



UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA
DIVISIÓN DE ESTUDIOS DE POSGRADO
DOCTORADO EN MODELACIÓN MATEMÁTICA

Extensión del modelo canónico lineal a tramos de alto nivel

TESIS
QUE PARA OBTENER EL GRADO DE:
DOCTOR EN MODELACIÓN MATEMÁTICA

PRESENTA:
M.I.E.C. VICTOR MANUEL TLAPA CARRERA

DIRECTOR DE TESIS:
DR. CUAUHTÉMOC HÉCTOR CASTAÑEDA ROLDÁN
CO-DIRECTOR DE TESIS:
DR. VÍCTOR MANUEL JIMÉNEZ FERNÁNDEZ

HUAJUAPAN DE LEÓN, OAXACA, MÉXICO. DICIEMBRE 2020.

Dedicatoria

A Desirée Carrera Martínez y Jesús Ángel López Ocampo.

Agradecimientos

En especial al Dr. Cuauhtémoc Héctor Castañeda Roldán y al Dr. Víctor Manuel Jiménez Fernández, por su entera confianza y apoyo a lo largo de este trabajo.

A mis revisores, Dra. Luz del Carmen Álvarez Marín, Dr. Virgilio Vázquez Hipólito, Dra. Orquídea Sánchez López, Dr. Evodio Muñoz Aguirre, Dr. José Margarito Hernández Morales, por sus valiosas observaciones, y en particular al Dr. Richard Jacinto Márquez Contreras, quien siempre mostró interés y una participación activa en el desarrollo de esta tesis.

Extensión del modelo canónico lineal a tramos de alto nivel
Victor Manuel Tlapa Carrera
Huajuapán de León, Oaxaca, México. Diciembre 2020

Índice general

Índice de Figuras	iii
Índice de Tablas	v
Glosario de abreviaturas y símbolos	v
1 Marco de referencia.	1
1.1 Introducción.	1
1.2 Definición del problema.	5
1.3 Justificación.	5
1.3.1 Pertinencia.	5
1.3.2 Relevancia.	5
1.4 Hipótesis.	6
1.5 Objetivos.	6
1.5.1 Objetivo general.	6
1.5.2 Objetivos particulares.	6
1.6 Metas.	6
1.7 Limitaciones.	7
1.8 Método.	7
2 Marco teórico.	9
2.1 Funciones lineales a tramos.	9
2.1.1 Regiones e intersecciones.	9
2.1.2 Funciones lineales a tramos y representación estándar.	13
2.1.3 Variación consistente.	16
2.1.4 Función salto.	18
2.1.5 Propiedad de degeneración mínima.	20
2.1.6 Otras propiedades de funciones PWL.	27
2.2 Representación canónica lineal a tramos.	28
2.2.1 Representación canónica lineal a tramos de alto nivel.	32
3 Particiones simpliciales para la representación HL-CPWL.	35
3.1 La función generadora γ .	36
3.2 Partición simplicial del hipercubo $[-1, 1]^n$.	38

3.2.1	Descripción del dominio.	38
3.2.2	Conjunto generador del espacio $PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$	42
3.3	Partición simplicial ortogonal, uniformemente espaciada, de $D_{\text{rec}} \subset \mathbb{R}^n$	46
3.3.1	Descripción del dominio.	46
3.3.2	Conjunto generador del espacio $PWL_{H_{D_{\text{rec}}}} [D_{\text{rec}}]$	49
3.4	Propiedades computacionales.	54
4	Resultados de este trabajo.	57
4.1	Generalización de la representación HL-CPWL con partición simplicial.	57
4.1.1	Descripción del dominio.	57
4.1.2	Conjunto generador del espacio $PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$	65
4.1.3	Propiedades computacionales.	73
4.1.4	Algoritmo para la construcción de una representación HL-CPWL de funciones $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$	74
4.2	Un método para extender la representación CPWL de una a dos dimensiones.	81
4.2.1	Proyección de representaciones CPWL.	81
5	Sobre el ajuste a una nube de puntos.	85
5.1	Nube de puntos.	85
5.2	Funciones continuas PWL a ajustar.	86
5.2.1	Particiones del dominio.	86
5.2.2	Conjuntos generadores.	89
5.3	Ajuste por mínimos cuadrados.	91
Discusión.		97
Conclusiones.	97
Trabajos futuros.	98
Referencias		99

Índice de Figuras

1.1	(a) Diagrama eléctrico de un diodo polarizado por una diferencia de potencial v , en el cual circula una corriente i . (b) Curva voltaje-corriente de un diodo usando un modelo PWL; el modelo Shockley y curva característica de un diodo ideal.	2
1.2	(a) Diagrama eléctrico de una resistencia convexa polarizada por una diferencia de potencial v , por la cual circula una corriente i . (b) Curva voltaje-corriente de una resistencia convexa.	3
1.3	(a) Diagrama eléctrico de un arreglo en paralelo de resistencias convexas más una fuente de corriente y una resistencia. (b) Curvas voltaje-corriente de cada rama del circuito. (c) Curva voltaje-corriente del circuito.	3
2.1	Ejemplo de partición H en \mathbb{R}^2	10
2.2	Ejemplo de región R en \mathbb{R}^2	11
2.3	Ejemplo de regiones R_i , $i \in \{1, \dots, 4\}$, adyacentes y no adyacentes, entre sí, en \mathbb{R}^2	12
2.4	(a) Intersecciones en \mathbb{R}^2 e (b) intersecciones en \mathbb{R}^3	13
2.5	Partición H de la función PWL $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ del Ejemplo 4.	14
2.6	Camino cerrado Γ en \mathbb{R}^2	18
2.7	Ejemplo gráfico de una función salto de orden 1 en \mathbb{R}^2	18
2.8	Salto de orden 1 y 2 en el plano.	19
2.9	(a - f) Intersecciones (o hiperplanos) de orden 1, $\hat{S}_1, \dots, \hat{S}_6$, respectivamente de la partición H dada para el Ejemplo 11.	21
2.10	(a - d) Intersecciones $\hat{S}_1^{(2)}, \dots, \hat{S}_4^{(2)}$, respectivamente (línea negra sólida), e hiperplanos, de la partición H , que las forman (superficie blanca rayada). Para el Ejemplo 11.	23
2.11	(a) Hiperplanos en H que forman a la intersección de degeneración mínima $\hat{S}_1^{(3)}$ de orden 3. (b) Intersecciones de degeneración mínima de orden 2 que forman a la intersección de degeneración mínima $\hat{S}_1^{(3)}$ de orden 3. Ejemplo 11.	24
3.1	Símplices en el primer cuadrante del hipercubo $[-1, 1]^2$	39
3.2	Partición $H_{[-1,1]^2}$ del hipercubo $[-1, 1]^2$	40
3.3	D_{eje} y partición $H_{D_{eje}}$ del Ejemplo 22.	47

4.1	(a) Vectores ξ_1 , ξ_2 y \mathbf{p} del Ejemplo 27. (b) Conjunto $D_{\text{eje}} \subset \mathbb{R}^2$ del Ejemplo 27.	58
4.2	(a) Hiperplanos S_1 , S_2 y $S_{1,2}$ generados por los vectores ξ_1 , ξ_2 y $\xi_1 + \xi_2$, respectivamente. (b) Partición $H_{D_{\text{eje}}}$ de D_{eje}	61
4.3	Función f del Ejemplo 33 (a) y vista sobre el plano x_1x_2	73
4.4	Algoritmo para la construcción de una representación HL-CPWL de funciones $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$	75
4.5	Proyecciones $\text{Proy}_{f((x, i))}xy$ en el plano xy (Plano de color gris). Funciones $f((x, i))$ (línea espaciada en el espacio xyz). Funciones $f((x, i))$ (línea sólida en el espacio xyz).	82
5.1	(a) Muestreo a. (b) Muestreo b. Puntos del plano x_1x_2 donde fueron evaluadas las funciones para obtener las nubes de puntos en el espacio de tres dimensiones. La línea sólida indica la frontera del rectángulo del que fueron tomados estos puntos.	86
5.2	(a) Partición H_1 de D_1 . (b) Partición H_2 de D_2 . Partición H_3 de D_3 . Para los tres casos, tanto los hiperplanos de cada partición como la frontera de los conjuntos que particionan se encuentran señalados por la línea espaciada de color gris. Se señala también, con una línea de color negro, la frontera del conjunto de donde se toman las nubes de puntos.	88
5.3	(a - c) Función $\text{sen}(x_1)$ (color negro) y el ajuste a las nubes de puntos, muestreo a, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (d - f) Nubes de puntos, muestreo a obtenido de la función $\text{sen}(x_1)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (g - i) Función $\text{sen}(x_1)$ (color negro) y el ajuste a las nubes de puntos, muestreo b, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (j - l) Nubes de puntos, muestreo b obtenido de la función $\text{sen}(x_1)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco).	94
5.4	(a - c) Función $\text{sen}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$ (color negro) y el ajuste a las nubes de puntos, muestreo a, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (d - f) Nubes de puntos, muestreo a obtenido de la función $\text{sen}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (g - i) Función $\text{sen}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$ (color negro) y el ajuste a las nubes de puntos, muestreo b, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (j - l) Nubes de puntos, muestreo b obtenido de la función $\text{sen}\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco).	95

Índice de Tablas

4.1	Valores de f en los vértices $V_{D_{\text{eje}}}$ del Ejemplo 33.	72
4.2	Obtención del arreglo de índices.	75
4.3	Obtención de un conjunto generador del espacio $PWL_{HD_{\text{gen}}}[D_{\text{gen}}]$	77
4.4	Interpolación de la función.	80
4.5	Puntos en D y su valor correspondiente $y_i = f((x_i, z_i))$	83
5.1	Bases que se ajustan a las nubes de puntos (columnas). Nubes de puntos obtenidas de las funciones $\text{sen}(x_1)$ y $\text{sen}(\frac{x_1}{2} + \frac{x_2}{2})$ por los muestreos a y b (filas). Va-lo-res de $E_{h,j}$ (intersección).	92

Glosario de abreviaturas y símbolos

α Vector normal a un hiperplano S .

A ordenado Tupla de índices ordenada como en la Definición 14.

B ordenado Tupla de índices ordenada como en la Definición 15.

CPWL Canónico lineal a tramos.

D Poliedro convexo en \mathbb{R}^n .

$\Delta^{(l)} \mathcal{J}_{\mu_i}(\mathbf{x})$ Función salto de orden k .

D_{gen} Subconjunto de \mathbb{R}^n definido por (4.1).

D_{rec} Subconjunto de \mathbb{R}^n rectangular definido en (3.16).

$\gamma^{(k)}$ Función generadora γ de orden k .

H Partición.

$H_{[-1,1]^n}$ Partición del hipercubo $H_{[-1,1]^n}$ formada por los hiperplanos (3.8).

$H_{D_{\text{gen}}}$ Partición de D_{gen} formada por los hiperplanos (4.13).

HL-CPWL Canónico lineal a tramos de alto nivel.

$H_{D_{\text{rec}}}$ Partición de D_{rec} formada por los hiperplanos (3.17).

\mathbf{J}_i Vector gradiente de una función PWL en una i -ésima región.

$\Lambda_{[-1,1]^n}$ Vector de funciones PWL las cuales generan el espacio $PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$.

$\Lambda_{D_{\text{gen}}}$ Vector de funciones PWL las cuales generan el espacio $PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$.

$\Lambda_{D_{\text{rec}}}$ Vector de funciones PWL las cuales generan el espacio $PWL_{H_{D_{\text{rec}}}} [D_{\text{rec}}]$.

$N_k(l)$ Número de intersecciones de orden l con degeneración mínima necesarias para formar una intersección de orden k con degeneración mínima.

$nl(f)$ Nivel de anidación de una función f .

$\pi_{\alpha_i, \alpha_j}^{-(k_i, k_j), \mathcal{P}}$ Representación de la forma (4.12) de la función que caracteriza a un hiperplano.

$\pi_{\alpha_q}^{(k_q), \mathcal{P}}$ Representación de la forma (4.12) de la función que caracteriza a un hiperplano.

$\pi_q^{(k_q)}$ Representación de la forma (3.7) de la función que caracteriza a un hiperplano.

$\pi_{i,j}^{+(k_i, k_j)}$ Representación de la forma (3.7) de la función que caracteriza a un hiperplano.

$\pi_{i,j}^{-(k_i, k_j)}$ Representación de la forma (3.7) de la función que caracteriza a un hiperplano.

Γ Parametrización de un camino cerrado en el dominio de una función PWL.

PWL Lineal a tramos.

$PWL_H [D]$ Espacio de funciones PWL definidas en D con partición H .

R Región dada por la Definición 4.

S Hiperplano.

$S^{(+)}$ Semiespacio abierto.

$S^{(-)}$ Semiespacio abierto.

$S^{(k)}$ Intersección de orden k .

$S^{[+]}$ Semiespacio cerrado.

$S^{[-]}$ Semiespacio cerrado.

$\hat{S}^{(k)}$ Intersección de mínima degeneración.

$\blacktriangle (\mathbf{x}_0, \dots, \mathbf{x}_n)$ Símplice dado por los punto $\mathbf{x}_0, \dots, \mathbf{x}_n \in \mathbb{R}^n$, como en la Definición 16.

$V_{[-1,1]^n}$ Conjunto irdenado de vértices formados por los símlices dados por la partición $H_{[-1,1]^n}$.

$V_{D_{\text{gen}}}$ Conjunto ordenado de vértices formados por los símlices dados por la partición $H_{D_{\text{gen}}}$.

$V_{D_{\text{rec}}}$ Conjunto ordenado de vértices formados por los símlices dados por la partición $H_{D_{\text{rec}}}$.

Capítulo 1

Marco de referencia.

1.1 Introducción.

Desde su primera aparición [1, 2], las representaciones explícitas de funciones lineales a tramos (PWL por sus siglas en inglés) han sido ampliamente utilizadas en áreas de la electrónica, tales como lo son el modelado de dispositivos eléctricos y el análisis de circuitos eléctricos no-lineales [3, 4]. Un ejemplo tradicional de su aplicación [5] es el modelado del comportamiento del diodo. En la Figura 1.1a se aprecia el diagrama de un diodo en polarización directa, es decir, que la diferencia de potencial aplicada al diodo produce una corriente, la cual va del ánodo al cátodo de este. En la Figura 1.1b se pueden ver tres curvas que modelan el comportamiento de la corriente que circula a través del diodo en relación con la diferencia de potencial o voltaje v entre sus terminales: la primera curva se ajusta al comportamiento de un diodo ideal, se puede ver que esta es una recta vertical, esto significa que a partir del punto en el que el diodo es polarizado directamente es tratado como un circuito cerrado (o un corto circuito), sin embargo esto no ocurre en un diodo real; la segunda curva está dada por el modelo de Shockley,

$$i = I_s (e^{v/v_t} - 1),$$

donde I_s es una constante en el orden de los microamperes, llamada corriente de saturación inversa, el parámetro $v_t = kT/q$ es llamado voltaje térmico y es directamente proporcional a la temperatura T , dada en grados K, k es la constante de Boltzmann y q es la carga de un electrón; la tercer curva es una aproximación por una función PWL al modelo de Shockley. Como se puede apreciar en las curvas dadas por el modelo de Shockley y la aproximación PWL, la corriente que circula a través del diodo empieza a crecer abruptamente a partir de un punto llamado voltaje de disparo (threshold voltage), este incrementa en relación al incremento en el voltaje aplicado a sus terminales. La primer ventaja que se puede observar que tiene el modelo PWL sobre el modelo de Shockley radica en su capacidad de representación, en este ejemplo el "punto de quiebre" es colocado en el voltaje de disparo, sin embargo, este puede variar tanto en su posición como en la cantidad de puntos de quiebre que se añaden a la curva, de tal forma que pueda tener un mejor ajuste a la curva de un diodo real, ya sea añadiendo una mayor cantidad de segmentos o cambiando el valor de la función en los puntos de quiebre. Así, la complejidad del modelo PWL se

encuentra dada en función de la precisión que se desea alcanzar. Ahora suponga que se

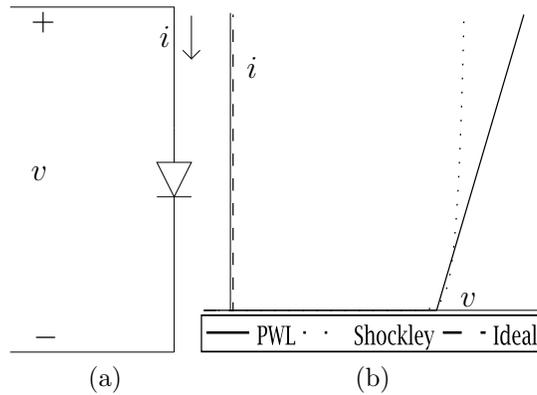


Figura 1.1: (a) Diagrama eléctrico de un diodo polarizado por una diferencia de potencial v , en el cual circula una corriente i . (b) Curva voltaje-corriente de un diodo usando un modelo PWL; el modelo Shockley y curva característica de un diodo ideal.

quiere representar el comportamiento de un dispositivo no-lineal formado por un arreglo en serie de una resistencia, una fuente de voltaje en corriente directa (DC) y un diodo ideal, como en la Figura 1.2a. Retomando la idea del comportamiento de un diodo ideal, el cual se considera como un corto circuito cuando la diferencia de potencial entre sus terminales es mayor que 0, el añadir una fuente de voltaje mueve el valor del voltaje de disparo al valor de E , es decir, por el diodo empezará a circular corriente cuando $v > E$, además, la resistencia en este arreglo tiene el efecto de modificar la curva de corriente-voltaje al aplicar v entre las terminales del dispositivo, este efecto se observa gráficamente en la Figura 1.2b, donde a diferencia del comportamiento de un diodo ideal, este dispositivo presenta una recta con una pendiente $G = 1/R$, $R > 0$, cuando v supera el voltaje de disparo dado por E . A este dispositivo se le llama resistencia convexa y su comportamiento es modelado por una función PWL. Ahora veamos un arreglo en paralelo formado por resistencias convexas, una resistencia y una fuente de corriente directa, al cual se le aplica un voltaje v entre sus terminales como en la Figura 1.3a. En la Figura 1.3b se ven en la misma gráfica las curvas voltaje-corriente de cada rama en el circuito, donde, por leyes de Ohm, estas están dadas por:

$$\begin{aligned} i_0 &= G_0 v, \\ i_1 &= \frac{1}{2} G_1 [|v - E_1| + (v - E_1)], \\ i_2 &= \frac{1}{2} G_2 [|v - E_2| + (v - E_2)], \end{aligned}$$

además, I es un valor que no varía con respecto a v . Por leyes de Kirchhoff de corriente, la suma de todas las corrientes que pasan por un nodo es igual a cero, entonces

$$i = -I - \frac{1}{2} (G_1 E_1 + G_2 E_2) + \left(G_0 + \frac{1}{2} G_1 + \frac{1}{2} G_2 \right) v + \frac{1}{2} G_1 |v - E_1| + \frac{1}{2} G_2 |v - E_2|.$$

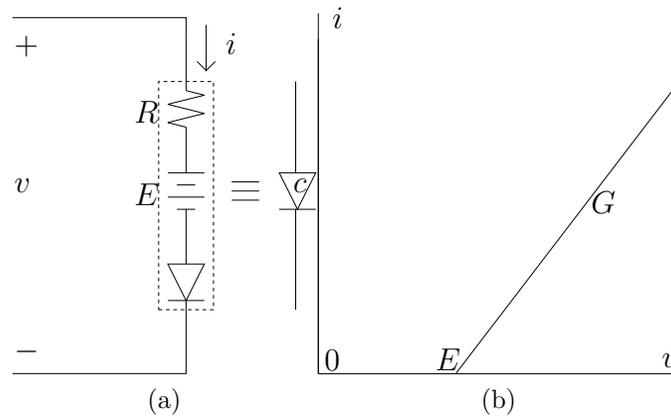


Figura 1.2: (a) Diagrama eléctrico de una resistencia convexa polarizada por una diferencia de potencial v , por la cual circula una corriente i . (b) Curva voltaje-corriente de una resistencia convexa.

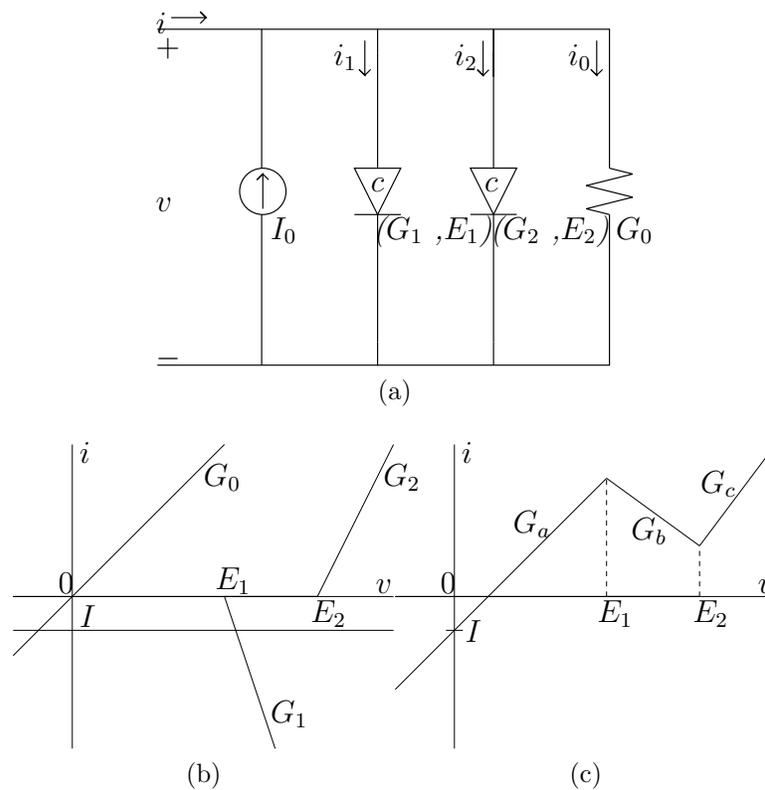


Figura 1.3: (a) Diagrama eléctrico de un arreglo en paralelo de resistencias convexas más una fuente de corriente y una resistencia. (b) Curvas voltaje-corriente de cada rama del circuito. (c) Curva voltaje-corriente del circuito.

