $G = ((g_{ij}))_{i,j=1,\dots,m}$ y $H = ((h_{ij}))_{i,j=1,\dots,n}$ las matrices correspondientes a los coeficientes de las métricas en M y N respectivamente. Entonces $dM(\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_m}(p)) = \sqrt{\det G(p)}$ y $dN(\frac{\partial}{\partial x_1}(q), \dots, \frac{\partial}{\partial x_n}(q)) = \sqrt{\det H(q)}$. Además, el volúmen del paralelepípedo generado por $D_p f \frac{\partial}{\partial x_1}(p), \dots, D_p f \frac{\partial}{\partial x_n}(p)$ es por definición $\mathcal{J}_f(p)$, de donde se deduce que

$$dN(\frac{\partial}{\partial x_1}(f(p)), \dots, \frac{\partial}{\partial x_m}(f(p))) = \mathcal{J}_f(p) dM|_{H_p}(\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_n}(p)),$$

donde notamos $dM|_{H_p}$ a la restricción de la forma de volúmen dM al subespacio H_p . Entonces descomponiendo dM en el espacio H_p y el complementario ortogonal ($dM = dM_{H_p} \wedge dM_{H_p^\perp}$), obtenemos:

$$\begin{aligned} \mathcal{J}_f(p) \sqrt{\det G(p)} &= \mathcal{J}_f(p) dM(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}, \frac{\partial}{\partial x_{n+1}}, \dots, \frac{\partial}{\partial x_m}) \\ &= \mathcal{J}_f(p) dM|_{H_p} \mathcal{J}_f(p) \sqrt{\det H(f(p))} (\frac{\partial}{\partial x_1}(p), \dots, \frac{\partial}{\partial x_n}(p)) dM|_{H_p^\perp}(\frac{\partial}{\partial x_{n+1}}(p), \dots, \frac{\partial}{\partial x_m}(p)) \\ &= \sqrt{\det H(f(p))} \sqrt{\det G_{(m-n)}(p)} \end{aligned}$$

donde $G_{(m-n)} = ((g_{ij}))_{i,j=n+1,\dots,m}$. Notando $x_h = (x_1, \dots, x_n)$ y $x_v = (x_{n+1}, \dots, x_m)$, e integrando en coordenadas obtenemos:

$$\begin{aligned} &\int_U \mathcal{J}_f \cdot \phi dM = \\ &= \int_{B_\varepsilon^n(0) \times B_\delta^{m-n}(0)} \mathcal{J}_f(\varphi^{-1}(x_h, x_v)) \phi(\varphi^{-1}(x_h, x_v)) \sqrt{\det G(\varphi^{-1}(x_h, x_v))} dx_1 \cdots dx_m = \\ &= \int_{B_\varepsilon^n(0) \times B_\delta^{m-n}(0)} \phi(\varphi^{-1}(x_h, x_v)) \sqrt{\det H(\psi^{-1}(x_h))} \sqrt{\det G_{(m-n)}(\varphi^{-1}(x_h, x_v))} dx_1 \cdots dx_m \\ &= \int_{B_\varepsilon^n(0)} \sqrt{\det H(\psi^{-1}(x_h))} \left(\int_{B_\delta^{m-n}(0)} \phi_{\psi^{-1}(x_h)}(\varphi^{-1}(x_h, x_v)) \cdot \right. \\ &\quad \left. \cdot \sqrt{\det [(G_{(m-n)})_{\psi^{-1}(x_h)}(\varphi^{-1}(x_h, x_v))] dx_{n+1} \cdots dx_m} \right) dx_1 \cdots dx_n \\ &= \int_{B_\varepsilon^n(0)} \bar{\phi}(\psi^{-1}(x_h)) \sqrt{\det H(\psi^{-1}(x_h))} dx_1 \cdots dx_n = \int_V \bar{\phi} dN \end{aligned}$$

obteniéndose el resultado para todo función $\phi : M \rightarrow \mathbb{R}$ con soporte incluido en U .

Veamos como se puede extender lo anterior a toda la variedad. Podemos encontrar un atlas $\{(U_\alpha, \varphi_\alpha)\}_{\alpha \in A}$ de M donde las parejas resultan de la forma local de las submersiones (donde $\{(V_\alpha, \psi_\alpha)\}_{\alpha \in A}$ es el atlas respectivo en N con esa propiedad local). Entonces, lo que probamos arriba es que la fórmula vale para $f|_{U_\alpha} : U_\alpha \rightarrow V_\alpha$. Sea $\{\rho_\alpha\}_{\alpha \in A}$ una partición de la unidad subordinada al cubrimiento $\{U_\alpha\}_{\alpha \in A}$. Entonces tenemos que