Por lo que las pendientes en la Figura 1.3c están dadas por

$$\begin{aligned}G_a &= G_0, \\G_b &= G_0 + G_1, \\G_c &= G_0 + G_1 + G_2.\end{aligned}$$

De los ejemplos anteriores, se pueden señalar las principales motivaciones para el uso de funciones PWL, las cuales son [6]:

- Tomar ventaja de técnicas establecidas para el análisis simbólico, numérico y gráfico de funciones que localmente son lineales o afines.
- Poseen una menor complejidad que otros modelos de representación no-lineal.
- Pueden ser eficientemente implementados en modelos computacionales.
- Poseen propiedades de aproximación universal, esto significa que cualquier comportamiento no-lineal, inclusive en funciones multivariantes, puede ser aproximado por una función lineal a tramos con una precisión arbitraria, la cual se alcanza al refinar la partición del dominio en regiones más pequeñas.

Es por esto que el uso de estos modelos se ha extendido a otras áreas del conocimiento como lo son: el procesamiento de imágenes [7,8], el estudio en sistemas de control predictivo [9, 10], ingeniería química [11], nuevas aplicaciones en el diseño de circuitos y sistemas electrónicos [12, 13], implementaciones en redes neuronales [14], modelado de superficies tridimensionales [15], entre otros.

Al esquema de construcción de funciones PWL basado en sumatorias de funciones de valor absoluto, abordado en el ejemplo del arreglo en paralelo de resistencias convexas, se le llama representación canónica lineal a tramos (CPWL por sus siglas en inglés) y fue el primer esquema de representación global para funciones PWL. Las limitaciones que este esquema presenta se encuentran relacionadas con la propiedad de variación consistente [16]. Para superar esta problemática, en [17] se propuso una representación basada en la anidación de funciones CPWL, llamada canónica lineal a tramos de alto nivel (HL-CPWL por sus siglas en inglés). En [18] se prueba que cualquier función PWL tiene una representación HL-CPWL, sin embargo, esta demostración es puramente teórica, por lo que no provee de una metodología constructiva. En [19] se presenta un esquema de construcción de representaciones HL-CPWL de funciones PWL con una partición formada por una intersección de degeneración mínima [20]. Sin embargo, como se señala en [21], el uso de particiones muy generales dificulta la obtención de una representación, debido a que para cada partición se debe revisar la propiedad de variación consistente si es que se quiere obtener una representación CPWL, o verificar la propiedad de intersección de degeneración mínima, si es que se quiere obtener una representación HL-CPWL. Para evitar estos inconvenientes, en [22] se propone una partición ortogonal, simplicial, igualmente distribuida del dominio de las funciones PWL, de tal manera que sea posible obtener una representación HL-CPWL. Hasta este momento este esquema de construcción ha sido el estándar usado en la mayoría de las aplicaciones donde se implementa una representación

HL-CPWL. Sin embargo, este esquema es rígido debido a la ortogonalidad y espaciamiento uniforme requerido en la partición.

1.2 Definición del problema.

Desarrollar un esquema de construcción de funciones lineales a tramos con una partición simplicial que no sea necesariamente uniformemente distribuida [22] es un problema que se ha abordado recientemente [21, 23], sin embargo, la propuesta de los autores no consigue superar del todo esta problemática, pues su construcción aún propone particiones simpliciales rígidas, es decir, particiones lineales a base de únicamente hiperplanos paralelos y ortogonales entre sí, además que sus pruebas se realizan únicamente para particiones de funciones PWL que van de \mathbb{R}^2 a \mathbb{R} .

1.3 Justificación.

1.3.1 Pertinencia.

En la literatura se han establecido diversas particiones lineales para la definición de funciones PWL. Hasta ahora las más convenientes por sus capacidades de representación sugieren un mallado regular [21–23] a base de triángulos (símplices) que cubren el dominio, sin embargo, estos esquemas son rígidos al tratarse de esquemas basados en mallas ortogonales uniformemente distribuidas, sin considerar las características de los datos o las superficies a las que se pretenden ajustar.

Con este trabajo se extiende la capacidad que tiene la representación canónica lineal a tramos de alto nivel de describir el comportamiento de funciones lineales a tramos, manteniendo ventajas computacionales dadas por la literatura [22], tales como el número de parámetros utilizados y el cálculo de coeficientes de dicha representación.

1.3.2 Relevancia.

Con la finalidad de mejorar los esquemas de construcción de funciones lineales a tramos, y sus cualidades de aproximación al modelar el comportamiento de una nube de puntos, el proporcionar una metodología que permita construir expresiones canónicas lineales a tramos con diferentes particiones lineales puede ayudar a establecer, entre todas sus variantes, la que mejor se ajuste a las necesidades del entorno en el que se aplique el modelo. Más aún, permitiría desarrollar un algoritmo que genere funciones HL-CPWL bajo diversas particiones lineales con un menor costo computacional (tiempo de cómputo y memoria).

Con esto como principal motivación, el trabajo a desarrollar buscará proponer un método que facilite la construcción de funciones lineales a tramos mejorando el desempeño y ampliando el ámbito de aplicación de técnicas existentes, así como encontrar la forma canónica de una función PWL que aproxime mejor a un conjunto de datos, definida sobre particiones menos restrictivas.

1.4 Hipótesis.

A partir de los métodos existentes para la construcción de funciones lineales a tramos [21–23], se puede establecer una metodología que considere diferentes esquemas de partición, no necesariamente a base de mallas ortogonales.

1.5 Objetivos.

1.5.1 Objetivo general.

Desarrollar un método y a partir de este un algoritmo que permita construir representaciones canónicas de funciones lineales a tramos que se ajusten a una nube de puntos dada sobre una malla no ortogonal, usando el menor número posible de parámetros y el menor uso de recursos computacionales (tiempo de cómputo y espacio en memoria).

1.5.2 Objetivos particulares.

1. Definir formalmente el dominio de las funciones PWL a representar.
2. Proponer una partición del dominio definido para las funciones PWL a representar.
3. Establecer un método para obtener una base del espacio de funciones lineales a tramos definidas en el dominio para las funciones PWL a representar, con la partición propuesta para este dominio.
4. Obtener un método que permita construir una representación HL-CPWL de funciones PWL en el espacio de funciones lineales a tramos definidas en el dominio dado y con particiones más generales a las existentes en la literatura.
5. Realizar pruebas numéricas del método obtenido.
6. Obtener una representación HL-CPWL, usando un menor o igual número de parámetros y recursos computacionales, de funciones PWL con esquemas de partición más generales a los esquemas de construcción de representaciones HL-CPWL existentes.
7. Mostrar la aplicación del esquema de construcción de representaciones PWL desarrollado al ajuste de una superficie.

1.6 Metas.

- Obtener un algoritmo para la construcción de una representación canónica de las funciones lineales a tramos con un dominio y particiones más generales a las dadas por los métodos de construcción de representaciones HL-CPWL existentes en la literatura.
- Establecer de manera formal las pruebas que garantizan la efectividad del algoritmo.

- Mostrar el desempeño de esta nueva propuesta con un ejemplo aplicado al ajuste de una nube de puntos.

1.7 Limitaciones.

En este trabajo se propone un nuevo esquema de construcción de representaciones canónicas lineales a tramos de alto nivel con partición simplicial, cuyas intersecciones entre hiperplanos de su partición formen intersecciones que tengan a lo más la propiedad de degeneración mínima, las cuales conserven las ventajas computacionales de reducir el número de parámetros para describir el comportamiento de una función continua lineal a tramos en relación con la representación estándar, optimizando así el tiempo del cómputo en la construcción de dichas representaciones. El problema de encontrar una representación canónica lineal a tramos de alto nivel con particiones más generales, las cuales tengan intersecciones con un mayor número de hiperplanos comparadas con las que poseen una intersección con degeneración mínima sigue siendo un problema abierto.

1.8 Método.

La metodología implementada para el desarrollo de este trabajo consta de: un estudio de los resultados principales existentes en la literatura relativos a las funciones lineales a tramos, donde se hace énfasis en el estudio de las propiedades de variación consistente y degeneración mínima, y la representación canónica de funciones lineales a tramos (Capítulo 2); posteriormente, se detalla el esquema de construcción de representaciones canónicas lineales a tramos de alto nivel con particiones simpliciales ortogonales igualmente distribuidas, esto con la finalidad de comprender y desarrollar un método de construcción sobre particiones más generales (Capítulo 3); a partir del desarrollo consecutivo de la descripción del dominio, descripción de una partición cuyas intersecciones posean la propiedad de degeneración mínima y las cuales sean del mismo orden que la dimensión del espacio de salida, ordenamiento de vértices en la partición, construcción de un conjunto generador del espacio de funciones lineales a tramos definidos e interpolación de funciones lineales a tramos, se hace un análisis computacional y se desarrolla un algoritmo que permita construir representaciones canónicas lineales a tramos de alto nivel con particiones más generales a las existentes en la literatura (Capítulo 4); además, se describe un método que permita aprovechar las capacidades de la representación canónica lineal a tramos de funciones lineales a tramos con un dominio de dimensión 1, de tal manera que esta pueda ser usada en la representación de funciones lineales a tramos con un dominio de dimensión 2 (Capítulo 4); posteriormente, se muestra un ejemplo de ajuste a una nube de puntos en el cual se propone que para algunos casos se puede dar una función con representación canónica lineal a tramos de alto nivel con un esquema de partición simplicial que requiera un menor número de parámetros de ajuste en comparación con las propuestas existentes en la literatura (Capítulo 5); para finalizar, se da una discusión, la cual incluye conclusiones y propuestas de trabajos futuros relacionados a lo abordado a lo largo del trabajo de investigación.

Capítulo 2

Marco teórico.

2.1 Funciones lineales a tramos.

A lo largo de esta sección se abordarán los conceptos necesarios para entender el comportamiento de las funciones lineales a tramos, principalmente aquellos utilizados para lograr cumplir con los objetivos y metas planteados en este trabajo, estos fueron mencionados y referenciados en la Introducción, sin embargo, aquí se desarrollarán con mayor detalle y usando ejemplos gráficos.

2.1.1 Regiones e intersecciones.

Definición 1. Definimos como *hiperplano* S a un subconjunto de \mathbb{R}^n tal que

$$S := \{ \mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{x} - \beta = 0 \}, \quad (2.1)$$

donde $\boldsymbol{\alpha} \in \mathbb{R}^n$ se dice que es un vector normal a S y $\beta \in \mathbb{R}$.

Definición 2. Llamaremos *partición* a una familia finita H de hiperplanos distintos en \mathbb{R}^n .

Ejemplo 1. En la Figura 2.1 se observa una partición $H = \{S_1, S_2, S_3, S_4\}$ formada por hiperplanos en \mathbb{R}^2 , donde

$$\begin{aligned} S_1 &= \{ \mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 - \frac{7}{2} = 0 \}, \\ S_2 &= \{ \mathbf{x} \in \mathbb{R}^2 : x_1 - x_2 + \frac{7}{2} = 0 \}, \\ S_3 &= \{ \mathbf{x} \in \mathbb{R}^2 : x_2 = 0 \}, \\ S_4 &= \{ \mathbf{x} \in \mathbb{R}^2 : \frac{2}{3}x_1 + x_2 + 2 = 0 \}. \end{aligned}$$

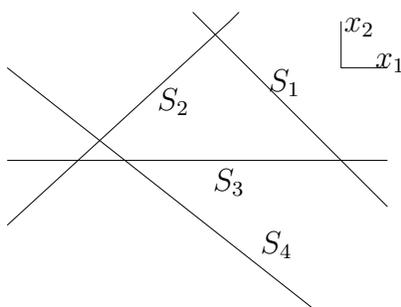


Figura 2.1: Ejemplo de partición H en \mathbb{R}^2 .

Un hiperplano es un subconjunto de \mathbb{R}^n de dimensión $n-1$. De modo que un hiperplano en \mathbb{R} es un punto en la recta, un hiperplano en \mathbb{R}^2 es una recta en el plano, como se vio en el ejemplo anterior, y un hiperplano en \mathbb{R}^3 es un plano en el espacio.

Definición 3. Un hiperplano S determina los semiespacios abiertos $S^{(+)}$ y $S^{(-)}$ definidos por

$$\begin{aligned} S^{(+)} &:= \{ \mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{x} - \beta > 0 \}, \\ S^{(-)} &:= \{ \mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{x} - \beta < 0 \}, \end{aligned}$$

y los semiespacios cerrados $S^{[+]}$ y $S^{[-]}$, dados por

$$\begin{aligned} S^{[+]} &:= \{ \mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{x} - \beta \geq 0 \}, \\ S^{[-]} &:= \{ \mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}^T \mathbf{x} - \beta \leq 0 \}. \end{aligned}$$

De la intersección finita, no vacía, entre semiespacios dados por una partición H se forman poliedros en \mathbb{R}^n .

Definición 4. Dada una partición finita $H = \{S_i, i \in I\}$ y $J \subset I$, llamamos *región* R_J a un conjunto no vacío de la forma

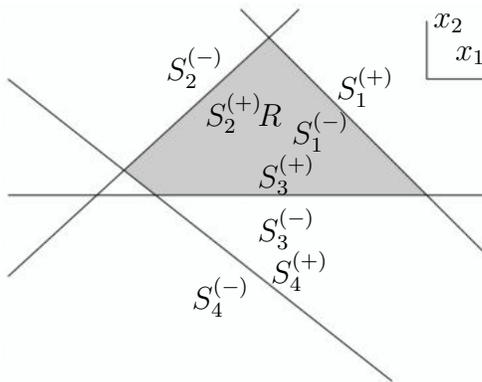
$$R_J := \left(\bigcap_{j \in J} S_j^{(+)} \right) \cap \left(\bigcap_{j \in J^c} S_j^{(-)} \right),$$

donde J^c es el complemento de J en I .

Podemos referirnos a una región determinada por H , sin especificar el conjunto J y representarla por R , simplemente.

Ejemplo 2. Continuando con el ejemplo anterior, si $I = \{1, 2, 3, 4\}$ y $J = \{2, 3, 4\}$, la región R se forma de la intersección $S_1^{(-)} \cap S_2^{(+)} \cap S_3^{(+)} \cap S_4^{(+)}$, como se ve en la Figura 2.2, donde

$$\begin{aligned} S_1^{(-)} &= \{ \mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 - \frac{7}{2} < 0 \}, \\ S_2^{(+)} &= \{ \mathbf{x} \in \mathbb{R}^2 : x_1 - x_2 + \frac{7}{2} > 0 \}, \\ S_3^{(+)} &= \{ \mathbf{x} \in \mathbb{R}^2 : x_2 > 0 \}, \\ S_4^{(+)} &= \{ \mathbf{x} \in \mathbb{R}^2 : \frac{2}{3}x_1 + x_2 + 2 > 0 \}. \end{aligned}$$

Figura 2.2: Ejemplo de región R en \mathbb{R}^2 .

Lema 1. Dada una partición H , vemos que:

- Dos regiones distintas R y R' , determinadas por H , son ajenas entre sí.
- Todo $\mathbf{x} \in \mathbb{R}^n$ tal que \mathbf{x} no está en ningún hiperplano de H , pertenece a alguna región determinada por H .
- Toda región determinada por H es un conjunto convexo.

Prueba. Consideremos una partición H de hiperplanos distintos S_i , $i \in I$, y dos subconjuntos de índices $J, J' \subset I$, $J \neq J'$.

Veamos que $R_J = \left(\bigcap_{j \in J} S_j^{(+)}\right) \cap \left(\bigcap_{j \in J^c} S_j^{(-)}\right) \neq \emptyset$ y $R_{J'} = \left(\bigcap_{j \in J'} S_j^{(+)}\right) \cap \left(\bigcap_{j \in J'^c} S_j^{(-)}\right) \neq \emptyset$ son regiones ajenas. Suponga que no es así, y sea $\mathbf{x} \in R_J$ y $\mathbf{x} \in R_{J'}$. Puesto que J y J' son distintas, asumamos sin pérdida de generalidad que existe un $i \in J$ tal que $i \in J'^c$. Por hipótesis $\mathbf{x} \in \left(\bigcap_{j \in J} S_j^{(+)}\right) \cap \left(\bigcap_{j \in J^c} S_j^{(-)}\right)$, así, $\mathbf{x} \in S_i^{(+)}$, pero además $\mathbf{x} \in \left(\bigcap_{j \in J'} S_j^{(+)}\right) \cap \left(\bigcap_{j \in J'^c} S_j^{(-)}\right)$, por lo que $\mathbf{x} \in S_i^{(-)}$, lo cual es una contradicción, pues $S_i^{(+)} \cap S_i^{(-)} = \emptyset$. Así R y $R_{J'}$ son ajenas entre sí.

Ahora probemos que todo $\mathbf{x} \in \mathbb{R}^n$ que no está en algún hiperplano de H pertenece a alguna región determinada por H . Vemos primero que, por hipótesis, \mathbf{x} está ya sea en $S_i^{(+)}$ o en $S_i^{(-)}$, para todo $i \in I$. Hacemos J el conjunto de todos los $i \in I$ tales que $\mathbf{x} \in S_i^{(+)}$, entonces $\mathbf{x} \in \left(\bigcap_{j \in J} S_j^{(+)}\right) \cap \left(\bigcap_{j \in J^c} S_j^{(-)}\right)$, por lo que $\mathbf{x} \in R_J$.

Sabemos que $S_i^{(+)}$ y $S_i^{(-)}$, $i \in I$, son conjuntos convexos y puesto que R es intersección finita de conjuntos convexos, R es un conjunto convexo [24].

■

Por lo anterior, abusando del lenguaje, se dice que H particiona a \mathbb{R}^n en regiones. Supongamos que H particiona a \mathbb{R}^n en las regiones R'_i , $i \in I$, y sea $D \subset \mathbb{R}^n$, $D \neq \emptyset$. Si para todo $i \in I$, $\text{int}(R'_i \cap D) \neq \emptyset$, donde $\text{int}(R'_i \cap D)$ indica el interior de $R'_i \cap D$, decimos que H particiona a D en las regiones $R_i = R'_i \cap D$, $i \in I$.

Para los propósitos de esta tesis, supondremos que D es un poliedro convexo. Bajo esta suposición, la Observación 1 es válida también para cualquier partición H de D y para las regiones determinadas por ella en D . En lo subsiguiente este será el contexto general para los resultados que se enuncian, siendo $D = \mathbb{R}^n$ un caso particular.

Dos regiones son adyacentes si se encuentran separadas por un hiperplano como se establece en la siguiente definición [17].

Definición 5. Dada una partición H , dos regiones R y R' son llamadas *adyacentes* si existe un punto $\mathbf{x} \in \mathbb{R}^n$ y un $r_0 > 0$, tal que para todo $0 < r \leq r_0$ la bola abierta con centro en \mathbf{x} y de radio r está contenida en la unión de las cerraduras de R y R' .

Dada una partición H la cual particiona a \mathbb{R}^n en k regiones, señalaremos cada una de estas regiones con un subíndice, de tal manera que R_i será la i -ésima región dada por H para $i \in \{1, \dots, k\}$.

Ejemplo 3. Sea $H = \{S_1, S_2\}$ la partición de \mathbb{R}^2 , donde $S_1 = \{\mathbf{x} \in \mathbb{R}^2 : x_1 - x_2 = 0\}$ y $S_2 = \{\mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 = 0\}$, mostrada en la Figura 2.3; de la definición 5, vemos que las regiones R_1 y R_3 no son adyacentes entre sí ya que no podemos encontrar un punto en \mathbb{R}^2 para el cual exista un $r_0 > 0$ tal que para todo $0 < r \leq r_0$ la bola abierta con centro en \mathbf{x} y de radio r está contenida en la unión de las cerraduras de R_1 y R_3 , lo mismo ocurre en las regiones R_2 y R_4 . Sin embargo, la región R_1 es adyacente a R_2 puesto que podemos encontrar un punto, supongamos $\mathbf{x} = (1 \ 1)^T$, y un $r_0 > 0$, supongamos $r = 1/4$, para los cuales para todo $0 < r \leq r_0$ la bola abierta con centro en \mathbf{x} y de radio r está contenida en la unión de las cerraduras de R_1 y R_2 . De igual forma R_1 es adyacente a R_4 y la región R_3 es adyacente a R_2 y a R_4 .

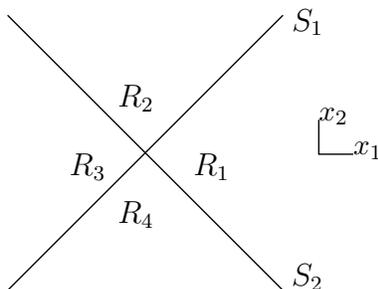


Figura 2.3: Ejemplo de regiones R_i , $i \in \{1, \dots, 4\}$, adyacentes y no adyacentes, entre sí, en \mathbb{R}^2 .

Así, los hiperplanos de una partición interactúan para formar regiones. Por otro lado, estos mismos hiperplanos se intersectan entre sí formando variedades lineales.

Definición 6. Sea H una partición en \mathbb{R}^n . Un hiperplano S de H , se dice que es una *intersección* de orden 1 y la denotaremos por $S^{(1)}$. Una variedad lineal de dimensión $(n - k)$, denotada por $S^{(k)}$, es una intersección de orden k si está dada por la intersección de $m \geq 2$ variedades lineales de orden $(k - 1)$, es decir

$$S^{(k)} := \bigcap_{i=1}^m S^{(k-1)}. \quad (2.2)$$

Para evitar confusiones entre las definiciones de hiperplano e intersección, nos referiremos a hiperplano cuando hablemos de intersecciones de orden 1, mientras que hablaremos de intersecciones siempre que la intersección sea de orden igual o mayor a 2. Sin embargo, seguiremos utilizando la misma notación que se ha utilizado hasta este momento para referirnos a estos dos tipos de objetos, únicamente agregaremos el índice superior para indicar el orden de la intersección cuando sea necesario. Además, se usarán subíndices para enumerar intersecciones de orden mayor.

Ejemplo 4. En la Figura 2.4a vemos que los hiperplanos $S_1^{(1)}$, $S_2^{(1)}$ y $S_3^{(1)}$ se intersectan entre sí formando una intersección $S_1^{(2)}$ de orden 2. También, en la Figura 2.4b, los hiperplanos $S_1^{(1)}$ y $S_2^{(1)}$ forman la intersección $S_1^{(2)}$; los hiperplanos $S_2^{(1)}$ y $S_3^{(1)}$ forman la intersección $S_2^{(2)}$, ambas señaladas por líneas punteadas, por último $S_1^{(2)}$ y $S_2^{(2)}$ forman la intersección $S_1^{(3)}$.

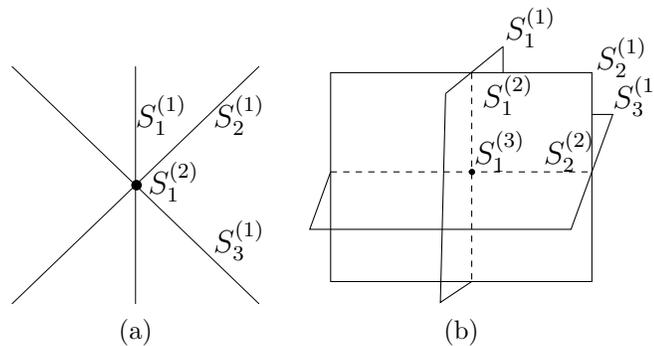


Figura 2.4: (a) Intersecciones en \mathbb{R}^2 e (b) intersecciones en \mathbb{R}^3 .

2.1.2 Funciones lineales a tramos y representación estándar.

Definición 7. Sean $D \subset \mathbb{R}^n$ y una partición H . Una función $f : D \rightarrow \mathbb{R}$ es *lineal a tramos* (PWL) sobre H si y solo si satisface las siguientes condiciones:

- D está particionado por H en un número finito de regiones R_i , $i \in \{1, \dots, k\}$, tales que

$$D := \bigcup_{i=1}^k \overline{R}_i, \quad (2.3)$$

donde \overline{R}_i es la cerradura de R_i en D .

- f es localmente representado por expresiones afines

$$f(\mathbf{x}) = \mathbf{J}_i^T \mathbf{x} + w_i, \quad \forall \mathbf{x} \in R_i, \quad (2.4)$$

donde $\mathbf{J}_i \in \mathbb{R}^n$ es el vector gradiente de f en R_i y $w_i \in \mathbb{R}$.

- Además, f es continua, esto es

$$\mathbf{J}_i^T \mathbf{x} + w_i = \mathbf{J}_j^T \mathbf{x} + w_j, \quad (2.5)$$

para todo \mathbf{x} en la intersección de la cerradura de dos regiones adyacentes R_i y R_j , $i, j \in \{1, \dots, k\}$.

Ejemplo 5. Sea H una partición dada por los hiperplanos

$$\begin{aligned} S_1 &= \{\mathbf{x} \in \mathbb{R}^2 : x_2 = 0\}, \\ S_2 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 - x_2 = 0\}, \\ S_3 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 = 0\}, \\ S_4 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 = 0\}, \end{aligned}$$

la cual particiona a \mathbb{R}^2 en regiones

$$\begin{aligned} R_1 &= S_1^{(+)} \cap S_2^{(+)} \cap S_3^{(+)} \cap S_4^{(+)}, & R_5 &= S_1^{(-)} \cap S_2^{(-)} \cap S_3^{(-)} \cap S_4^{(-)}, \\ R_2 &= S_1^{(+)} \cap S_2^{(-)} \cap S_3^{(+)} \cap S_4^{(+)}, & R_6 &= S_1^{(-)} \cap S_2^{(+)} \cap S_3^{(-)} \cap S_4^{(-)}, \\ R_3 &= S_1^{(+)} \cap S_2^{(-)} \cap S_3^{(-)} \cap S_4^{(+)}, & R_7 &= S_1^{(-)} \cap S_2^{(+)} \cap S_3^{(+)} \cap S_4^{(-)}, \\ R_4 &= S_1^{(+)} \cap S_2^{(+)} \cap S_3^{(-)} \cap S_4^{(-)}, & R_8 &= S_1^{(-)} \cap S_2^{(+)} \cap S_3^{(+)} \cap S_4^{(+)}, \end{aligned}$$

señaladas en la Figura 2.5.

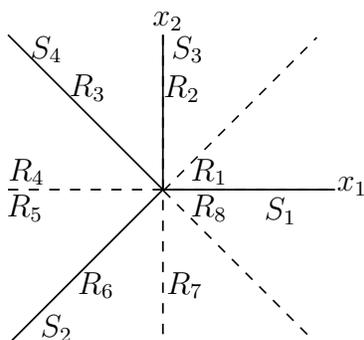


Figura 2.5: Partición H de la función PWL $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ del Ejemplo 4.

La función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

$$f(\mathbf{x}) = \begin{cases} 2x_1 + x_2 + 1 & \text{si } \mathbf{x} \in \overline{R}_1 \cup \overline{R}_2, \\ 4x_1 + x_2 + 1 & \text{si } \mathbf{x} \in \overline{R}_3, \\ \frac{5}{2}x_1 - \frac{1}{2}x_2 + 1 & \text{si } \mathbf{x} \in \overline{R}_4 \cup \overline{R}_5, \\ 2x_1 + 1 & \text{si } \mathbf{x} \in \overline{R}_6 \cup \overline{R}_7 \cup \overline{R}_8. \end{cases}$$

está bien definida y es continua. En efecto es PWL. Primero, $\mathbb{R}^2 = \bigcup_{i=1}^8 \overline{R}_i$, por lo que se cumple con (2.3). Ahora veamos que sucede en las regiones adyacentes separadas por cada uno de los hiperplanos.

- Regiones adyacentes separadas por el hiperplano S_1 : f es continua en $\overline{R}_4 \cup \overline{R}_5$; supongamos $x_2 = 0$, entonces, $f(\mathbf{x}) = 2x_1 + x_2 + 1 = 2x_1 + 1$, $\mathbf{x} \in \overline{R}_1$, además, $f(\mathbf{x}) = 2x_1 + 1$, $\mathbf{x} \in \overline{R}_8$.
- Regiones adyacentes separadas por el hiperplano S_2 : f es continua en $\overline{R}_1 \cup \overline{R}_2$; supongamos $x_1 = x_2$, entonces, $f(\mathbf{x}) = \frac{5}{2}x_1 - \frac{1}{2}x_2 + 1 = 2x_1 + 1$, $\mathbf{x} \in \overline{R}_5$, además, $f(\mathbf{x}) = 2x_1 + 1$, $\mathbf{x} \in \overline{R}_6$.
- Regiones adyacentes separadas por el hiperplano S_3 : f es continua en $\overline{R}_6 \cup \overline{R}_7$; supongamos $x_1 = 0$, entonces, $f(\mathbf{x}) = 2x_1 + x_2 + 1 = x_2 + 1$, $\mathbf{x} \in \overline{R}_2$, además, $f(\mathbf{x}) = 4x_1 + x_2 + 1 = x_2 + 1$, $\mathbf{x} \in \overline{R}_3$.
- Regiones adyacentes separadas por el hiperplano S_4 : f es continua en $\overline{R}_7 \cup \overline{R}_8$; supongamos $x_1 = -x_2$, entonces, $f(\mathbf{x}) = 4x_1 + x_2 + 1 = 3x_1 + 1$, $\mathbf{x} \in \overline{R}_3$, además, $f(\mathbf{x}) = \frac{5}{2}x_1 - \frac{1}{2}x_2 + 1 = 3x_1 + 1$, $\mathbf{x} \in \overline{R}_4$.

La condición de continuidad de una función f PWL impone una relación entre los vectores gradientes de f en regiones adyacentes y el vector normal al hiperplano que las separa, esta relación se enuncia en el siguiente lema [17].

Lema 2. Sean f una función PWL continua, \mathbf{J}_i y \mathbf{J}_j sus vectores gradientes en dos regiones adyacentes R_i y R_j , respectivamente, separadas por un hiperplano S y $\boldsymbol{\alpha}$ un vector normal a S . Existe un único $c \in \mathbb{R}$ tal que

$$\mathbf{J}_i - \mathbf{J}_j = c\boldsymbol{\alpha}. \quad (2.6)$$

Prueba. Por hipótesis $f = \mathbf{J}_i^T \mathbf{x} + w_i$ para todo $\mathbf{x} \in R_i$ y $f = \mathbf{J}_j^T \mathbf{x} + w_j$ para todo $\mathbf{x} \in R_j$, además, por la condición de continuidad

$$\mathbf{J}_i^T \mathbf{x} + w_i = \mathbf{J}_j^T \mathbf{x} + w_j, \quad (2.7)$$

para todo $\mathbf{x} \in S$.

Asumamos sin pérdida de generalidad que S contiene al vector $\mathbf{0}$ (en caso contrario basta hacer una traslación de S), por (2.7) tenemos $\mathbf{J}_i^T \mathbf{0} + w_i = \mathbf{J}_j^T \mathbf{0} + w_j$, de aquí que $w_i = w_j$, por lo que

$$(\mathbf{J}_i^T - \mathbf{J}_j^T) \mathbf{x} = \mathbf{0}$$

para todos los puntos en S . De aquí que $(\mathbf{J}_i^T - \mathbf{J}_j^T)$ es ortogonal a S , de donde se sigue la conclusión del lema. ■

Ejemplo 6. Continuando con la función dada en el Ejemplo 5, los vectores gradiente de f asociados a cada región en su dominio son:

$$\begin{aligned} \mathbf{J}_1 &= (2 \ 1)^T, & \mathbf{J}_2 &= (2 \ 1)^T, & \mathbf{J}_3 &= (4 \ 1)^T, & \mathbf{J}_4 &= \left(\frac{5}{2} \ -\frac{1}{2}\right)^T, \\ \mathbf{J}_5 &= \left(\frac{5}{2} \ -\frac{1}{2}\right)^T, & \mathbf{J}_6 &= (2 \ 0)^T, & \mathbf{J}_7 &= (2 \ 0)^T, & \mathbf{J}_8 &= (2 \ 0)^T. \end{aligned}$$

Además, los vectores

$$\boldsymbol{\alpha}_1 = (0 \ 1)^T, \quad \boldsymbol{\alpha}_2 = (1 \ -2)^T, \quad \boldsymbol{\alpha}_3 = (1 \ 0)^T, \quad \boldsymbol{\alpha}_4 = (1 \ 1)^T,$$

son normales a los hiperplanos S_1, \dots, S_4 , respectivamente.

Por lo que la relación entre cada par de vectores gradiente de regiones adyacentes con el vector normal al hiperplano que las separa está dada por:

$$\begin{aligned} \mathbf{J}_1 - \mathbf{J}_8 &= c_1 \boldsymbol{\alpha}_1, & \mathbf{J}_2 - \mathbf{J}_1 &= c_2 \boldsymbol{\alpha}_2, & \mathbf{J}_3 - \mathbf{J}_2 &= c_3 \boldsymbol{\alpha}_3, & \mathbf{J}_4 - \mathbf{J}_3 &= c_4 \boldsymbol{\alpha}_4, \\ \mathbf{J}_4 - \mathbf{J}_5 &= c_5 \boldsymbol{\alpha}_1, & \mathbf{J}_5 - \mathbf{J}_6 &= c_6 \boldsymbol{\alpha}_2, & \mathbf{J}_6 - \mathbf{J}_7 &= c_7 \boldsymbol{\alpha}_3, & \mathbf{J}_7 - \mathbf{J}_8 &= c_8 \boldsymbol{\alpha}_4, \end{aligned}$$

donde

$$c_1 = -1, \quad c_2 = 0, \quad c_3 = 2, \quad c_4 = -\frac{3}{2}, \quad c_5 = 0, \quad c_6 = \frac{1}{4}, \quad c_7 = 0, \quad c_8 = 0. \quad (2.8)$$

2.1.3 Variación consistente.

En el Ejemplo 6 se observa que se verifica el Lema 2, pero aún para pares distintos de regiones adyacentes separadas por un mismo hiperplano, los valores de c no necesariamente coinciden. Se dice que una función PWL que muestra consistentemente esta coincidencia satisface la condición dada por la siguiente definición.

Definición 8. Sea f una función continua PWL, se dice que f tiene *variación consistente* si

$$\mathbf{J}_{R_{j,1,+}} - \mathbf{J}_{R_{j,1,-}} = \dots = \mathbf{J}_{R_{j,m,+}} - \mathbf{J}_{R_{j,m,-}} = c_j \boldsymbol{\alpha}_j, \quad (2.9)$$

donde $\boldsymbol{\alpha}_j$ es un vector normal al j -ésimo hiperplano S_j en la partición H , $\mathbf{J}_{R_{j,i,+}}$ y $\mathbf{J}_{R_{j,i,-}}$ son los vectores gradiente de cada i -ésimo par de regiones distintas, adyacentes y separadas por S_j , $i \in \{1, \dots, m\}$.

Ejemplo 7. Continuando con el Ejemplo 6, vemos que esta función no cumple con la variación consistente, pues, si nos fijamos en en las regiones adyacentes, separadas por el hiperplano S_1 , es decir, los pares de regiones R_1 y R_8 ; y R_4 y R_5 , de (2.8) vemos que $c_1 = -1$ y $c_5 = 0$, por lo que $c_1 \neq c_5$, así, f no cumple con (2.9).

Ejemplo 8. Sea $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la función continua PWL con la partición H descrita en el Ejemplo 3, donde $\alpha_1 = (1 \ -1)^T$ es un vector normal al hiperplano S_1 y $\alpha_2 = (1 \ 1)^T$ es un vector normal al hiperplano S_2 y dada por

$$f(\mathbf{x}) = \begin{cases} x_1 - 2x_2 + 4 & \text{si } \mathbf{x} \in \overline{R_1}, \\ -x_1 + 2 & \text{si } \mathbf{x} \in \overline{R_2}, \\ 3x_1 + 4x_2 - 2 & \text{si } \mathbf{x} \in \overline{R_3}, \\ 5x_1 + 2x_2 & \text{si } \mathbf{x} \in \overline{R_4}. \end{cases}$$

Los vectores \mathbf{J}_i asociados a cada región R_i , $i \in \{1, \dots, 4\}$, son los siguientes:

$$\mathbf{J}_1 = (1 \ -2)^T, \quad \mathbf{J}_2 = (-1 \ 0)^T, \quad \mathbf{J}_3 = (3 \ 4)^T \text{ y } \mathbf{J}_4 = (5 \ 2)^T.$$

Por lo que la relación entre cada par de vectores gradiente de regiones adyacentes con el vector normal al hiperplano que las separa está dada por:

$$\mathbf{J}_2 - \mathbf{J}_1 = c_1 \alpha_1, \quad \mathbf{J}_3 - \mathbf{J}_2 = c_2 \alpha_2, \quad \mathbf{J}_3 - \mathbf{J}_4 = c_3 \alpha_1, \quad \mathbf{J}_4 - \mathbf{J}_1 = c_4 \alpha_2.$$

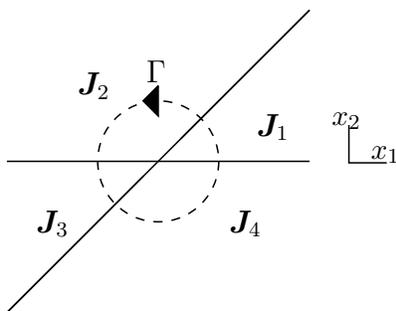
donde

$$c_1 = -1, \quad c_2 = 1, \quad c_3 = -1, \quad c_4 = 1. \quad (2.10)$$

Si nos fijamos en la relación entre las regiones adyacentes, separadas por el hiperplano S_1 , es decir, los pares de regiones R_1 y R_2 ; y R_3 y R_4 , de (2.10) vemos que $c_1 = c_3$, además, si nos fijamos en la relación entre las regiones adyacentes, separadas por el hiperplano S_2 , es decir, los pares de regiones R_2 y R_3 ; y R_4 y R_1 , de (2.10) vemos que $c_2 = c_4$, por lo que f cumple con (2.9), así, f tiene variación consistente.

Con esto podemos señalar que no todas las funciones PWL tienen variación consistente y verificar esta condición exige un análisis de cada una de las regiones adyacentes a cada hiperplano.

A partir de aquí podemos extender la noción de continuidad dada anteriormente, para ello veamos lo siguiente. Sea $\Gamma : [0, 1] \rightarrow \mathbb{R}^n$ la parametrización de un camino cerrado en el dominio de una función continua PWL f (ver Figura 2.6). A lo largo de este camino veremos que la función gradiente de f es constante al interior de cada una de las regiones que Γ atraviesa. Pensemos en esto como el comportamiento de la función $T_\Gamma : [0, 1] \rightarrow \mathbb{R}^n$ tal que $T_\Gamma(s) = \left. \frac{\partial f}{\partial \mathbf{x}_i} \right|_{\mathbf{x}=\Gamma(s)}$, T_Γ será constante a tramos y tendrá saltos de tamaño $c\alpha$. Puesto que $T_\Gamma(0) = T_\Gamma(1)$, la suma de los saltos en T_Γ debe ser cero. A esta propiedad la llamaremos propiedad del camino cerrado [17].

Figura 2.6: Camino cerrado Γ en \mathbb{R}^2 .

2.1.4 Función salto.

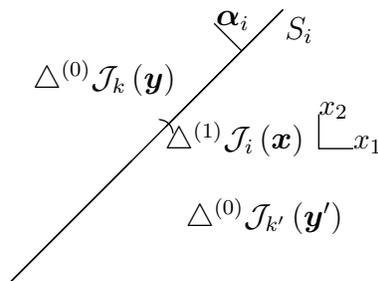
Como se sugiere en la idea de un camino cerrado, el salto entre regiones adyacentes separadas por un hiperplano en una partición de una función continua PWL se puede establecer como la diferencia entre los gradientes de las funciones asociadas a estas regiones, evaluados en puntos equidistantes al hiperplano que los separa.

Definición 9. Sea f una función continua PWL con una partición H y S_i un i -ésimo hiperplano que separa dos regiones adyacentes R_k y $R_{k'}$, sea α_i un vector normal a S_i . Para todo $\mathbf{x} \in S_i$ donde existe un $\epsilon > 0$ tal que, dado un $\delta \in (0, \epsilon)$, los puntos $\mathbf{y} = \mathbf{x} + \delta\alpha$ y $\mathbf{y}' = \mathbf{x} - \delta\alpha$ son puntos en R_k y $R_{k'}$, respectivamente, los saltos de primer orden en $S_i^{(1)}$ están dados por

$$\Delta^{(1)} \mathcal{J}_i(\mathbf{x}) := \Delta^{(0)} \mathcal{J}_k(\mathbf{y}) - \Delta^{(0)} \mathcal{J}_{k'}(\mathbf{y}'), \quad (2.11)$$

donde $\Delta^{(0)} \mathcal{J}_k(\mathbf{x}) := \mathbf{J}_k^T \mathbf{x}$, $\mathbf{x} \in R_k$.

Ejemplo 9. En la Figura 2.7 podemos ver que el hiperplano S_i en \mathbb{R}^2 separa dos regiones R_k y $R_{k'}$ las cuales tienen asociada una función salto de orden 0 colocada sobre la región a la cual hacen referencia. A su vez se encuentra señalada la función salto de orden uno en el hiperplano S_i , en la dirección del vector α_i , dada por $\Delta^{(1)} \mathcal{J}_i(\mathbf{x}) = \Delta^{(0)} \mathcal{J}_k(\mathbf{y}) - \Delta^{(0)} \mathcal{J}_{k'}(\mathbf{y}')$.

Figura 2.7: Ejemplo gráfico de una función salto de orden 1 en \mathbb{R}^2 .

Estos saltos de primer orden son constantes a tramos, es decir, tienen el mismo valor hasta que cruzan por una intersección de orden mayor, lo que nos da una noción de una función salto de orden mayor. Sea $S_i^{(l)}$ una intersección la cual contiene a una intersección $S_j^{(l+1)}$, por definición de intersección (Definición 6), $S_j^{(l+1)}$ es un conjunto afín, de dimensión $(n - l - 1)$, $k > l$, del conjunto afín $S_i^{(l)}$ de dimensión $(n - l)$, por lo que su complemento ortogonal es de dimensión 1. Sea $\{\sigma \boldsymbol{\eta}_{i,j}^{(l)} : \sigma \in \mathbb{R}, \boldsymbol{\eta}_{i,j}^{(l)} \in S_i^{(l)}\}$ el complemento ortogonal de $S_j^{(l+1)}$ en $S_i^{(l)}$.

Definición 10. Sea $S_j^{(l)}$ un subconjunto propio de $S_i^{(l-1)}$ y $\boldsymbol{\eta}_{i,j}^{(l-1)}$ un elemento en el complemento ortogonal de $S_j^{(l)}$ en $S_i^{(l-1)}$. Sea $\mathbf{x} \in S_j^{(l)}$, tal que existe un $\epsilon > 0$ tal que para todo $\delta \in (0, \epsilon)$, los puntos $\mathbf{y} = \mathbf{x} + \delta \boldsymbol{\eta}_{i,j}^{(l-1)}$ y $\mathbf{y}' = \mathbf{x} - \delta \boldsymbol{\eta}_{i,j}^{(l-1)}$ son de regiones adyacentes separadas por $S_j^{(l)}$. La función salto $\Delta^{(l)} \mathcal{J}_{\mu_l}(\mathbf{x})$ de orden l en $S_j^{(l)}$ está dada por

$$\Delta^{(l)} \mathcal{J}_{\mu_l}(\mathbf{x}) := \Delta^{(l-1)} \mathcal{J}_{\mu_{l-1}}(\mathbf{y}) - \Delta^{(l-1)} \mathcal{J}_{\mu_{l-1}}(\mathbf{y}'), \quad (2.12)$$

donde $\mu_{l-1} := \mu_{l-2}, i$ y $\mu_l := \mu_{l-1}, j$.

Para evitar confusiones, aclaramos que la función μ_l es una función que permite anidar l número de índices. De esta forma se indican los saltos en las intersecciones de orden menor al salto de orden l .