$$\begin{aligned}
\int_M \mathcal{J}_f \cdot \phi \, dM &= \sum_{\alpha \in A} \int_{U_\alpha} \rho_\alpha \mathcal{J}_f \cdot \phi \, dM = \sum_{\alpha \in A} \int_{V_\alpha} \overline{(\rho_\alpha \phi)} \, dN \\
&= \sum_{\alpha \in A} \int_{q \in V_\alpha} \int_{f^{-1}(q) \cap U_\alpha} \rho_\alpha \phi \, dM_q \, dN \\
&= \sum_{\alpha \in A} \int_{q \in V_\alpha} \int_{f^{-1}(q)} \rho_\alpha \phi \, dM_q \, dN \\
&= \sum_{\alpha \in A} \int_N \int_{f^{-1}(q)} \rho_\alpha \phi \, dM_q \, dN \\
&= \int_N \int_{f^{-1}(q)} \sum_{\alpha \in A} \rho_\alpha \phi \, dM_q \, dN = \int_N \bar{\phi} \, dN
\end{aligned}$$

concluyendo la prueba. □

Bibliografía

- [1] D. Armentano y M. Wschebor: Random systems of polynomial equations. The expected number of roots under smooth analysis, preprint (2006).
- [2] J.-M Azais y M. Wschebor: A self contained proof of the Rice formula for random fields. <http://www.lsp.ups-tlse.fr/Azais/publi/completeproof.pdf>.
- [3] J.-M Azais y M. Wschebor: On the Roots of a Random System of Equations. The Theorem of Shub and Smale and Some Extensions, *Found. Comput. Math.*, (2005), 125-144.
- [4] J.-M Azais y M. Wschebor: Rice Formula. A complete proof. <http://www.cmat.edu.uy/~wschebor>.
- [5] L. Blum, F. Cucker, M. Shub y S. Smale: *Complexity and real computation*, Springer, 1998.
- [6] A. T. Bharucha-Reid y M. Sambandham: *Random Polynomials*, Probability and Mathematical Statistics, Academic Press, Orlando, FL, 1986.
- [7] Bloch y Pólya: On the number of real roots of a random algebraic equation, *Proc. Cambridge Philos. Soc.* **33**,(1932),102-114.
- [8] M.-P. do Carmo: *Geometria Riemanniana*, Projecto Euclides, IMPA, Rio de Janeiro, 1988.
- [9] I. Chavel: *Isoperimetric Inequalities, Differential Geometric and Analytic Perspectives*, Cambridge University Press, 2001.
- [10] F. Cucker y R. Wong: *The collected papers of Stephen Smale*, Vol. III, Singapore University Press; River Edge, NJ : World Scientific, 2000.
- [11] F. Cucker y M. Wschebor: *On the Expected Condition Number of Linear Programming Problems*, *Numerische Mathematik*, 94, 3, (2003) 419-478.
- [12] A. Edelman y E. Kostlan: How many zeros of a random polynomial are real?, *Bull. Amer. Math. Soc. (N.S.)* **32**,(1), (1995), 1-37.
- [13] L. Evans y R. Gariepy: *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.
- [14] S. Gallot, D. Hulin y J. Lafontaine: *Riemannian geometry*, Universitext, Berlin : Springer, 2004.

- [15] M. Kac: On the average number of real roots of a random algebraic equation, *Bull. Amer. Math. Soc.* **49**,(1943),314-320.
- [16] : F. Kirwain *Complex Algebraic Curves*, Cambridge University Press 1992.
- [17] E. Kostlan: On the expected number of real roots of a system of random polynomial equations in *Found. Comp. Math.*, (Hong-Kong 2000), World Scientific River Edge, NJ, pp. 149-188 (2002).
- [18] Littlewood y Offord: On the number of real roots of a random algebraic equation, *J. London Math. Soc.* **13**,(1938),288-295.
- [19] Littlewood y Offord: On the roots of certain algebraic equation, *Proc. London Math. Soc.* **35**,(1939),133-148.
- [20] D. Mumford: *Algebraic Geometry I. Complex Projective Varieties*, Springer-Verlag, Berlin Heidelberg New York 1970.
- [21] T. Sakai: *Riemannian Geometry*, Translations of mathematical monographs, 149, Providence, RI : American Mathematical Society, 1996.
- [22] M. Shub y S. Smale: Complexity of Bezout's theorem, II. Volumes and probabilities, in *Computational Algebraic Geometry* (Nice, 1992), *Progr. Math*, vol.109, pp. 267-285, Birkhäuser Boston, Boston, MA, 1993.
- [23] S. Smale: Newton's methods estimates from data at one point, in R. Ewing, K. Gross, and C. Martin (Eds) *The Merging of Disciplines: New Directions in Pure, Applied, and Computational Mathematics*. Springer-Verlag. 1986.
- [24] M. Spivak: *A comprehensive Introduction to Differential Geometry*, Vol. V, 2nd ed., Publish or Perish, Wilmington, DE, 1979.
- [25] M. Wschebor: On the Kostlan-Shub-Smale model for random polynomial systems. Variance of the number of roots., *J. Complexity*, **21**,(6), (2005), 773-789.