Ejemplo 10. En la Figura 2.8, la función salto $\Delta^{(2)} \mathcal{J}_{1,1}(\mathbf{x})$ de orden 2 en la intersección $S_1^{(2)}$ se forma por la diferencia entre la función salto de orden 1 evaluada en un punto en $S_1^{(1)}$ en la dirección de $\boldsymbol{\eta}_{1,1}^{(1)}$ y la función salto de orden 1 evaluada en un punto en $S_1^{(1)}$ en la dirección de $-\boldsymbol{\eta}_{1,1}^{(1)}$, es decir, $\Delta^{(2)} \mathcal{J}_{1,1}(\mathbf{x}) = \Delta^{(1)} \mathcal{J}_1(\mathbf{y}) - \Delta^{(1)} \mathcal{J}_1(\mathbf{y}')$.

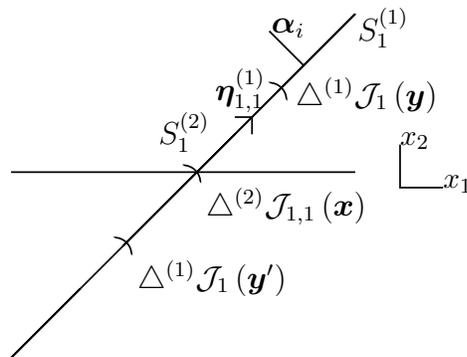


Figura 2.8: Saltos de orden 1 y 2 en el plano.

2.1.5 Propiedad de degeneración mínima.

Como se destaca en [20], no en toda intersección $S^{(k)}$ de orden k se puede encontrar una función salto de orden k tal que $\Delta^{(k)} \mathcal{J}_{\mu_i}(\mathbf{x}) \neq 0$.

Definición 11. Toda intersección de orden k donde se pueda encontrar una función salto $\Delta^{(k)} \mathcal{J}_{\mu_k}(\mathbf{x}) \neq 0$ es llamada *degenerada*. Además, se dice que una intersección degenerada de orden k es de *degeneración mínima*, y se denota por $\hat{S}^{(k)}$, si esta puede expresarse como la intersección de algún número l , y no de menos, de intersecciones $\hat{S}^{(l)}$, $1 \leq l < k$.

En [20] se señala una relación que existe entre la propiedad de variación consistente y la degeneración mínima. Para todas las funciones continuas PWL con dominio en \mathbb{R} , todas sus intersecciones tienen variación consistente, pero también, para todo par de regiones adyacentes se tiene $\Delta^{(1)} \mathcal{J}_{\mu(1)}(\mathbf{x}) \neq 0$ y esto hace que todo hiperplano sea de degeneración mínima [20]. Cuando la dimensión del espacio de salida de una función continua PWL es igual a 2, en [17], se probó que se necesitan de tres hiperplanos intersectándose para formar una intersección con degeneración mínima $\hat{S}^{(2)}$. La estructura de una intersección de degeneración mínima de orden mayor es ampliamente estudiada en [20], donde se establece el siguiente lema.

Lema 3. Una intersección $\hat{S}^{(k)}$, con $k \geq 3$, está formada por $k + 1$ intersecciones de degeneración mínima de orden $k - 1$. En general, está formada por

$$N_k(l) := \binom{k+1}{l+1} \quad (2.13)$$

intersecciones con degeneración mínima de orden l , para $1 \leq l < k$.

Prueba. Ver [20]. ■

De (2.13) obtenemos la identidad

$$N_k(l) = N_{k-1}(l) + N_{k-1}(l-1). \quad (2.14)$$

Si definimos $N_k := N_k(1)$, por (2.14), tenemos $N_k = N_{k-1}(1) + N_{k-1}(1-1) = N_{k-1} + \binom{k-1+1}{0+1} = N_{k-1} + \binom{k}{1}$, por lo que

$$N_k = N_{k-1} + k. \quad (2.15)$$

Ejemplo 11. Consideremos una función continua PWL $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ sobre una partición H formada por los hiperplanos (ver Figura 2.9):

$$\begin{aligned} \hat{S}_1 &= \{\mathbf{x} \in \mathbb{R}^3 : 2x_1 = 0\}, \\ \hat{S}_2 &= \{\mathbf{x} \in \mathbb{R}^3 : 3x_2 = 0\}, \\ \hat{S}_3 &= \{\mathbf{x} \in \mathbb{R}^3 : -8x_1 + 2x_2 + 4x_3 = 0\}, \\ \hat{S}_4 &= \{\mathbf{x} \in \mathbb{R}^3 : 2x_1 - x_2 = 0\}, \\ \hat{S}_5 &= \{\mathbf{x} \in \mathbb{R}^3 : -6x_1 + x_2 + 2x_3 = 0\}, \\ \hat{S}_6 &= \{\mathbf{x} \in \mathbb{R}^3 : -2x_1 + x_3 = 0\}. \end{aligned} \quad (2.16)$$

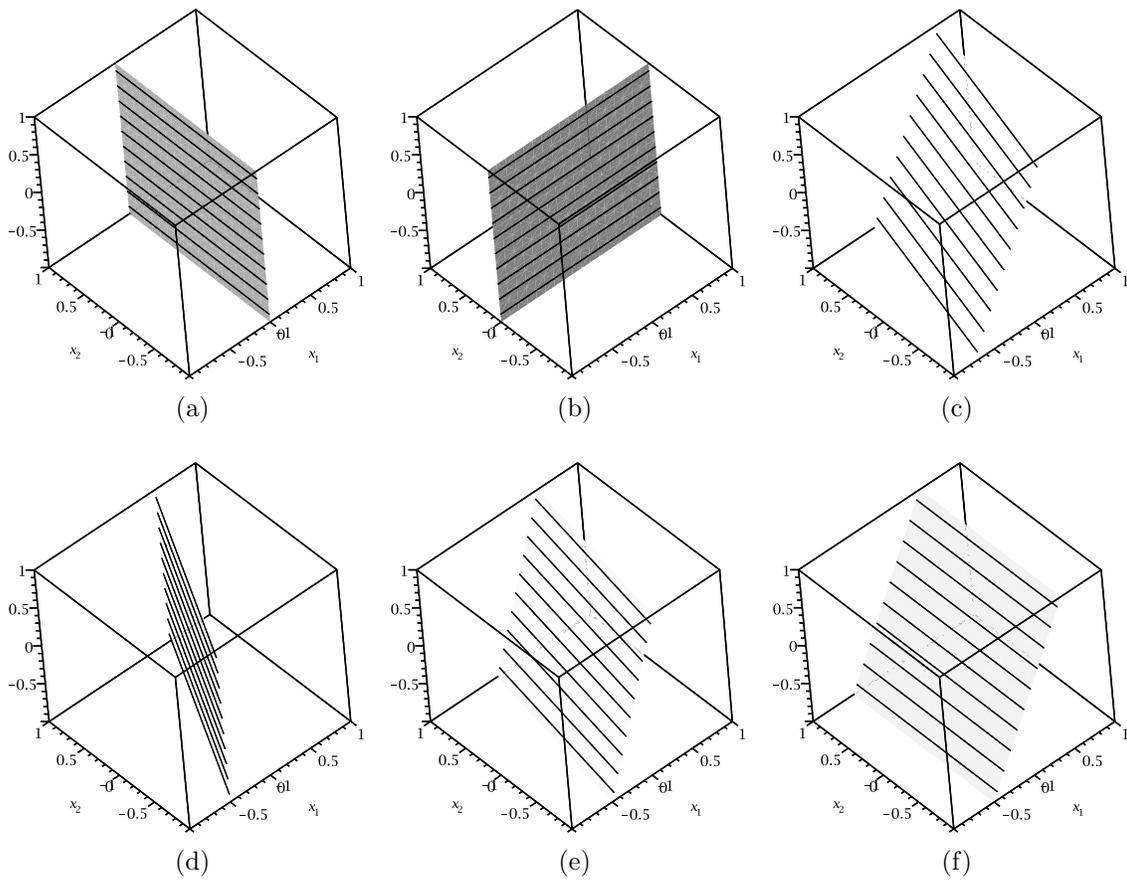


Figura 2.9: (a - f) Intersecciones (o hiperplanos) de orden 1, $\hat{S}_1, \dots, \hat{S}_6$, respectivamente de la partición H dada para el Ejemplo 11.

Se verá que la intersección de estos 6 semiplanos es una intersección de degeneración mínima de orden 3, que a su vez está formada por 6 intersecciones de degeneración mínima de orden 1 de acuerdo con la ecuación $N_3 = \binom{3+1}{1+1} = \binom{4}{2} = 6$ dada en (2.13).

Además, de acuerdo con (2.13), una intersección de degeneración mínima de orden 3 puede expresarse como una intersección de $N_3(2) = \binom{3+1}{2+1} = \binom{4}{3} = 4$ intersecciones de degeneración mínima de orden 2. Así, veremos que de los 6 hiperplanos en H se forman 4 intersecciones de degeneración mínima de orden 2. Como se mencionó anteriormente, una intersección de degeneración mínima de orden 2 puede expresarse como una intersección de 3 intersecciones de degeneración mínima de orden 1, es decir, por la intersección de 3 hiperplanos de H , además, una intersección de orden 2, en un espacio de dimensión 3, es una variedad lineal de dimensión 1, descrita por $\{\sigma\zeta : \sigma \in \mathbb{R}, \zeta \in \mathbb{R}^n\}$, así, para definir las intersecciones de degeneración mínima de orden 2 que formen una intersección de degeneración mínima de orden 3 basta con encontrar los vectores ζ que den solución a sistemas formados por tres hiperplanos en H :

- Supongamos $\zeta = (0 \ 0 \ 1)^T$. Se puede ver que ζ da solución al sistema

$$\begin{aligned} 2x_1 &= 0, \\ 3x_2 &= 0, \\ 2x_1 - x_2 &= 0. \end{aligned}$$

Así, $\sigma\zeta \in \hat{S}_1 \cap \hat{S}_2 \cap \hat{S}_4$, $\sigma \in \mathbb{R}$, por lo que podemos suponer que $\hat{S}_1^{(2)} = \hat{S}_1 \cap \hat{S}_2 \cap \hat{S}_4$ (ver Figura 2.10a) sea de degeneración mínima.

- Supongamos $\zeta = (0 \ -2 \ 1)^T$. Se puede ver que ζ da solución al sistema

$$\begin{aligned} 2x_1 &= 0, \\ -8x_1 + 2x_2 + 4x_3 &= 0, \\ -6x_1 + x_2 + 2x_3 &= 0. \end{aligned}$$

Así, $\sigma\zeta \in \hat{S}_1 \cap \hat{S}_3 \cap \hat{S}_5$, $\sigma \in \mathbb{R}$, por lo que podemos suponer que $\hat{S}_2^{(2)} = \hat{S}_1 \cap \hat{S}_3 \cap \hat{S}_5$ (ver Figura 2.10b) sea de degeneración mínima.

- Supongamos $\zeta = (1 \ 0 \ 2)^T$. Se puede ver que ζ da solución al sistema

$$\begin{aligned} 3x_2 &= 0, \\ -8x_1 + 2x_2 + 4x_3 &= 0, \\ -2x_1 + x_3 &= 0. \end{aligned}$$

Así, $\sigma\zeta \in \hat{S}_2 \cap \hat{S}_3 \cap \hat{S}_6$, $\forall \sigma \in \mathbb{R}$, por lo que podemos suponer que $\hat{S}_3^{(2)} = \hat{S}_2 \cap \hat{S}_3 \cap \hat{S}_6$ (ver Figura 2.10c) sea de degeneración mínima.

- Supongamos $\zeta = (1 \ 2 \ 2)^T$. Se puede ver que ζ da solución al sistema

$$\begin{aligned} 2x_1 - x_2 &= 0, \\ -6x_1 + x_2 + 2x_3 &= 0, \\ -2x_1 + x_3 &= 0. \end{aligned}$$

Así, $\sigma\zeta \in \hat{S}_4 \cap \hat{S}_5 \cap \hat{S}_4$, $\forall \sigma \in \mathbb{R}$, por lo que podemos suponer que $\hat{S}_4^{(2)} = \hat{S}_4 \cap \hat{S}_5 \cap \hat{S}_4$ (ver Figura 2.10c) sea de degeneración mínima.

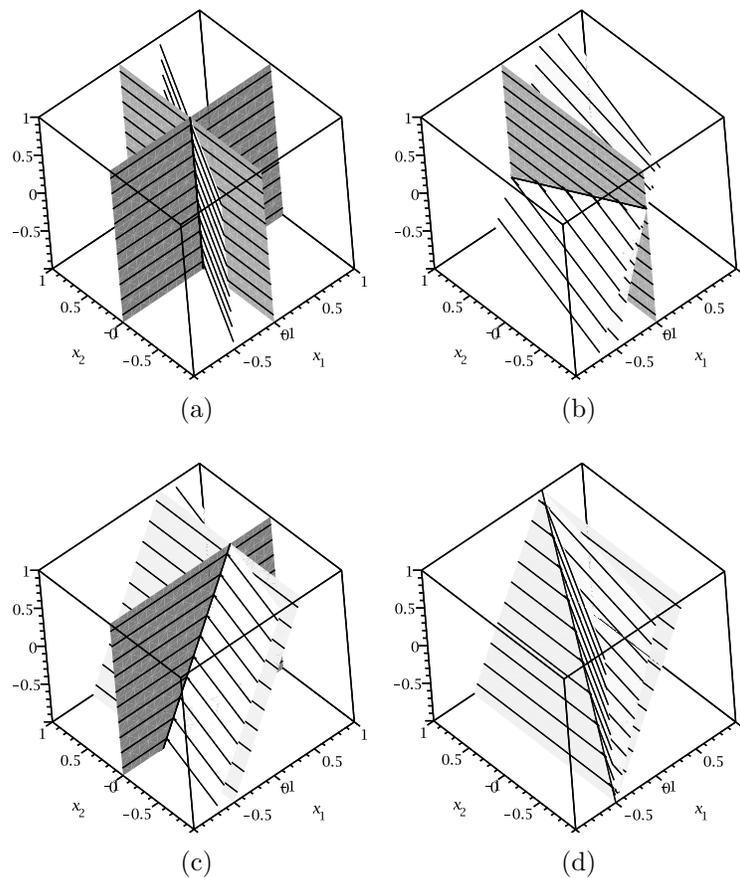


Figura 2.10: (a - d) Intersecciones $\hat{S}_1^{(2)}, \dots, \hat{S}_4^{(2)}$, respectivamente (línea negra sólida), e hiperplanos, de la partición H , que las forman (superficie blanca rayada). Para el Ejemplo 11.

Así, considerando $\hat{S}_1^{(3)} = \bigcap_{i=1}^4 \hat{S}_i^{(2)} = \bigcap_{i=1}^6 \hat{S}_i^{(1)}$ (ver Figura 2.11), podemos suponer que esta sea una intersección de degeneración mínima, pues cumple con las condiciones descritas en el Lema 3.

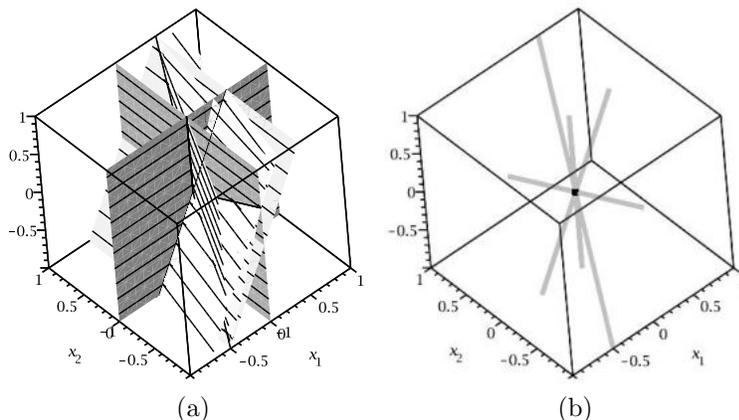


Figura 2.11: (a) Hiperplanos en H que forman a la intersección de degeneración mínima $\hat{S}_1^{(3)}$ de orden 3. (b) Intersecciones de degeneración mínima de orden 2 que forman a la intersección de degeneración mínima $\hat{S}_1^{(3)}$ de orden 3. Ejemplo 11.

La relación (2.15) establece que para formar una intersección con degeneración mínima $\hat{S}^{(k)}$, se necesitan el mismo número de hiperplanos que se necesitan para formar una intersección de degeneración mínima $\hat{S}^{(k-1)}$ más k hiperplanos, todos intersectándose en $\hat{S}^{(k)}$. Por esto, en [20] se establece el siguiente lema.

Lema 4. *Sea $\hat{S}^{(k)}$ una intersección con degeneración mínima de orden k , formada por $(k+1)$ intersecciones con degeneración mínima de orden $(k-1)$. Dada una de estas intersecciones a la cual denominamos por $\hat{S}_{\mu_{k-1}}^{(k-1)}$ formada de la intersección de los hiperplanos $\hat{S}_1^{(1)}, \dots, \hat{S}_{N_{k-1}}^{(1)}$, con vectores normales $\omega_1, \dots, \omega_{N_{k-1}}$, respectivamente. Sea α_l , $l \in \{1, \dots, k\}$ un vector normal al l -ésimo hiperplano de k hiperplanos que se intersectan con $\hat{S}_{\mu_{k-1}}^{(k-1)}$ para formar a $\hat{S}^{(k)}$, entonces, todo vector ω_i , $i \in \{1, \dots, N_k\}$ puede escribirse como una combinación lineal de algún par de vectores α_j y α_q , $j, q \in \{1, \dots, k\}$, $j \neq q$.*

Prueba. Ver [20]. ■

En [20] se señala que la elección de $\hat{S}_{\mu_{k-1}}^{(k-1)}$ no supone ninguna pérdida de generalidad, es decir, el Lema 4 funciona para cualquier intersección con degeneración mínima de orden $k-1$ en $\hat{S}^{(k)}$.

Ejemplo 12. Continuando con la función dada en el Ejemplo 11. De la elección de $\hat{S}_i^{(2)}$, $i \in \{1, \dots, 4\}$ se desprenden 4 casos de los cuales solo tomaremos 2 como ejemplo.

i) Dada la intersección $\hat{S}_1^{(2)}$ entonces podemos seleccionar

$$\alpha_1 = (-8 \ 2 \ 4)^T, \quad \alpha_2 = (-6 \ 1 \ 2)^T \text{ y } \alpha_3 = (-2 \ 0 \ 1)^T,$$

como vectores normales a los hiperplanos \hat{S}_3 , \hat{S}_5 y \hat{S}_6 , respectivamente. Además, escogemos los vectores

$$\boldsymbol{\omega}_1 = (2 \ 0 \ 0)^T, \quad \boldsymbol{\omega}_2 = (0 \ 3 \ 0)^T \text{ y } \boldsymbol{\omega}_3 = (2 \ -1 \ 0)^T,$$

normales a los hiperplanos \hat{S}_1 , \hat{S}_2 y \hat{S}_4 , respectivamente. Por Lema, 4, existen $c_{i,j}$, $0 \leq i, j, \leq 3$, tales que

$$\boldsymbol{\omega}_1 = c_{1,1}\boldsymbol{\alpha}_1 + c_{1,2}\boldsymbol{\alpha}_2, \quad \boldsymbol{\omega}_2 = c_{2,1}\boldsymbol{\alpha}_1 + c_{2,3}\boldsymbol{\alpha}_3, \quad \boldsymbol{\omega}_3 = c_{3,2}\boldsymbol{\alpha}_2 + c_{3,3}\boldsymbol{\alpha}_3.$$

Lo cual se verifica en efecto con los valores; $c_{1,1} = \frac{1}{2}$ y $c_{1,2} = -1$; $c_{2,1} = \frac{3}{2}$ y $c_{2,3} = -6$; $c_{3,2} = -1$ y $c_{3,3} = 2$.

ii) Dada la intersección $\hat{S}_4^{(2)}$ entonces podemos seleccionar

$$\boldsymbol{\alpha}_1 = (2 \ 0 \ 0)^T, \quad \boldsymbol{\alpha}_2 = (0 \ 3 \ 0)^T \text{ y } \boldsymbol{\alpha}_3 = (-8 \ 2 \ 4)^T,$$

como vectores normales a los hiperplanos \hat{S}_1 , \hat{S}_2 y \hat{S}_3 , respectivamente. Además, escogemos los vectores

$$\boldsymbol{\omega}_1 = (2 \ -1 \ 0)^T, \quad \boldsymbol{\omega}_2 = (-6 \ 1 \ 2)^T \text{ y } \boldsymbol{\omega}_3 = (-2 \ 0 \ 1)^T,$$

normales a los hiperplanos \hat{S}_4 , \hat{S}_5 y \hat{S}_6 , respectivamente. Por Lema 4, existen $c_{i,j}$, $0 \leq i, j, \leq 3$, tales que

$$\boldsymbol{\omega}_1 = c_{1,1}\boldsymbol{\alpha}_1 + c_{1,2}\boldsymbol{\alpha}_2, \quad \boldsymbol{\omega}_2 = c_{2,1}\boldsymbol{\alpha}_1 + c_{2,3}\boldsymbol{\alpha}_3, \quad \boldsymbol{\omega}_3 = c_{3,2}\boldsymbol{\alpha}_2 + c_{3,3}\boldsymbol{\alpha}_3.$$

Lo cual se verifica en efecto con los valores; $c_{1,1} = 1$ y $c_{1,2} = -\frac{1}{3}$; $c_{2,1} = -1$ y $c_{2,3} = \frac{1}{2}$; $c_{3,2} = -\frac{1}{6}$ y $c_{3,3} = \frac{1}{4}$.

En [19] se establece una forma de representar a todos los vectores $\boldsymbol{\alpha}_i$, $1 \in \{1, \dots, k\}$, de modo que estos cumplan con el Lema 4 manteniendo uniformidad en los coeficientes que los multiplican como se enuncia a continuación.

Lema 5. *Sea $\hat{S}^{(k)}$ una intersección con degeneración mínima de orden k , la cual está formada por $(k+1)$ intersecciones con degeneración mínima de orden $(k-1)$, dada una de estas intersecciones a la cual denominamos por $\hat{S}_{\mu_{k-1}}^{(k-1)}$ la cual se forma de la intersección de los hiperplanos $\hat{S}_1^{(1)}, \dots, \hat{S}_{N_{k-1}}^{(1)}$, los cuales tienen vectores normales $\boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_{N_{k-1}}$, respectivamente y sean $\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_k$ vectores normales uno a uno al resto de los hiperplanos que se intersectan en $\hat{S}^{(k)}$, entonces, todo vector $\boldsymbol{\omega}_l$, $l \in \{1, \dots, N_{k-1}\}$, es de la forma*

$$\boldsymbol{\omega}_l = \rho_l (a_i \boldsymbol{\alpha}_i - a_j \boldsymbol{\alpha}_j), \quad (2.17)$$

donde $i, j \in \{1, \dots, k\}$, $i \neq j$.

Prueba. Por Lema 4, todo $\boldsymbol{\omega}_l$, $l \in \{1, \dots, k-1\}$ puede ser escrito $\boldsymbol{\omega}_l = c_{l,1}\boldsymbol{\alpha}_1 + c_{l,l+1}\boldsymbol{\alpha}_{l+1}$, suponga

$$a_1 = 1, \quad \rho_1 = c_{1,1} \quad \text{y} \quad a_2 = -\frac{c_{1,2}}{\rho_1}.$$

Más aún, suponga

$$\rho_l = \frac{c_{l,1}}{a_1} \quad \text{y} \quad a_{l+1} = -\frac{c_{l,l+1}}{\rho_l},$$

con $l \in \{2, \dots, k-1\}$. Con esta elección de coeficientes podemos escribir

$$\begin{aligned} \boldsymbol{\omega}_1 &= \rho_1 (a_1\boldsymbol{\alpha}_1 - a_2\boldsymbol{\alpha}_2), \\ &\vdots \\ \boldsymbol{\omega}_{k-1} &= \rho_{k-1} (a_1\boldsymbol{\alpha}_1 - a_k\boldsymbol{\alpha}_k). \end{aligned} \tag{2.18}$$

Sea $\boldsymbol{\omega}_p$, $p \in \{k-1, \dots, N_{k-1}\}$, el cual asumiremos, por el Lema 4, que es de la forma

$$\boldsymbol{\omega}_p = c_{p,i}\boldsymbol{\alpha}_i + c_{p,j}\boldsymbol{\alpha}_j,$$

con $i, j \in \{1, \dots, k\}$, $i \neq j$. Puesto que $\boldsymbol{\omega}_p^T \boldsymbol{x} = 0$, $\boldsymbol{x} \in \hat{S}_{\mu_{k-1}}^{(k-1)}$, tenemos

$$c_{p,i}\boldsymbol{\alpha}_i^T \boldsymbol{x} = -c_{p,j}\boldsymbol{\alpha}_j^T \boldsymbol{x}, \tag{2.19}$$

además,

$$\begin{aligned} a_1\boldsymbol{\alpha}_1\boldsymbol{x} &= a_i\boldsymbol{\alpha}_i\boldsymbol{x}, \\ a_1\boldsymbol{\alpha}_1\boldsymbol{x} &= a_j\boldsymbol{\alpha}_j\boldsymbol{x}, \end{aligned}$$

por lo que

$$a_i\boldsymbol{\alpha}_i\boldsymbol{x} = a_j\boldsymbol{\alpha}_j\boldsymbol{x}. \tag{2.20}$$

De (2.19) y (2.20) tenemos $c_{p,i} = ma_i$ y $-c_{p,j} = ma_j$, así $m = c_{p,i}/a_i = -c_{p,j}/a_j$, haciendo $m = \rho_k$, por lo que $\boldsymbol{\omega}_k = \rho_k(a_i\boldsymbol{\alpha}_i - a_j\boldsymbol{\alpha}_j)$. ■

De la prueba del Lema 5 se establece que una posible elección de coeficientes que cumplan con la relación (2.17) es:

$$\begin{cases} a_1 = 1, \\ a_{l+1} = -\frac{c_{l,l+1}}{\rho_l}, \quad \rho_l = \frac{c_{l,1}}{a_1} & \text{si } 1 \leq l \leq k-1, \\ \rho_l = \frac{c_{l,i}}{a_i} = -\frac{c_{l,j}}{a_j} & \text{si } k-1 < l \leq N_{k-1}. \end{cases} \tag{2.21}$$

Ejemplo 13. Continuando con el inciso ii) del Ejemplo 12, los vectores asignados

$$\boldsymbol{\alpha}_1 = (2 \ 0 \ 0)^T, \quad \boldsymbol{\alpha}_2 = (0 \ 3 \ 0)^T \quad \text{y} \quad \boldsymbol{\alpha}_3 = (-8 \ 2 \ 4)^T,$$

son normales a los hiperplanos \hat{S}_1 , \hat{S}_2 y \hat{S}_3 , respectivamente y los vectores

$$\boldsymbol{\omega}_1 = (2 \ -1 \ 0)^T, \quad \boldsymbol{\omega}_2 = (-6 \ 1 \ 2)^T \text{ y } \quad \boldsymbol{\omega}_3 = (-2 \ 0 \ 1)^T,$$

son normales a los hiperplanos \hat{S}_4 , \hat{S}_5 y \hat{S}_6 , por lo que la elección de coeficientes

$$\begin{aligned} a_1 &= 1, \quad \rho_1 = \frac{c_{1,1}}{a_1} = \frac{1}{1} = 1, \\ a_2 &= -\frac{c_{1,2}}{\rho_1} = -\frac{\frac{1}{3}}{1} = -\frac{1}{3}, \quad \rho_2 = \frac{c_{2,1}}{a_1} = \frac{-1}{1} = -1, \\ a_3 &= -\frac{c_{2,3}}{\rho_2} = -\frac{\frac{1}{2}}{-1} = \frac{1}{2}, \text{ y } \rho_3 = \frac{c_{3,2}}{a_2} = \frac{-\frac{1}{6}}{-\frac{1}{3}} = -\frac{1}{2}, \end{aligned}$$

satisfacen (2.17), es decir,

$$\begin{aligned} \boldsymbol{\omega}_1 &= \rho_1 (a_1 \boldsymbol{\alpha}_1 - a_2 \boldsymbol{\alpha}_2), \\ \boldsymbol{\omega}_2 &= \rho_2 (a_1 \boldsymbol{\alpha}_1 - a_3 \boldsymbol{\alpha}_3), \\ \boldsymbol{\omega}_3 &= \rho_3 (a_2 \boldsymbol{\alpha}_2 - a_3 \boldsymbol{\alpha}_3). \end{aligned}$$

2.1.6 Otras propiedades de funciones PWL.

Ahora que se han revisado las principales propiedades relacionadas con el dominio de funciones continuas PWL que se usarán a lo largo de este trabajo, veremos dos propiedades relacionadas a la interacción entre funciones continuas PWL.

Lema 6. Sean $f_1 : D \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, donde cada componente de f_1 es una función PWL con partición H_1 , y $f_2 : \mathbb{R}^m \rightarrow \mathbb{R}$ una función PWL, entonces $f(\mathbf{x}) = f_2 \circ f_1$ es también una función PWL en D .

Prueba. Ver [25]. ■

Lema 7. Sean f_1 y f_2 funciones continuas PWL con un mismo dominio D y partición H_1 y H_2 , respectivamente. Dados $\sigma_1, \sigma_2 \in \mathbb{R}$, la combinación lineal

$$f(\mathbf{x}) = \sigma_1 f_1(\mathbf{x}) + \sigma_2 f_2(\mathbf{x}),$$

con $\mathbf{x} \in D$, es también una función continua PWL, con partición

$$H = H_1 \cup H_2.$$

Prueba. Ver [25]. ■

Sea $PWL_H[D]$ el conjunto de funciones continuas PWL $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, definidas en D con partición lineal H . Dados $f, g \in PWL_H[D]$ y $r \in \mathbb{R}$, definimos, respectivamente, la adición y la multiplicación por escalar en elementos de este conjunto de la siguiente manera:

$$\begin{aligned} (f + g)(\mathbf{x}) &= f(\mathbf{x}) + g(\mathbf{x}), \quad \forall \mathbf{x} \in D, \\ (r \cdot f)(\mathbf{x}) &= r \cdot f(\mathbf{x}), \quad \forall \mathbf{x} \in D. \end{aligned} \tag{2.22}$$

Teorema 1. El conjunto $PWL_H[D]$ con la adición y multiplicación por escalar definidos en (2.22) es un espacio vectorial lineal.

Prueba. Ver [22]. ■

2.2 Representación canónica lineal a tramos.

Como se vio anteriormente, la representación estándar de funciones continuas PWL, de acuerdo a la Definición 7, requiere de cierta información que computacionalmente se tiene que almacenar, como lo es la representación de los hiperplanos en la partición, la descripción de las regiones que estos determinan y el valor de la función en cada región. Con el objetivo de proporcionar una representación de funciones PWL compacta en comparación con la estándar, en [2] se presenta la representación CPWL como se define a continuación.

Definición 12. Sea f una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$, se dice que f tiene una representación *canónica lineal a tramos* (CPWL) si puede ser escrita de la forma

$$f(\mathbf{x}) = a + \mathbf{b}^T \mathbf{x} + \sum_{i=1}^h c_i |\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i|, \quad (2.23)$$

donde $a, c_i, \beta_i \in \mathbb{R}$ y $\mathbf{b}, \boldsymbol{\alpha}_i, \mathbf{x} \in \mathbb{R}^n$.

Ejemplo 14. Definimos la función $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que

$$f = 1 + 2x_1 + x_2 + |x_1 - x_2 + 1| - 2|x_1 + x_2 - 1|$$

Esta puede ser escrita de la forma (2.23), con $a = 1$, $\mathbf{b} = (2 \ 1)^T$, $\mathbf{x} = (x_1 \ x_2)^T$, $h = 2$, $c_1 = 1$, $c_2 = -2$, $\boldsymbol{\alpha}_1 = (1 \ -1)^T$, $\beta_1 = -1$, $\boldsymbol{\alpha}_2 = (1 \ 1)^T$ y $\beta_2 = 1$.

A esta representación se le llama canónica lineal a tramos, sin embargo, hasta este punto no se ha probado que funciones de la forma (2.23) sean también PWL. Para ello vemos el siguiente lema [2].

Lema 8. *Sea f una función con representación CPWL, f es también continua PWL.*

Prueba. Sea f una función con representación CPWL, probaremos que f es PWL sobre la partición H formada por los hiperplanos

$$\{\mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}_i^T \mathbf{x} - \beta = 0\},$$

$i \in \{1, \dots, h\}$. Para obtener su expresión por región (2.4) a partir de (2.23) veamos lo siguiente. Sea R_j una región formada por los hiperplanos en H , el gradiente de f en la región R_j es

$$\mathbf{J}_j^T = \mathbf{b}^T + \sum_{i=1}^h c_i (\text{sgn}(\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i)) \boldsymbol{\alpha}_i^T, \quad (2.24)$$

para todo $\mathbf{x} \in R_j$ y el operador $\text{sgn}(\cdot)$ es la función signo tal que $\text{sgn}(a) = 1$ si $a \geq 0$ y $\text{sgn}(a) = -1$ si $a < 0$, $a \in \mathbb{R}$.

Ahora, sustituyendo (2.23) y (2.24) en (2.4) tenemos

$$\begin{aligned}
w_j &= f(\mathbf{x}) - \mathbf{J}_j^T \mathbf{x} = \\
&= a + \mathbf{b}^T \mathbf{x} + \sum_{i=1}^h c_i |\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i| - \left(\mathbf{b}^T + \sum_{k=1}^h c_k (\operatorname{sgn}(\boldsymbol{\alpha}_k^T \mathbf{x} - \beta_k)) \boldsymbol{\alpha}_k^T \right) \mathbf{x} = \\
&= a + \mathbf{b}^T \mathbf{x} + \sum_{i=1}^h c_i (\operatorname{sgn}(\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i)) (\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i) - \mathbf{b}^T \mathbf{x} - \sum_{k=1}^h c_k (\operatorname{sgn}(\boldsymbol{\alpha}_k^T \mathbf{x} - \beta_k)) \boldsymbol{\alpha}_k^T \mathbf{x} \\
&= a - \sum_{i=1}^h c_i (\operatorname{sgn}(\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i)) \beta_i,
\end{aligned}$$

así,

$$w_i = a - \sum_{i=1}^h c_i (\operatorname{sgn}(\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i)) \beta_i. \quad (2.25)$$

Así, de (2.24) y (2.25) se desprende que f es PWL sobre H . ■

De esta forma, podemos ver que todas las funciones representadas por (2.23) son continuas PWL y las funciones asociados a cada región pueden obtenerse por (2.24) y (2.25), además, en este tipo de representación los hiperplanos de la partición aparecen de forma explícita, por lo que no es necesario almacenar la información de cada región, lo cual representa un ahorro en la memoria utilizada para representar funciones PWL, esta propiedad será vista con mayor detalle en el Capítulo 3. Sin embargo, no todas las funciones continuas PWL poseen una representación CPWL. La condición para la existencia de una representación CPWL de funciones continuas PWL es dada en [16] y se enuncia a continuación.

Teorema 2. *Una función PWL con partición H tiene una representación CPWL sí y solo sí esta tiene una variación consistente.*

Prueba. Sea f una función continua PWL con variación consistente, cuya partición H se forma de los hiperplanos

$$S_i = \{\mathbf{x} \in \mathbb{R}^n : b_i (\boldsymbol{\alpha}_i^T \mathbf{x} - \beta) = 0\},$$

$i \in \{1, \dots, m\}$, y $b_i \in \{-1, 1\}$ tales que existe una región R_0 , determinada por H , dada por $R_0 = \bigcap_{i=1}^m S_i^{(+)}$.

De la Definición 8 tenemos que existe un c_i tal que

$$\mathbf{J}_{R_{i,1,+}} - \mathbf{J}_{R_{i,1,-}} = \dots = \mathbf{J}_{R_{i,n_i,+}} - \mathbf{J}_{R_{i,n_i,-}} = c_i \boldsymbol{\alpha}_i,$$

$i \in \{1, \dots, m\}$, con n_i pares de regiones adyacentes separadas por el hiperplano S_i . Probaremos que f tiene una expresión canónica

$$f(\mathbf{x}) = a + \mathbf{b}^T \mathbf{x} + \sum_{i=1}^h \frac{1}{2} c_i |\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i|.$$

Dado el anterior R_0 , tenemos que para todo $\mathbf{x} \in R_0$, $\mathbf{x} \in \{\mathbf{x} \in \mathbb{R}^n : \boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i > 0\}$, para todo $i \in \{1, \dots, m\}$.

Por (2.4) tenemos que

$$f(\mathbf{x}) = \mathbf{J}_0^T \mathbf{x} + a_0, \quad (2.26)$$

para todo $\mathbf{x} \in R_0$. Hagamos

$$\begin{aligned} \mathbf{b}^T &= \mathbf{J}_0^T - \frac{1}{2} \sum_{i=1}^m c_i \boldsymbol{\alpha}_i^T \\ a &= a_0 + \frac{1}{2} \sum_{i=1}^m \beta_i c_i, \end{aligned} \quad (2.27)$$

donde c_i está dado por la variación consistente, es decir, los c_i son tales que (2.9) se cumple. Sea \hat{f} la función obtenida al sustituir en (2.23) a \mathbf{b}^T y a dados por (2.27). Así:

$$\begin{aligned} \hat{f}(\mathbf{x}) &= a_0 - \frac{1}{2} \sum_{i=1}^m \beta_i c_i + \left(\mathbf{J}_0^T - \frac{1}{2} \sum_{i=1}^m c_i \boldsymbol{\alpha}_i^T \right) \mathbf{x} + \frac{1}{2} \sum_{i=1}^m c_i |\boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i| \\ &= a_0 + \mathbf{J}_0^T \mathbf{x} + \frac{1}{2} \sum_{i=1}^m \beta_i c_i - c_i \boldsymbol{\alpha}_i^T \mathbf{x} + c_i \boldsymbol{\alpha}_i^T \mathbf{x} - \beta_i c_i = a_0 + \mathbf{J}_0^T \mathbf{x} = f(\mathbf{x}), \end{aligned} \quad (2.28)$$

para todo $\mathbf{x} \in R_0$.

Ahora sea $R_1 = S_k^{(-)} \cap \left(\bigcap_{i=1, i \neq k}^m S_i^{(+)} \right)$ una región adyacente a R_0 , separada por un k -ésimo hiperplano, $k \in \{1, \dots, m\}$. Por (2.4), (2.24) y (2.25) tenemos que

$$\hat{f}(\mathbf{x}) = \left(a + \frac{1}{2} \beta_k c_k - \frac{1}{2} \sum_{i=1, i \neq k}^m \beta_i c_i \right) + \left(\mathbf{b}^T - \frac{1}{2} c_k \boldsymbol{\alpha}_k^T + \frac{1}{2} \sum_{i=1, i \neq k}^m c_i \boldsymbol{\alpha}_i^T \right)$$

y por (2.27),

$$\hat{f}(\mathbf{x}) = (a_0 + \beta_k c_k) + (\mathbf{J}_0^T - c_k \boldsymbol{\alpha}_k^T). \quad (2.29)$$

Por (2.6) tenemos

$$\mathbf{J}_1 = \mathbf{J}_0 - c_k \boldsymbol{\alpha}_k^T \quad (2.30)$$

y por (2.5),

$$\mathbf{J}_0^T \mathbf{x} + a_0 = \mathbf{J}_1^T \mathbf{x} + a_1, \quad (2.31)$$

para todo \mathbf{x} en la intersección de las cerraduras de R_0 y R_1 . Sustituyendo (2.30) en (2.31) tenemos

$$a_1 = a_0 + (\mathbf{J}_0^T - \mathbf{J}_1^T) \mathbf{x},$$

además, por la relación $\boldsymbol{\alpha}_k^T \mathbf{x} = \beta_k$, con $\mathbf{x} \in S_k$, se sigue que

$$a_1 = a_0 + \beta_k c_k. \quad (2.32)$$

Sustituyendo (2.31) y (2.32) en (2.29), tenemos

$$\hat{f}(\mathbf{x}) = (a_0 + \beta_k c_k) + (\mathbf{J}_k^T - c_k \boldsymbol{\alpha}_k^T) \mathbf{x} = \mathbf{J}_1^T \mathbf{x} + a_1 = f(\mathbf{x}),$$

para todo $\mathbf{x} \in R_1$.

Puesto que R_1 es arbitrario, $\hat{f}(\mathbf{x}) = f(\mathbf{x})$ para toda región adyacente a R_0 . Repitiendo el proceso para toda región adyacente de R_1 y aplicando inducción sobre la región, se tiene que $\hat{f}(\mathbf{x}) = f(\mathbf{x})$ para todo \mathbf{x} en el dominio de f . Por lo tanto, f tiene una representación CPWL.

Ahora, sea f una función con representación CPWL y partición H con m hiperplanos distintos S_1, \dots, S_m . Supongamos que S_k , $k \in \{1, \dots, m\}$, separa a q pares distintos de regiones $R_{k,p,+}$, $R_{k,p,-}$, $p \in \{1, \dots, q\}$ adyacentes. Así, $\alpha_k^T \mathbf{x} - \beta_k \geq 0$ para todo $\mathbf{x} \in R_{k,p,+}$ y $\alpha_k^T \mathbf{x} - \beta_k \leq 0$ para todo $\mathbf{x} \in R_{k,p,-}$. Por lo que podemos escribir a f como

$$f(\mathbf{x}) = a + \mathbf{b}^T \mathbf{x} + (c'_k \alpha_k^T \mathbf{x} - \beta_k d_k) \operatorname{sgn}(\alpha_k^T \mathbf{x} - \beta_k) + \sum_{i=1, i \neq k}^m c'_i |\alpha_i^T \mathbf{x} - \beta_i|,$$

y los vectores gradiente serán

$$\begin{aligned} \mathbf{J}_{k,p,+}^T &= \mathbf{b}^T + c'_k \alpha_k^T + \sum_{i=1, i \neq k}^m d_i \alpha_i^T (\operatorname{sgn}(\alpha_k^T \mathbf{x} - \beta_k)), \\ \mathbf{J}_{k,p,-}^T &= \mathbf{b}^T - c'_k \alpha_k^T + \sum_{i=1, i \neq k}^m d_i \alpha_i^T (\operatorname{sgn}(\alpha_k^T \mathbf{x} - \beta_k)), \end{aligned}$$

para todo $\mathbf{x} \in R_{k,p,+}$ y para todo $\mathbf{x} \in R_{k,p,-}$, respectivamente. Así,

$$\mathbf{J}_{k,p,+}^T - \mathbf{J}_{k,p,-}^T = c'_k \alpha_k^T - (-c'_k \alpha_k^T) = 2c'_k \alpha_k^T = c_k \alpha_k^T,$$

donde $2c'_k = c_k$. De aquí que los coeficientes para la representación CPWL,

$$c'_k = \frac{c_k}{2}, \quad (2.33)$$

donde c_k es tal que se cumple con (2.9), se obtienen por la propiedad de variación consistente. Puesto que el hiperplano k y las regiones adyacentes separadas por este son arbitrarios, f presenta variación consistente. ■

Así, para determinar los coeficientes usados para dar una representación CPWL (2.23) de una función f PWL con una partición H formada por m hiperplanos, en [2] se hacen las siguientes observaciones:

- Para obtener el vector \mathbf{b} veamos primero que, de la prueba del Teorema 2, para un \mathbf{x} en una región R de la partición H de f se tiene que

$$\mathbf{J}^T = \mathbf{b}^T + \frac{1}{2} \sum_{i=1}^m c_i \alpha_i^T \operatorname{sgn}(\alpha_i^T \mathbf{x} - \beta_i). \quad (2.34)$$

Suponga R_+ tal como se define a R_0 en la prueba del Teorema 2, se tiene que para todo \mathbf{x} en R_+ , $\operatorname{sgn}(\alpha_i^T \mathbf{x} - \beta_i) = 1$, para todo $i \in \{1, \dots, m\}$. Además, sea R_- tal que para todo \mathbf{x} en R_- , $\operatorname{sgn}(\alpha_i^T \mathbf{x} - \beta_i) = -1$, para todo $i \in \{1, \dots, m\}$. De lo anterior se sigue que

$$\mathbf{b}^T = \frac{1}{2} (\mathbf{J}_+^T + \mathbf{J}_-^T), \quad (2.35)$$

dónde \mathbf{J}_+ y \mathbf{J}_- son las funciones gradiente de f en R_+ y R_- , respectivamente.

- Para determinar los coeficientes c_k , $k \in \{1, \dots, m\}$, veamos que para dos puntos en regiones adyacentes $R_{k,+}$ y $R_{k,-}$ cuyos valores $\text{sgn}(\boldsymbol{\alpha}_j^T \mathbf{x} - \beta_j)$, $j \in \{1, \dots, m\}$, difieren únicamente en el valor correspondiente a $j = k$, si evaluamos (2.34) en algún punto de las regiones adyacentes $R_{k,+}$ y $R_{k,-}$, tenemos que

$$c_k = \pm \frac{1}{2} \boldsymbol{\alpha}_k^T (\mathbf{J}_{i,+} + \mathbf{J}_{i,-}) / \boldsymbol{\alpha}_k^T \boldsymbol{\alpha}_k, \quad (2.36)$$

donde el signo \pm es escogido si la sucesión $(\text{sgn}(\boldsymbol{\alpha}_j^T \mathbf{x} - \beta_j))$, $j \in \{1, \dots, m\}$ asociada a $R_{k,+}$ tiene un 1 en la k -ésima posición.

- Finalmente, para obtener el valor de a , de evaluar (2.23) en $\mathbf{0}$ se sigue que

$$a = f(\mathbf{0}) - \sum_{i=1}^m c_i |\beta_i|. \quad (2.37)$$

De (2.27) y (2.33) en la prueba del Teorema 2 se obtienen los vectores y coeficientes para la representación CPWL de una función PWL con variación consistente. Sin embargo, si f tiene una representación (2.23), entonces f permite distintas representaciones canónicas

$$f(\mathbf{x}) = a + \mathbf{b}^T \mathbf{x} + \sum_{i=1}^h \frac{c_i}{k} |k \boldsymbol{\alpha}_i^T \mathbf{x} - k \beta_i|,$$

con $k \neq 0$. Por esto, en [16] se da el siguiente teorema de unicidad.

Teorema 3. *Si una función continua PWL tiene una representación CPWL, esta representación es única para parámetros fijos de su partición.*

Prueba. Ver [16]. ■

2.2.1 Representación canónica lineal a tramos de alto nivel.

La variación consistente de una función PWL es una condición necesaria para la existencia de la representación CPWL de la misma. Como se mencionó [16], todas las funciones continuas PWL con un dominio en \mathbb{R} cumplen con esta propiedad. Conforme la dimensión del espacio de salida de funciones continuas PWL incrementa, también lo hace la interacción de sus intersecciones, por lo que se torna complicado verificar que cumplan con esta propiedad. Para superar esta problemática en [17] se propuso una representación basada en la anidación de funciones CPWL, la cual garantiza la existencia de una representación canónica de funciones continuas PWL con un dominio en \mathbb{R}^2 e intersecciones a lo más con degeneración mínima. A partir de aquí se introduce la idea de una representación CPWL de orden mayor, la cual se define a continuación.

Definición 13. La representación (2.23) es también llamada representación canónica de primer nivel. Dado un entero $k > 1$, diremos que una función $g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ tiene una representación canónica de k -ésimo nivel si puede ser escrita en la forma

$$g(\mathbf{x}) = g_0(\mathbf{x}) + \sum_{j=1}^l d_j |g_j(\mathbf{x})|, \quad (2.38)$$

donde $d_j \in \mathbb{R}$, $j \in \{1, \dots, l\}$, las funciones $g_m : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $m \in \{0, \dots, l\}$ son funciones con representación canónica de a lo más $(k-1)$ -ésimo nivel y al menos una función g_q , $q \in \{1, \dots, l\}$ es una función con representación canónica de $(k-1)$ -ésimo nivel.

A esta representación se le llama representación *canónica lineal a tramos de alto nivel* (HL-CPWL), una función de esta forma es también PWL. Sea \mathcal{F} el conjunto de todas las funciones continuas PWL que van de D a \mathbb{R} . En [18] se prueba lo siguiente.

Teorema 4. *Para toda función $f \in \mathcal{F}$ existe un entero $k \geq 1$ tal que f puede ser representado por una función HL-CPWL de k -ésimo nivel.*

Prueba. Ver [18]. ■

Esta prueba no provee de una construcción de dicha representación, por lo que brindar de un método que permita obtener los coeficientes de la representación HL-CPWL de todas las funciones en \mathcal{F} sigue siendo un problema abierto. En [19] se propuso un esquema de construcción de la representación HL-CPWL de funciones PWL con un dominio en \mathbb{R}^n e intersecciones a lo más con degeneración mínima. Sin embargo, como se señala en [21], el uso de particiones arbitrarias para funciones PWL resulta inconveniente para obtener un algoritmo para la construcción de su representación HL-CPWL o CPWL, puesto que para obtener dichas representaciones es necesario verificar la propiedad de degeneración mínima (para obtener una representación HL-CPWL) o de variación consistente (para obtener una representación CPWL). Es por esto que en [22] se propone por primera vez un esquema de construcción que se basa en una partición uniforme. Estos esquemas de construcción serán abordados en el siguiente capítulo.

Capítulo 3

Particiones simpliciales para la representación HL-CPWL.

Antes de continuar con este capítulo, detallaremos la forma en que se ordenarán los índices en el resto de este documento, así como la noción de símplice.

En lo sucesivo utilizaremos índices $i \in \{i_1, \dots, i_2\}$, donde $i_1, i_2 \in \mathbb{Z}$ y $i_1 \leq i_2$. La notación $i \in \{i_1, \dots, i_2\} \uparrow$ indica que el índice i toma valores incrementando de uno en uno de i_1 hasta i_2 y se llamará tupla de índices a un conjunto ordenado de índices. Además, conjuntos de tuplas se ordenarán como se enuncia en las siguientes definiciones.

Definición 14. Un conjunto de tuplas de k índices $\{r_1, \dots, r_k\}$, $r_i \in \{1, \dots, n\}$, $i \in \{1, \dots, k\}$, donde $r_1 \in \{1, \dots, n - k + 1\}$ y $r_j \in \{r_{j-1} + 1, \dots, n - k + j\}$ es A ordenado si está ordenado de la siguiente forma:

- r_1 toma valores en $\{1, \dots, n - k + 1\} \uparrow$
- Para r_1 fijo, r_2 toma valores en $\{r_1 + 1, \dots, n - k + 2\} \uparrow$
- Para r_{k-1} fijo, r_k toma valores en $\{r_{k-1} + 1, \dots, n\} \uparrow$

Ejemplo 15. Sea $n = 6$ y $k = 4$, las tuplas de índices A ordenadas, con $r_1 \in \{1, \dots, 3\} \uparrow$, $r_2 \in \{r_1 + 1, \dots, 4\} \uparrow$, $r_3 \in \{r_2 + 1, \dots, 5\} \uparrow$, $r_4 \in \{r_3 + 1, \dots, 6\} \uparrow$, son: $\{1, 2, 3, 4\}$, $\{1, 2, 3, 5\}$, $\{1, 2, 3, 6\}$, $\{1, 2, 4, 5\}$, $\{1, 2, 4, 6\}$, $\{1, 2, 5, 6\}$, $\{1, 3, 4, 5\}$, $\{1, 3, 4, 6\}$, $\{1, 3, 5, 6\}$, $\{1, 4, 5, 6\}$, $\{2, 3, 4, 5\}$, $\{2, 3, 4, 6\}$, $\{2, 3, 5, 6\}$, $\{2, 4, 5, 6\}$, $\{3, 4, 5, 6\}$.

Definición 15. Un conjunto de tuplas $\{s_1, \dots, s_n\}$ de n índices $s_1 \in \{a_1, \dots, b_1\}$, \dots , $s_n \in \{a_n, \dots, b_n\}$, con $a_i, b_i \in \mathbb{Z}$ y $a_i \leq b_i$, $\forall i \in \{1, \dots, n\}$ es B ordenado si está ordenado de la siguiente forma:

- s_1 toma valores entre $\{a_1, \dots, b_1\} \uparrow$
- Para s_1 fijo, s_2 toma valores entre $\{a_2, \dots, b_2\} \uparrow$
- Para s_{n-1} fijo, s_n toma valores entre $\{a_n, \dots, b_n\} \uparrow$

Ejemplo 16. Las tuplas de índices de la forma $\{s_1, s_2, s_3\}$, con $s_1 \in \{3, 4, 5\}$, $s_2 \in \{5, 6\}$ y $s_3 \in \{5, 6, 7\}$, están B ordenadas de la siguiente forma: $\{3, 5, 5\}$, $\{3, 5, 6\}$, $\{3, 5, 7\}$, $\{3, 6, 6\}$, $\{3, 6, 7\}$, $\{4, 5, 5\}$, $\{4, 5, 6\}$, $\{4, 5, 7\}$, $\{4, 6, 6\}$, $\{4, 6, 7\}$, $\{5, 5, 5\}$, $\{5, 5, 6\}$, $\{5, 5, 7\}$, $\{5, 6, 6\}$, $\{5, 6, 7\}$.

Definición 16. Sean $\mathbf{x}_0, \dots, \mathbf{x}_n \in \mathbb{R}^n$, un *símplice* se define como

$$\blacktriangle(\mathbf{x}_0, \dots, \mathbf{x}_n) = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_{i=0}^n \mu_i \mathbf{x}_i \right\}, \quad (3.1)$$

donde $0 \leq \mu_i \leq 1$, $i \in \{0, \dots, n\}$ y $\sum_{i=0}^n \mu_i = 1$.

Un símplice se dice que es propio si no está contenido en hiperplanos de dimensión $(n - 1)$. En el esquema de construcción de la representación HL-CPWL de funciones continuas PWL solo se consideran a los símplices propios.

3.1 La función generadora γ .

Sean \mathcal{C} el espacio de funciones continuas que van de $D \subset \mathbb{R}^n$ a \mathbb{R} , una función $\zeta : \mathbb{R} \rightarrow \mathbb{R}$ y un conjunto de funciones $\mathfrak{S} = \{f_i \in \mathcal{C}, i = 1, 2, \dots, p\}$. Definimos la función generadora $\gamma_\zeta : \mathfrak{S} \times \mathfrak{S} \rightarrow \mathcal{C}$ por

$$\gamma_\zeta(f_1, f_2) = \frac{1}{4} [\zeta(\zeta(-f_1) + f_2) - \zeta(-f_1 + \zeta(f_2)) + \zeta(-f_1) + \zeta(f_2) - \zeta(-f_1 + f_2)].$$

Para nuestros propósitos, usaremos $\zeta(\mathbf{x}) = |\mathbf{x}|$. Para simplificar la notación, abreviaremos $\gamma_{|\cdot|}(f_i, f_j) = \gamma(f_i, f_j)$ y el punto de evaluación de las funciones en la entrada de la función generadora será omitido cuando no sea necesario especificarlo. Así

$$\gamma(f_1, f_2) = \frac{1}{4} (||-f_1| + f_2| - |-f_1 + |f_2|| + |-f_1| + |f_2| - |-f_1 + f_2|), \quad (3.2)$$

se sigue que

$$\gamma(f_1, f_2) = \begin{cases} f_1 & \text{si } 0 \leq f_1 \leq f_2, \\ f_2 & \text{si } 0 \leq f_2 \leq f_1, \\ 0 & \text{si } f_1 < 0 \text{ o } f_2 < 0. \end{cases} \quad (3.3)$$

Definimos la función $\gamma^{(k)} : \mathfrak{S}^k \rightarrow \mathcal{C}$, con $k \geq 0$, tal que

$$\begin{cases} \gamma^{(0)}(f_1) = f_1, \\ \gamma^{(1)}(f_1) = \gamma(f_1, f_1), \\ \gamma^{(2)}(f_1, f_2) = \gamma(f_1, f_2), \\ \vdots \\ \gamma^{(k)}(f_1, \dots, f_k) = \gamma(f_1, \gamma^{(k-1)}(f_2, \dots, f_k)). \end{cases} \quad (3.4)$$

De (3.3) y (3.4) tenemos

$$\gamma^{(k)}(f_1, \dots, f_k) = \begin{cases} f_1 & \text{si } 0 \leq f_1 \leq \gamma^{(k-1)}(f_2, \dots, f_k), \\ \gamma^{(k-1)}(f_2, \dots, f_k) & \text{si } 0 \leq \gamma^{(k-1)}(f_2, \dots, f_k) \leq f_1, \\ 0 & \text{si } f_{r_i} < 0 \text{ para alg\u00fan } i \in \{1, \dots, k\}. \end{cases} \quad (3.5)$$

Definici\u00f3n 17. El *nivel de anidaci\u00f3n* de una funci\u00f3n $nl(f)$ es igual al n\u00famero de funciones absolutas anidadas en su expresi\u00f3n.

As\u00ed, las funciones $\gamma^{(k)}(f_1, \dots, f_k)$ van anidando funciones de valor absoluto en su expresi\u00f3n, por lo que en [22] se establece el siguiente lema.

Lema 9. Dado $f = \gamma^{(k)}(f_1, \dots, f_k)$, entonces $nl(f) = k$, si $nl(f_i) = 0$ para toda $i \in \{1, \dots, k\}$.

Prueba. La prueba se hace por inducci\u00f3n en el orden de anidaci\u00f3n $nl(f)$ de la funci\u00f3n $f = \gamma^{(k)}(f_1, \dots, f_k)$. Sea $k = 0$, vemos que $\gamma^{(0)}(f_1)$ no tiene valores absolutos, por lo que tiene nivel de anidaci\u00f3n 0. Supongamos $k = 1$, de (3.2) y (3.4) tenemos que

$$\begin{aligned} \gamma^{(1)}(f_1) &= \frac{1}{4} (||-f_1| + f_1| - |-f_1 + |f_1|| + |-f_1| + |f_1| - |-f_1 + f_1|) = \\ &= \frac{1}{4} (||-f_1| + f_1| - |-f_1 + |f_1|| + 2|f_1|), \end{aligned} \quad (3.6)$$

note que si $f_1 \geq 0$ tenemos que $|-f_1| + f_1 = 2f_1 \geq 0$ y $-f_1 + |f_1| = 0$ y si $f_1 < 0$ tenemos que $|-f_1| + f_1 = 0$ y $-f_1 + |f_1| = -2f_1 \geq 0$. As\u00ed $||-f_1| + f_1| \equiv |-f_1| + f_1$ y $|-f_1 + |f_1|| \equiv -f_1 + |f_1|$ por lo que (3.6) se reduce a

$$\gamma^{(1)}(f_1) = \frac{1}{4} (|-f_1| + f_1 + f_1 - |f_1| + 2|f_1|) = \frac{1}{4} (f_1 + |f_1|),$$

la cual tiene nivel de anidaci\u00f3n igual a 1. Suponga $k = 2$, por inspecci\u00f3n podemos ver que $\gamma^{(2)}(f_1, f_2)$ tiene nivel de anidaci\u00f3n igual a 2. Suponga $k > 2$, asumamos primero que $\gamma^{(k-1)}(f_2, \dots, f_k)$ tiene nivel de anidaci\u00f3n igual a $(k - 1)$, por (3.5) tenemos que $\gamma^{(k)}(f_1, \dots, f_k) = \gamma(f_1, \gamma^{(k-1)}(f_2, \dots, f_k))$, adem\u00e1s, $\gamma^{(k-1)}(f_2, \dots, f_k) \geq 0$, por lo que

$$\begin{aligned} \gamma^{(k)}(f_1, \dots, f_k) &= \frac{1}{4} (||-f_1| + \gamma^{(k-1)}(f_2, \dots, f_k)| - |-f_1 + |\gamma^{(k-1)}(f_2, \dots, f_k)|| + |-f_1| + \\ &= \frac{1}{4} (|-f_1| + \gamma^{(k-1)}(f_2, \dots, f_k) - \\ &= \frac{1}{2} (|-f_1| + \gamma^{(k-1)}(f_2, \dots, f_k) - |-f_1 + \gamma^{(k-1)}(f_2, \dots, f_k)|), \end{aligned}$$

la cual tiene nivel de anidaci\u00f3n igual a k . ■

Para representar las funciones asociadas a los hiperplanos utilizados en las particiones que se presentan en este cap\u00edtulo usaremos la siguiente notaci\u00f3n. Dados los enteros q, i, j, k_q, k_i y k_j , definimos

$$\begin{aligned}
 \text{a)} \quad & \pi_q^{(k_q)}(\mathbf{x}) = x_q - k_q, \\
 \text{b)} \quad & \pi_{i,j}^{-(k_i,k_j)}(\mathbf{x}) = (x_i - k_i) - (x_j - k_j), \\
 \text{c)} \quad & \pi_{i,j}^{+(k_i,k_j)}(\mathbf{x}) = (x_i - k_i) + (x_j - k_j),
 \end{aligned} \tag{3.7}$$

donde x_i representa al i -ésimo componente del vector $\mathbf{x} \in \mathbb{R}^n$. $\pi_q^{(0)}$ y $\pi_{i,j}^{\pm(0,0)}$ serán abreviados π_q y $\pi_{i,j}^{\pm}$, respectivamente. Además, el punto de evaluación será omitido cuando no sea necesario especificarlo, es decir, $\pi_q^{(k_q)}(\mathbf{x}) = \pi_q^{(k_q)}$, $\pi_{i,j}^{-(k_i,k_j)}(\mathbf{x}) = \pi_{i,j}^{-(k_i,k_j)}$ y $\pi_{i,j}^{+(k_i,k_j)}(\mathbf{x}) = \pi_{i,j}^{+(k_i,k_j)}$.

Suponga $\mathfrak{S} = \{1, \pi_1(\mathbf{x}), \dots, \pi_h(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\}$, donde $\pi_1(\mathbf{x}), \dots, \pi_h(\mathbf{x})$ son funciones de la forma (3.7). Entonces, las funciones $\gamma^{(k)} : \mathfrak{S}^k \rightarrow \mathcal{C}$ son representaciones HL-CPWL de la forma (2.38).

3.2 Partición simplicial del hipercubo $[-1, 1]^n$.

3.2.1 Descripción del dominio.

Sea $[-1, 1]^n$ el dominio de las funciones PWL que trataremos en esta sección, con partición $H_{[-1,1]^n}$ formada por los hiperplanos

$$\begin{aligned}
 & \{\mathbf{x} \in \mathbb{R}^n : \pi_q(\mathbf{x}) = 0\}, \\
 & \{\mathbf{x} \in \mathbb{R}^n : \pi_{i,j}^+(\mathbf{x}) = 0\}, \\
 & \{\mathbf{x} \in \mathbb{R}^n : \pi_{i,j}^-(\mathbf{x}) = 0\},
 \end{aligned} \tag{3.8}$$

donde $q \in \{1, \dots, n\}$, $i \in \{1, \dots, n-1\}$, $j \in \{i+1, \dots, n\}$ y las funciones $\pi_q(\mathbf{x})$, $\pi_{i,j}^+(\mathbf{x})$ y $\pi_{i,j}^-(\mathbf{x})$ son de la forma (3.7).

Veamos que $H_{[-1,1]^n}$ particiona a $[-1, 1]^n$ en símlices propios. Para ello recordemos que un símlice se define por sus $(n+1)$ vértices. Así, definimos todos los vértices de los símlices en el hipercubo escogiendo un vértice en común, el origen, y el resto de los n vértices como la suma de vectores canónicos. De modo que cada símlice

$$\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n}), \tag{3.9}$$

tiene sus vértices definidos por

$$\mathbf{v}_{r_k} = \sum_{i=0}^k (-1)^{c_{r_i}} \mathbf{e}_{r_i}, \quad k \in \{1, \dots, n\}, \tag{3.10}$$

donde \mathbf{e}_{r_i} es el r_i -ésimo vector canónico de \mathbb{R}^n , $r_i \in \{1, \dots, n\}$, $r_i \neq r_j$ si $i \neq j$, $c_{r_i} \in \{0, 1\}$ y $\mathbf{v}_{r_0} = \mathbf{e}_{r_0} = \mathbf{0}$.

Ejemplo 17. Los vértices $\mathbf{0}$, \mathbf{e}_1 , \mathbf{e}_2 , $\mathbf{e}_1 + \mathbf{e}_2$, forman a los símlices $\blacktriangle(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2)$ y $\blacktriangle(\mathbf{0}, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2)$ que aparecen en la Figura 3.1, además, se encuentran en el primer cuadrante del hipercubo $[-1, 1]^2$.

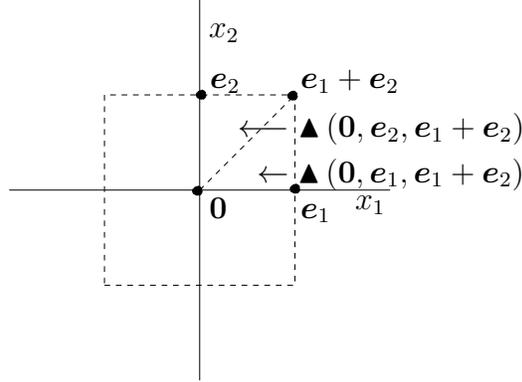


Figura 3.1: Símplices en el primer cuadrante del hipercubo $[-1, 1]^2$.

Como se ha visto anteriormente, una partición segmenta el dominio de una función en regiones, las cuales son poliedros convexos. Para el caso de la partición $H_{[-1,1]^n}$, en [22] se prueba que estas regiones son símplices propios, tal como se ve a continuación.

Lema 10. *La partición $H_{[-1,1]^n}$ y los puntos $\partial[-1, 1]^n$ en la frontera del hipercubo $[-1, 1]^n$ subdividen al hipercubo $[-1, 1]^n$ en símplices de la forma $\blacktriangle(v_{r_0}, \dots, v_{r_n})$.*

Prueba. Para probar el lema veremos que cada borde (combinación convexa de dos vértices) de un símplice resulta de la intersección entre hiperplanos de la partición y los puntos en la frontera del hipercubo. Sea \mathbf{y} un punto en el borde de un símplice $\blacktriangle(v_{r_0}, \dots, v_{r_n})$, este puede ser expresado como combinación convexa de cualesquiera dos vértices, de modo que

$$\mathbf{y} = \alpha \sum_{i=0}^{k_1} (-1)^{c_{r_i}} \mathbf{e}_{r_i} + (1 - \alpha) \sum_{i=0}^{k_2} (-1)^{c_{r_i}} \mathbf{e}_{r_i}, \quad (3.11)$$

donde $k_1, k_2 \in \{0, \dots, n\}$, $k_1 \neq k_2$, $r_i \in \{0, \dots, n\}$, $r_i \neq r_j$ si $i \neq j$, $c_{r_i} \in \{0, 1\}$ y $\alpha \in [0, 1]$. Supongamos sin pérdida de generalidad que $k_2 > k_1$ y llamemos y_{r_i} al r_i -ésimo componente de \mathbf{y} , se sigue que:

- (i) $y_{r_i} = (-1)^{c_{r_i}}$, $\forall i \in \{1, \dots, k_1\}$,
- (ii) $y_{r_i} = (1 - \alpha)(-1)^{c_{r_i}}$, $\forall i \in \{k_1 + 1, \dots, k_2\}$,
- (iii) $y_{r_i} = 0$, $\forall i \in \{k_2 + 1, \dots, n\}$.

Definimos la función $\lambda(c_{r_i}, c_{r_j}) = \frac{(-1)^{c_{r_i}}}{(-1)^{c_{r_j}}}$, tal que

$$\lambda(c_{r_i}, c_{r_j}) = \begin{cases} 1 & \text{si } c_{r_i} = c_{r_j}, \\ -1 & \text{si } c_{r_i} \neq c_{r_j}. \end{cases}$$

Por (i), $y_{r_i} = \lambda(c_{r_i}, c_{r_j}) y_{r_j} = \pm y_{r_j} = \pm 1$, $\forall i, j \in \{1, \dots, k_1\}$, $i \neq j$, por lo que \mathbf{y} pertenece al hiperplano $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{r_i, r_j}^{\pm}(\mathbf{x}) = x_{r_i} \pm x_{r_j} = 0 \right\}$, $i, j \in \{1, \dots, k_1\}$, $i \neq j$, el cual está en la partición $H_{[-1,1]^n}$. Además, puesto que $y_{r_i} = \pm 1$, $i \in \{1, \dots, k_1\}$, tenemos que \mathbf{y} pertenece

a $\partial[-1, 1]^n$. Por (ii), $y_{r_i} = \lambda(c_{r_i}, c_{r_j})y_{r_j} = \pm y_{r_j} = \pm(1 - \alpha)$, $\forall i, j \in \{k_1 + 1, \dots, k_2\}$, $i \neq j$, por lo que \mathbf{y} pertenece al hiperplano $\{\mathbf{x} \in \mathbb{R}^n : \pi_{r_i, r_j}^\pm(\mathbf{x}) = x_{r_i} \pm x_{r_j} = 0\}$, $i, j \in \{k_1 + 1, \dots, k_2\}$, $i \neq j$, el cual está en la partición $H_{[-1, 1]^n}$. Por último, de (iii) tenemos que $y_{r_i} = y_{r_j} = 0$, $\forall i, j \in \{k_2 + 1, \dots, n\}$, $i \neq j$, por lo que \mathbf{y} pertenece al hiperplano $\{\mathbf{x} \in \mathbb{R}^n : \pi_{r_i, r_j}^\pm(\mathbf{x}) = x_{r_i} \pm x_{r_j} = 0\}$, $i, j \in \{k_1 + 1, \dots, k_2\}$, $i \neq j$, el cual está en la partición $H_{[-1, 1]^n}$. Además, como $y_{r_i} = 0$, $\forall i \in \{k_2 + 1, \dots, n\}$, \mathbf{y} pertenece al hiperplano $\{\mathbf{x} \in \mathbb{R}^n : \pi_{r_i}(\mathbf{x}) = x_{r_i} = 0\}$, $i \in \{k_2 + 1, \dots, n\}$, el cual está en la partición $H_{[-1, 1]^n}$. Por lo que cada borde (combinación convexa de dos vértices) de un símplex resulta de la intersección entre hiperplanos de la partición $H_{[-1, 1]^n}$ y los puntos en la frontera del hiper-cubo $[-1, 1]^n$. Así, la partición $H_{[-1, 1]^n}$ y los puntos $\partial[-1, 1]^n$ en la frontera del hiper-cubo $[-1, 1]^n$ particionan al hiper-cubo $[-1, 1]^n$ en símplexes de la forma $\blacktriangle(v_{r_0}, \dots, v_{r_n})$. ■

Ejemplo 18. Sea $H_{[-1, 1]^2}$ la partición formada por los hiperplanos

$$\begin{aligned} S_1 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 = 0\}, \\ S_2 &= \{\mathbf{x} \in \mathbb{R}^2 : x_2 = 0\}, \\ S_3 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 + x_2 = 0\}, \\ S_4 &= \{\mathbf{x} \in \mathbb{R}^2 : x_1 - x_2 = 0\}, \end{aligned}$$

la cual particiona a $[-1, 1]^2$ en símplexes

$$\begin{aligned} \blacktriangle(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, \mathbf{e}_2, -\mathbf{e}_1 + \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, -\mathbf{e}_1, -\mathbf{e}_1 + \mathbf{e}_2), \\ \blacktriangle(\mathbf{0}, -\mathbf{e}_1, -\mathbf{e}_1 - \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, -\mathbf{e}_2, -\mathbf{e}_1 - \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, -\mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2), \quad \blacktriangle(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 - \mathbf{e}_2), \end{aligned}$$

señalados en la Figura 3.2.

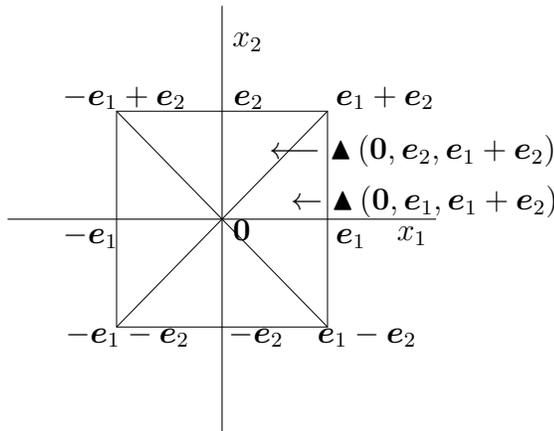


Figura 3.2: Partición $H_{[-1, 1]^2}$ del hiper-cubo $[-1, 1]^2$.

Por razones numéricas que serán vistas en la siguiente subsección, los vértices de los símplexes que particionan a $[-1, 1]^n$ serán ordenados en un conjunto $V_{[-1, 1]^n}$ de la siguiente manera: primero, los vértices son separados por clases, cada clase está dada por el

número de elementos diferentes de 0 en cada vértice; también, los vértices pertenecientes a una misma clase $k \in \{0, \dots, n\}$ son ordenados en un vector $V_{[-1,1]^n}^{(k)}$ como se muestra a continuación:

- Vértices de clase 0 ($V_{[-1,1]^n}^{(0)}$): vértices sin elementos diferentes de 0, $\mathbf{x} = \mathbf{0}$
- Vértices de clase 1 ($V_{[-1,1]^n}^{(1)}$): $2 \binom{n}{1}$ vértices con un elemento diferente de 0:

$$\begin{cases} (x_1, \dots, x_n) : x_{r_1} = 1 & x_{r_2} = 0 & \forall r_2 \neq r_1, \\ (x_1, \dots, x_n) : x_{r_1} = -1 & x_{r_2} = 0 & \forall r_2 \neq r_1, \end{cases}$$

donde $r_i \in \{1, \dots, n\} \uparrow$.

- Vértices de clase $2 \leq k \leq n$ ($V_{[-1,1]^n}^{(k)}$): $2^k \binom{n}{k}$ vértices con k elementos diferentes de 0:

$$(x_1, \dots, x_n) : x_{r_j} = (-1)^{c_j^h}, \forall j \in \{1, \dots, k\} \text{ y } x_{r_i} = 0 \quad \forall r_i \neq r_j, \quad (3.12)$$

donde las tuplas de índices $\{r_1, \dots, r_k\}$, $r_1 \in \{1, \dots, n - k + 1\}$, $r_j \in \{r_{j-1}, \dots, n - k + j\}$, con $j \in \{2, \dots, k\} \uparrow$, son A ordenados. Además, para una tupla fija de índices $\{r_1, \dots, r_k\}$, las diferentes combinaciones de signo de las variables son tomadas por la elección de los coeficientes $c_j^h = \lceil (2^k + 1 - h) / (2^{k-j}) \rceil$, con $h \in \{1, \dots, 2^k\} \uparrow$ y $\lceil a \rceil$ nos da el entero mayor que a más cercano.

Así, $V_{[-1,1]^n}$ se forma de la concatenación de los vectores de vértices de clase k , $V_{[-1,1]^n}^{(k)}$, $k \in \{0, \dots, n\} \uparrow$.

Ejemplo 19. De acuerdo con lo anterior, los vertices de los s mplices formados por la partici3n $H_{[-1,1]^3}$ en el hipercubo $[-1, 1]^3$ se ordenan de la siguiente forma

- $V_{[-1,1]^3}^{(0)}$: (0 0 0).
- $V_{[-1,1]^3}^{(1)}$: (1 0 0), (-1 0 0), (0 1 0), (0 -1 0), (0 0 1), (0 0 -1).
- $V_{[-1,1]^3}^{(2)}$: (1 1 0), (1 -1 0), (-1 1 0), (-1 -1 0), (1 0 1), (1 0 -1), (-1 0 1), (-1 0 -1), (0 1 1), (0 1 -1), (0 -1 1), (0 -1 -1).
- $V_{[-1,1]^3}^{(3)}$: (1 1 1), (1 1 -1), (1 -1 1), (1 -1 -1), (-1 1 1), (-1 1 -1), (-1 -1 1), (-1 -1 -1).

3.2.2 Conjunto generador del espacio $PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$.

Sea $\mathfrak{S}_{[-1,1]^n} = \{1, \pi_{r_1}, -\pi_{r_1}, r_1 \in \{1, \dots, n\}\}$, con π_{r_1} y $-\pi_{r_1}$ como en a) de (3.7) y la función generadora $\gamma^{(k)}(\cdot, \dots, \cdot)$ definida en (3.5), un vector de funciones $\mathbf{\Lambda}_{[-1,1]^n}$ es definido y ordenado de la siguiente manera: primero, las funciones son separadas por su nivel de anidación $k \in \{0, \dots, n\}$ en vectores $\mathbf{\Lambda}_{[-1,1]^n}^{(k)}$, además, funciones con un mismo nivel de anidación son ordenadas como se muestra a continuación:

- Funciones con nivel de anidación igual a 0 ($\mathbf{\Lambda}_{[-1,1]^n}^{(0)}$): la constante $\gamma^{(0)}(1) = 1$.
- Funciones con nivel de anidación igual a 1 ($\mathbf{\Lambda}_{[-1,1]^n}^{(1)}$): $2 \binom{n}{1}$ funciones de la forma

$$\begin{cases} \gamma^{(1)}(\pi_{r_1}), \\ \gamma^{(1)}(-\pi_{r_1}), \end{cases}$$

$$r_1 \in \{1, \dots, n\} \uparrow.$$

- Funciones con nivel de anidación igual a $2 \leq k \leq n$ ($\mathbf{\Lambda}_{[-1,1]^n}^{(k)}$): $2^k \binom{n}{k}$ funciones de la forma

$$\gamma^{(k)}\left(\left(-1\right)^{c_1^h} \pi_{r_1}, \dots, \left(-1\right)^{c_k^h} \pi_{r_k}\right), \quad (3.13)$$

donde las tuplas de índices $\{r_1, \dots, r_k\}$, $r_1 \in \{1, \dots, n - k + 1\}$, $r_j \in \{r_{j-1}, \dots, n - k + j\}$, con $j \in \{2, \dots, k\} \uparrow$, son A ordenados. Además, para una tupla fija de índices $\{r_1, \dots, r_k\}$, las diferentes combinaciones de signo de las variables son tomadas por la elección de los coeficientes $c_q^h = \lceil (2^k + 1 - h) / (2^{k-q}) \rceil$, con $h \in \{1, \dots, 2^k\} \uparrow$, $q \in \{1, \dots, k\}$ y $\lceil a \rceil$ nos da el entero mayor que a más cercano.

Así, $\mathbf{\Lambda}_{[-1,1]^n}$ se forma de la concatenación de los vectores de funciones con nivel de anidación k , $\mathbf{\Lambda}_{[-1,1]^n}^{(k)}$, $k \in \{0, \dots, n\} \uparrow$.

Ejemplo 20. De acuerdo con lo anterior, las funciones PWL en $\mathbf{\Lambda}_{[-1,1]^3}$ se ordenan de la siguiente forma:

- $\mathbf{\Lambda}_{[-1,1]^3}^{(0)}$: $\gamma^{(0)}(1)$.
- $\mathbf{\Lambda}_{[-1,1]^3}^{(1)}$: $\gamma^{(1)}(\pi_1), \gamma^{(1)}(-\pi_1), \gamma^{(1)}(\pi_2), \gamma^{(1)}(-\pi_2), \gamma^{(1)}(\pi_3), \gamma^{(1)}(-\pi_3)$.
- $\mathbf{\Lambda}_{[-1,1]^3}^{(2)}$: $\gamma^{(2)}(\pi_1, \pi_2), \gamma^{(2)}(\pi_1, -\pi_2), \gamma^{(2)}(-\pi_1, \pi_2), \gamma^{(2)}(-\pi_1, -\pi_2), \gamma^{(2)}(\pi_1, \pi_3), \gamma^{(2)}(\pi_1, -\pi_3), \gamma^{(2)}(-\pi_1, \pi_3), \gamma^{(2)}(-\pi_1, -\pi_3), \gamma^{(2)}(\pi_2, \pi_3), \gamma^{(2)}(\pi_2, -\pi_3), \gamma^{(2)}(-\pi_2, \pi_3), \gamma^{(2)}(-\pi_2, -\pi_3)$.
- $\mathbf{\Lambda}_{[-1,1]^3}^{(3)}$: $\gamma^{(3)}(\pi_1, \pi_2, \pi_3), \gamma^{(3)}(\pi_1, \pi_2, -\pi_3), \gamma^{(3)}(\pi_1, -\pi_2, \pi_3), \gamma^{(3)}(\pi_1, -\pi_2, -\pi_3), \gamma^{(3)}(-\pi_1, \pi_2, \pi_3), \gamma^{(3)}(-\pi_1, \pi_2, -\pi_3), \gamma^{(3)}(-\pi_1, -\pi_2, \pi_3), \gamma^{(3)}(-\pi_1, -\pi_2, -\pi_3)$.

Lema 11. *El número de vértices en $V_{[-1,1]^n}$ es igual al número de funciones en $\Lambda_{[-1,1]^n}$.*

Prueba. Ver [22]. ■

Como se mencionó anteriormente, las funciones en $\Lambda_{[-1,1]^n}$ son también funciones HL-CPWL, por lo que están definidas en una partición, la cual se especifica en el siguiente lema [22].

Lema 12. *Las funciones HL-CPWL en $\Lambda_{[-1,1]^n}$ están definidos en la partición $H_{[-1,1]^n}$ y ninguna otra componente fue añadida a la partición.*

Prueba. Sea $\Lambda_{[-1,1]^n}$ un vector de funciones construidas como en (3.13). Puesto que una función PWL está definida por sus funciones locales afines (2.4), y estas cambian cuando se cruza un hiperplano que separa regiones adyacentes, aquí se probará que las funciones en $\Lambda_{[-1,1]^n}$ cambian su expresión solamente cuando se cruza un hiperplano de $H_{[-1,1]^n}$, esto se verá por inducción en el nivel de anidación de las funciones en $\Lambda_{[-1,1]^n}$.

Sean $\gamma^{(1)}(\pi_i)$, $\gamma^{(1)}(-\pi_i)$, $i \in \{1, \dots, n\}$, funciones en $\Lambda_{[-1,1]^n}^{(1)}$, reduciendo la expresión (3.2) tenemos que $\gamma^{(1)}(\pi_i) = \frac{1}{2}(\pi_i + |\pi_i|)$ y $\gamma^{(1)}(-\pi_i) = \frac{1}{2}(\pi_i - |\pi_i|)$. Sea $\mathbf{x} \in \mathbb{R}^n$, tenemos que

$$\begin{aligned} \gamma^{(1)}(\pi_i) &= \begin{cases} \pi_i & \text{si } \pi_i \geq 0, \\ 0 & \text{si } \pi_i < 0, \end{cases} \\ \gamma^{(1)}(-\pi_i) &= \begin{cases} 0 & \text{si } \pi_i \geq 0, \\ \pi_i & \text{si } \pi_i < 0, \end{cases} \end{aligned}$$

por lo que hay un cambio de expresión en $\gamma^{(1)}(\pi_i)$ y en $\gamma^{(1)}(-\pi_i)$ cuando se cruza el hiperplano $\{\mathbf{x} \in \mathbb{R}^n : \pi_i(\mathbf{x}) = 0\}$, $i \in \{1, \dots, n\}$, el cual es un hiperplano en la partición $H_{[-1,1]^n}$.

Supongamos que las funciones HL-CPWL en $\Lambda_{[-1,1]^n}^{(j)}$ de la forma (3.13) están definidas en $H_{[-1,1]^n}$. Sea

$$\gamma^{(j+1)}\left((-1)^{c_1^i} \pi_{r_1}, \dots, (-1)^{c_{j+1}^i} \pi_{r_{j+1}}\right),$$

con $r_1 \in \{1, \dots, n - (j + 1) + 1\}$, $r_k \in \{r_{k-1}, \dots, n - (j + 1) + k\}$, para $k \in \{2, \dots, (j + 1)\} \uparrow$ y $c_q^h = \lceil (2^{(j+1)} + 1 - h) / (2^{(j+1)-q}) \rceil$, con $h \in \{1, \dots, 2^{(j+1)}\} \uparrow$, $q \in \{1, \dots, j + 1\}$. De (3.5) tenemos que $\gamma^{j+1}((-1)^{c_1^i} \pi_{r_1}, \dots, (-1)^{c_{j+1}^i} \pi_{r_{j+1}}) =$

- a) $(-1)^{c_1^i} \pi_{r_1}$ si $0 \leq (-1)^{c_1^i} \pi_{r_1} \leq \gamma^{(j)}\left((-1)^{c_2^i} \pi_{r_2}, \dots, (-1)^{c_{j+1}^i} \pi_{r_{j+1}}\right)$,
- b) $\gamma^{(j)}\left((-1)^{c_2^i} \pi_{r_2}, \dots, (-1)^{c_{j+1}^i} \pi_{r_{j+1}}\right)$ si $0 \leq \gamma^{(j)}\left((-1)^{c_2^i} \pi_{r_2}, \dots, (-1)^{c_{j+1}^i} \pi_{r_{j+1}}\right) \leq (-1)^{c_1^i} \pi_{r_1}$,
- c) 0 si $(-1)^{c_k^i} \pi_{r_k} < 0$ para algún $k \in \{1, \dots, j + 1\}$,

de donde se desprenden los siguientes casos:

- (i) Supongamos que ocurre b), por hipótesis tenemos que $\gamma^{j+1}(\cdot, \dots, \cdot) = \gamma^j(\cdot, \dots, \cdot)$ está definida en la partición $H_{[-1,1]^n}$.
- (ii) Supongamos que ocurre c), tenemos que $(-1)^{c_k} \pi_{r_k}(\mathbf{x}) = (-1)^{c_k} x_{r_k} = 0$, por lo que hay un cambio cuando se cruza $\{\mathbf{x} \in \mathbb{R}^n : \pi_{r_k}(\mathbf{x}) = 0\}$, con $r_k \in \{1, \dots, j+1\}$, el cual es un hiperplano en la partición $H_{[-1,1]^n}$.
- (iii) Supongamos que ocurre a) y b), tenemos que $\gamma^{(j)} \left((-1)^{c_2} \pi_{r_2}, \dots, (-1)^{c_{j+1}} \pi_{r_{j+1}} \right) = (-1)^{c_1} \pi_{r_1}$, por (3.5) tenemos

$$\gamma^{(j)} \left((-1)^{c_2} \pi_{r_2}, \dots, (-1)^{c_{j+1}} \pi_{r_{j+1}} \right) = \begin{cases} (-1)^{c_2} \pi_{r_2}, \\ \vdots \\ (-1)^{c_{j+1}} \pi_{r_{j+1}}, \end{cases}$$

en la región determinada por $\{\mathbf{x} \in \mathbb{R}^n : (-1)^{c_k} \pi_{r_k}(\mathbf{x}) > 0\}$, $k \in \{2, \dots, j+1\}$, por lo que los cambios de expresión ocurren cuando $(-1)^{c_1} \pi_{r_1}(\mathbf{x}) = (-1)^{c_v} \pi_{r_v}(\mathbf{x})$, es decir, $(-1)^{c_1} x_{r_1} = (-1)^{c_v} x_{r_v}$, cuando se cruza el hiperplano $\{\mathbf{x} \in \mathbb{R}^n : (-1)^{c_k} \pi_{r_1, r_v}^{-\lambda_{v,1}(i)}(\mathbf{x}) = x_{r_1} - \lambda_{v,1}(i) x_{r_v} = 0\}$, con $v \in \{2, \dots, j+1\}$ y $\lambda_{v_1, v_2}(i) = \frac{(-1)^{c_{v_1}}}{(-1)^{c_{v_2}}} \in \{-1, 1\}$, el cual es un hiperplano en la partición $H_{[-1,1]^n}$.

Por lo que todas las funciones HL-CPWL en $\Lambda_{[-1,1]^n}$ están definidos en la partición $H_{[-1,1]^n}$ y ninguna otra componente fue añadida a la partición. ■

Por el Lema 11, $\Lambda_{[-1,1]^n}$ tiene el mismo número de funciones que $V_{[-1,1]^n}$ tiene de vértices. Así, sea $A_{[-1,1]^n} \in \mathbb{R}^{m \times m}$ la matriz cuyas filas son obtenidas de la evaluación de las funciones $\Lambda_{[-1,1]^n}$ en los vértices de $V_{[-1,1]^n}$, es decir,

$$A_{[-1,1]^n} = \begin{bmatrix} \Lambda_{[-1,1]^n}_1(V_{[-1,1]^n}_1) & \dots & \Lambda_{[-1,1]^n}_m(V_{[-1,1]^n}_1) \\ \vdots & & \vdots \\ \Lambda_{[-1,1]^n}_1(V_{[-1,1]^n}_m) & \dots & \Lambda_{[-1,1]^n}_m(V_{[-1,1]^n}_m) \end{bmatrix}, \quad (3.14)$$

donde $\Lambda_{[-1,1]^n}_i$ es la i -ésima función en $\Lambda_{[-1,1]^n}$, $i \in \{1, \dots, m\}$, y $V_{[-1,1]^n}_j$ es el j -ésimo vértice en $V_{[-1,1]^n}$, $j \in \{1, \dots, m\}$. Las ventajas numéricas que, como se mencionó antes, tienen las definiciones constructivas del conjunto $V_{[-1,1]^n}$ y el vector $\Lambda_{[-1,1]^n}$, se ven en el siguiente lema [22].

Lema 13. *La matriz $A_{[-1,1]^n}$ es no singular.*

Prueba. Para probar este lema veremos que la matriz $A_{[-1,1]^n}$ es triangular inferior y sus elementos en la diagonal principal son diferentes de 0. Sea \mathbf{x} un vértice en $V_{[-1,1]^n}^{(k)}$, donde la r'_j -ésima componente de \mathbf{x} es

$$x_{r'_j} = (-1)^{c_j^{i'}}, \quad j \in \{1, \dots, k\} \quad y \quad x_{r'_q} = 0 \quad \forall r'_q \neq r'_j,$$

con $r'_1 \in \{1, \dots, n - k + 1\}$, $r'_j \in \{r'_{j-1}, \dots, n - k + j\}$, si $j \in \{2, \dots, k\}$, además, $c_j^{i'} = \lceil (2^k + 1 - i') / (2^{k-j}) \rceil$, con $i' \in \{1, \dots, 2^k\} \uparrow$. Y sea $g(\mathbf{x})$ una función en $\Lambda_{[-1,1]^n}^{(l)}$ tal que

$$g(\mathbf{x}) = \gamma^{(l)} \left((-1)^{c_1^{i''}} \pi_{r_1''}, \dots, (-1)^{c_l^{i''}} \pi_{r_l''} \right),$$

con $r''_1 \in \{1, \dots, n - l + 1\}$, $r''_j \in \{r''_{j-1}, \dots, n - l + j\}$, con $j \in \{2, \dots, l\}$, además $c_j^{i''} = \lceil (2^l + 1 - i'') / (2^{l-j}) \rceil$, con $i'' \in \{1, \dots, 2^l\}$. Así, $g(\mathbf{x})$ es un elemento en la diagonal de $A_{[-1,1]^n}$ sí y solo sí

- $k = l$,
- $r'_j = r''_j, \forall j \in \{1, \dots, k\}$,
- $i' = i''$.

Se sigue que, si $g(\mathbf{x})$ es un elemento en la diagonal de $A_{[-1,1]^n}$, entonces $(-1)^{c_j^{i''}} \pi_{r_j''}(\mathbf{x}) = (-1)^{c_j^{i'}}$, es decir que $(-1)^{c_j^{i''}} x_{r_j''} = (-1)^{c_j^{i'}}$, por lo que $x_{r_j''} = 1, \forall j \in \{1, \dots, k\}$. Así, $\gamma^{(k)}(1, \dots, 1) = 1$, por (3.5) tenemos que $g(\mathbf{x}) = 1$.

Suponga $k = l$, $r''_j = r'_j, \forall j \in \{1, \dots, k\}$ y $i'' \neq i'$, se sigue que $(-1)^{c_j^{i''}} \neq (-1)^{c_j^{i'}}$ para al menos un $j = j^*$ por lo que $(-1)^{c_{j^*}^{i''}} \pi_{r_{j^*}''}(\mathbf{x}) = (-1)^{c_{j^*}^{i''}} x_{r_{j^*}''} = -1$, por (3.5) tenemos que $g(\mathbf{x}) = 0$. Ahora, suponga $k = l$ y $r''_j \neq r'_j$, para algún $j = j^* \in \{1, \dots, k\}$, se sigue que $(-1)^{c_{j^*}^{i''}} \pi_{r_{j^*}''}(\mathbf{x}) = (-1)^{c_{j^*}^{i''}} x_{r_{j^*}''} = 0$ y por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Por último, suponga que $l > k$. Por hipótesis existe un r_{j^*} , con $j^* \in \{1, \dots, l\}$ tal que $r''_{j^*} \neq r'_{j^*}, \forall j \in \{1, \dots, k\}$, así, $(-1)^{c_{j^*}^{i''}} \pi_{r_{j^*}''}(\mathbf{x}) = (-1)^{c_{j^*}^{i''}} x_{r_{j^*}''} = 0$ y por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Por lo que que la matriz $A_{[-1,1]^n}$ es triangular inferior y sus elementos en la diagonal principal son diferentes de 0, en consecuencia es no singular. ■

Ejemplo 21. Continuando con el Ejemplo 18, para el cual $V_{[-1,1]^2} =$

$$\{(0 \ 0), (1 \ 0), (-1 \ 0), (0 \ 1), (0 \ -1), (1 \ 1), (1 \ -1), (-1 \ 1), (-1 \ -1)\}$$

y $\Lambda_{[-1,1]^2} =$

$$\left[\begin{array}{cccccc} \gamma^{(0)}(1) & \gamma^{(1)}(\pi_1) & \gamma^{(1)}(-\pi_1) & \gamma^{(1)}(\pi_2) & \gamma^{(1)}(-\pi_2) & \gamma^{(2)}(\pi_1, \pi_2) & \gamma^{(2)}(\pi_1, -\pi_2) \\ \gamma^{(2)}(-\pi_1, \pi_2) & & & & & & \end{array} \right]^T,$$

tenemos que

$$A_{[-1,1]^2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Sea $\mathbf{c}_{[-1,1]^n}^{(i)} \in \mathbb{R}^{q_i}$, con $q_i = 2^i \binom{n}{i}$, $i \in \{0, \dots, n\}$ y $q_0 = 1$, un vector de coeficientes $\mathbf{c}_{[-1,1]^n} \in \mathbb{R}^{\sum_{i=0}^n q_i}$, es decir, $\mathbf{c}_{[-1,1]^n} = [\mathbf{c}_{[-1,1]^n}^{(0)} \dots \mathbf{c}_{[-1,1]^n}^{(n)}]$. Así, funciones de la forma

$$g(\mathbf{x}) = \mathbf{c}_{[-1,1]^n}^T \mathbf{\Lambda}_{[-1,1]^n}, \quad (3.15)$$

$g : [-1, 1]^n \rightarrow \mathbb{R}$ son funciones HL-CPWL definidas sobre la partición $H_{[-1,1]^n}$. Por lo que se verá en el siguiente teorema [22], (3.15) nos proporciona una representación general de las funciones en $PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$.

Teorema 5. *Las funciones en el vector $\mathbf{\Lambda}_{[-1,1]^n}$ generan el espacio de funciones $PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$*

Prueba. Sea $f \in PWL_{H_{[-1,1]^n}} [[-1, 1]^n]$ y el vector $\mathbf{\Lambda}_{[-1,1]^n}$ de funciones HL-CPWL definidas sobre la partición $H_{[-1,1]^n}$, la prueba se basa en encontrar una elección de parámetros tales que f pueda ser únicamente representada por (3.15), puesto que los cambios en la función ocurren en la intersección de sus fronteras, f queda definida por los valores que adquiere en sus vértices [22], se pretende encontrar un vector $\mathbf{c}_{[-1,1]^n}$ tal que $\mathbf{b}_{[-1,1]^n} = A_{[-1,1]^n} \mathbf{c}_{[-1,1]^n}$, donde $A_{[-1,1]^n}$ es la matriz (3.14) y $\mathbf{b}_{[-1,1]^n}$ es el vector que resulta de evaluar f en los vértices de $V_{[-1,1]^n}$. Por Lema 13, $A_{[-1,1]^n}$ es no singular, por lo que existe su matriz inversa $A_{[-1,1]^n}^{-1}$, en consecuencia es posible encontrar un único $\mathbf{c}_{[-1,1]^n}$ tal que $\mathbf{c}_{[-1,1]^n} = A_{[-1,1]^n}^{-1} \mathbf{b}_{[-1,1]^n}$, el cual satisface (3.15). ■

3.3 Partición simplicial ortogonal, uniformemente espaciada, de $D_{\text{rec}} \subset \mathbb{R}^n$.

3.3.1 Descripción del dominio.

Sea $D_{\text{rec}} \subset \mathbb{R}^n$

$$D_{\text{rec}} := \left\{ (x_1 \ \dots \ x_n) : 0 \leq x_i \leq m_i \delta, i \in \{1, \dots, n\} \right\}, \quad (3.16)$$

el dominio de las funciones PWL que trataremos en esta sección, donde δ es el tamaño de paso de una malla que particiona a D_{rec} y $m_i \in \mathbb{Z}^+$. La partición $H_{D_{\text{rec}}}$ de D_{rec} consideradas en esta sección están formadas por los hiperplanos:

$$\begin{cases} \mathbf{x} \in \mathbb{R}^n : \pi^{(k_q \delta)}(\mathbf{x}) = x_q - k_q \delta = 0, \\ \mathbf{x} \in \mathbb{R}^n : \pi_{i,j}^{-,(0,k_j \delta)}(\mathbf{x}) = x_i - (x_j - k_j \delta) = 0, \\ \mathbf{x} \in \mathbb{R}^n : \pi_{i,j}^{-,(k_i \delta, 0)}(\mathbf{x}) = (x_i - k_i \delta) - x_j = 0, \end{cases} \quad (3.17)$$

donde $q \in \{1, \dots, n\}$, $k_q \in \{1, \dots, m_q - 1\}$, $k_i \in \{1, \dots, m_i - 1\}$, $k_j \in \{1, \dots, m_j - 1\}$, $i \in \{1, \dots, n - 1\}$ y $j \in \{i + 1, \dots, n\}$.

Ejemplo 22. Sea

$$D_{\text{eje}} = \{ \mathbf{x} \in \mathbb{R}^2 : 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 2 \},$$

con $\delta = 1$, $m_1 = 3$ y $m_2 = 2$. $H_{D_{\text{eje}}}$ está dada por los siguientes hiperplanos:

$$\begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_1^{(0)} = x_1 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_1^{(1)} = x_1 - 1 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_1^{(2)} = x_1 - 2 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_2^{(0)} = x_2 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_2^{(1)} = x_2 - 1 = 0, \end{cases} \quad \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{1,2}^{-,(0,0)} = x_1 - x_2 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_{1,2}^{-,(0,1)} = x_1 - (x_2 - 1) = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_{1,2}^{-,(1,0)} = (x_1 - 1) - x_2 = 0, \\ \mathbf{x} \in \mathbb{R}^2 : \pi_{1,2}^{-,(2,0)} = (x_1 - 2) - x_2 = 0, \end{cases}$$

como se observa en la Figura 3.3.

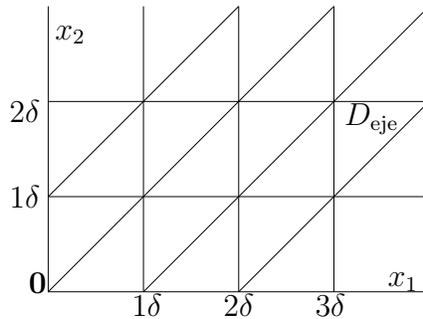


Figura 3.3: D_{eje} y partición $H_{D_{\text{eje}}}$ del Ejemplo 22.

Así, D_{eje} se subdivide en $\prod_{i=1}^n m_i$ hipercubos congruentes a $[0, \delta]^n$. cada uno de ellos particionado en forma similar a la del hipercubo $[-1, 1]^n$ visto en la sección anterior. Así, la partición forma símlices en cada uno de estos hipercubos como lo establece el siguiente lema [22].

Lema 14. La partición $H_{D_{\text{rec}}}$ y los puntos ∂D_{rec} en la frontera de D_{rec} particionan a D_{rec} en símlices que se obtienen al escalar los símlices de la forma $\blacktriangle(v_{r_0}, \dots, v_{r_n})$, (3.1), por un factor δ y desplazarlos sumándoles un vector \mathbf{p} .

Prueba. La prueba sigue la demostración del Lema 10. Probaremos que los hipercubos de D_{rec} son una traslación del hipercubo $\delta[0, 1]^n$.

Sea $\mathbf{p} \in \mathbb{R}^n$, con r_i -ésima componente $p_{r_i} \in \{0, \dots, m_{r_i} - 1\}$ (con esta elección de las componentes de \mathbf{p} , se ubica a este vector en la intersección de los hiperplanos de la partición $H_{D_{\text{rec}}}$) y sean $\delta \in \mathbb{R}$ y la función homeomórfica $T_{\mathbf{p}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ definida por

$$\mathbf{z} = T_{\mathbf{p}}(\mathbf{x}) = \mathbf{x} + \delta\mathbf{p}, \quad (3.18)$$

la cual traslada $[0, \delta]^n$ a $\delta\mathbf{p} + [0, \delta]^n$, de modo que $\mathbf{x} = \mathbf{z} - \delta\mathbf{p}$ con componentes $x_{r_i} = z_{r_i} - \delta p_{r_i}$,

Por el Lema 10, los hiperplanos $\{\mathbf{x} \in \mathbb{R}^n : \pi_q = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^n : \pi_{i,j}^- = 0\}$, con $q \in \{1, \dots, n\}$, $i \in \{1, \dots, n-1\}$ y $j \in \{i+1, \dots, n\}$, particionan a $[0, \delta]^n$, resta probar que la imagen de estos hiperplanos bajo la función $T_{\mathbf{p}}$ pertenecen a la partición $H_{D_{\text{rec}}}$, por lo que veremos los siguientes casos:

- (i) Dado el hiperplano $\{\mathbf{x} \in \mathbb{R}^n : \pi_{r_i} = x_{r_i} = 0\}$ y (3.18), tenemos, $z_{r_i} - \delta p_{r_i} = x_{r_i}$, por lo que $\mathbf{z} \in \left\{ \mathbf{z} \in \mathbb{R}^n : \pi_{r_i}^{(\delta p_{r_i})} = 0 \right\}$, con $r_i \in \{1, \dots, n\}$, $p_{r_i} \in \{0, \dots, m_{r_i}\}$, el cual es un hiperplano en la partición $H_{D_{\text{rec}}}$.
- (ii) Dados los hiperplanos $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{r_i, r_j}^- = x_{r_i} - x_{r_j} = 0 \right\}$ y (3.18), tenemos, $(z_{r_i} - \delta p_{r_i}) - (z_{r_j} - \delta p_{r_j}) = x_{r_i} - x_{r_j}$, se siguen dos casos:
 - Si $p_{r_i} \geq p_{r_j}$. Dado un $\lambda_{r_i} = p_{r_i} - p_{r_j}$, tenemos que $\mathbf{z} \in \left\{ \mathbf{z} \in \mathbb{R}^n : \pi_{r_i, r_j}^{-(\delta \lambda_{r_i}, 0)} = 0 \right\}$, con $r_i \in \{1, \dots, n-1\}$, $r_j \in \{r_i + 1, \dots, n\}$, $\lambda_{r_i} \in \{0, \dots, m_{r_i}\}$, el cual es un hiperplano en la partición $H_{D_{\text{rec}}}$.
 - Si $p_{r_j} \geq p_{r_i}$. Dado un $\lambda_{r_j} = p_{r_j} - p_{r_i}$, tenemos que $\mathbf{z} \in \left\{ \mathbf{z} \in \mathbb{R}^n : \pi_{r_i, r_j}^{-(0, \delta \lambda_{r_j})} = 0 \right\}$, con $r_i \in \{1, \dots, n-1\}$, $r_j \in \{r_i + 1, \dots, n\}$, $\lambda_{r_j} \in \{0, \dots, m_{r_j}\}$, el cual es un hiperplano en la partición $H_{D_{\text{rec}}}$.

Por lo que la partición $H_{D_{\text{rec}}}$ y los puntos ∂D_{rec} en la frontera de D_{rec} particionan a D_{rec} en símlices $\delta(\blacktriangle(v_{r_0}, \dots, v_{r_n}) + \mathbf{p})$. ■

Ejemplo 23. Continuando con el Ejemplo 22, los símlices resultantes de esta partición son de la forma $\mathbf{p} + \blacktriangle(\mathbf{0}, \mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2)$ y $\mathbf{p} + \blacktriangle(\mathbf{0}, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2)$, con $\mathbf{p} = (p_1, p_2)$, $p_1 \in \{0, 1, 2\}$ y $p_2 \in \{0, 1\}$ (Figura 3.3).

Por razones numéricas que serán vistas en la siguiente subsección, los vértices de los símlices que particionan a D_{rec} serán ordenados en un vector $V_{D_{\text{rec}}}$ de la siguiente manera: primero, los vértices son separados por clases, cada clase está dada por el número de elementos diferentes de 0 en cada vértice; también, los vértices pertenecientes a una misma clase $r \in \{0, \dots, n\}$ son ordenados en un vector $V_{D_{\text{rec}}}^{(r)}$ como se muestra a continuación:

- Vértices de clase 0 ($V_{D_{\text{rec}}}^{(0)}$): vértices sin elementos diferentes de 0, $\mathbf{x} = \mathbf{0}$.

- Vértices de clase 1 ($V_{D_{\text{rec}}}^{(1)}$): $\sum_{k_1=1}^n m_{k_1}$ vértices con un elemento diferente de 0:

$$\mathbf{x} = q_{k_1} \delta \mathbf{e}_{k_1},$$

donde $k_1 \in \{1, \dots, n\} \uparrow$. Además, para cada k_1 fijo, tenemos $q_{k_1} \in \{1, \dots, m_{k_1}\} \uparrow$.

- Vértices de clase $2 \leq r \leq n$ ($V_{D_{\text{rec}}}^{(r)}$): $\left(\sum_{k_1=1}^{n-r+1} m_{k_1} \sum_{k_2=k_1+1}^{n-r+2} m_{k_2} \dots \sum_{k_r=k_{r-1}+1}^n m_{k_r} \right)$ vértices con r elementos diferentes de 0:

$$\mathbf{x} = \sum_{p=1}^r q_{k_p} \delta \mathbf{e}_{k_p}, \quad (3.19)$$

donde las tuplas $\{k_1, \dots, k_r\}$, $k_1 \in \{1, \dots, n-r+1\}$ y $k_j \in \{k_{j-1}+1, \dots, n-r+j\}$, $\forall j \in \{2, \dots, r\}$ son A ordenadas. Además, cada tupla fija de índices $\{k_1, \dots, k_r\}$, $q_{k_i} \in \{1, \dots, m_{k_i}\}$, con $i \in \{1, \dots, r\}$, es B ordenada.

Así, $V_{D_{\text{rec}}}$ se forma de la concatenación de los vectores de vértices de clase r , $V_{D_{\text{rec}}}^{(r)}$, $r \in \{0, \dots, n\} \uparrow$.

Ejemplo 24. Continuando con el Ejemplo 22, los vértices de la partición $H_{D_{\text{eje}}}$ en D_{eje} se ordenan de la siguiente manera:

- $V_{D_{\text{eje}}}^{(0)}$: $(0 \ 0)$.
- $V_{D_{\text{eje}}}^{(1)}$: $(1 \ 0)$, $(2 \ 0)$, $(3 \ 0)$, $(0 \ 1)$, $(0 \ 2)$.
- $V_{D_{\text{eje}}}^{(2)}$: $(1 \ 1)$, $(1 \ 2)$, $(2 \ 1)$, $(2 \ 2)$, $(3 \ 1)$, $(3 \ 2)$.

3.3.2 Conjunto generador del espacio $PWL_{H_{D_{\text{rec}}}} [D_{\text{rec}}]$.

Sea $\mathfrak{S}_{D_{\text{rec}}} = \left\{ 1, \pi_k^{(j_k \delta)}, k \in \{1, \dots, n\}, j_k \in \{0, \dots, m_k - 1\} \right\}$, con $\pi_k^{(j_k \delta)}$ como en a) de (3.7) y la función generadora $\gamma^{(k)}(\cdot, \dots, \cdot)$ definida en (3.5). Un vector de funciones $\Lambda_{D_{\text{rec}}}$ es definido y ordenado de la siguiente manera: primero, las funciones son separadas por su nivel de anidación $l \in \{0, \dots, n\}$ en un vector $\Lambda_{D_{\text{rec}}}^{(l)}$, además, funciones con un mismo nivel de anidación son ordenadas como se muestra a continuación:

- Funciones con nivel de anidación igual a 0 ($\Lambda_{D_{\text{rec}}}^{(0)}$): la constante $\gamma^{(0)}(1) = 1$.
- Funciones con nivel de anidación igual a 1 ($\Lambda_{D_{\text{rec}}}^{(1)}$): $\sum_{k_1=1}^n m_{k_1}$ funciones de la forma:

$$\gamma^{(1)} \left(\pi_{k_1}^{(j_{k_1} \delta)} \right),$$

con $k_1 \in \{1, \dots, n\} \uparrow$. Además, para cada k_1 fijo, tenemos $j_{k_1} \in \{0, \dots, m_{k_1} - 1\} \uparrow$.

- Funciones con nivel de anidación igual a $2 \leq l \leq n$ ($\Lambda_{D_{\text{rec}}}^{(l)}$): $\left(\sum_{k_1=1}^{n-l+1} m_{k_1} \sum_{k_2=k_1+1}^{n-l+2} m_{k_2} \dots \sum_{k_l=k_{l-1}+1}^n m_{k_l} \right)$ funciones de la forma:

$$\gamma^{(l)} \left(\pi_{k_1}^{(j_{k_1} \delta)}, \dots, \pi_{k_l}^{(j_{k_l} \delta)} \right), \quad (3.20)$$

donde las tuplas de índices $\{k_1, \dots, k_l\}$, $k_1 \in \{1, \dots, n-l+1\}$, $k_h \in \{k_{h-1}, \dots, n-l+h\}$, $\forall h \in \{2, \dots, l\}$ son A ordenados. Además, para cada tupla fija de índices $\{k_1, \dots, k_l\}$, $j_{k_i} \in \{0, \dots, m_{k_i} - 1\}$, con $i \in \{1, \dots, n\}$, son B ordenados.

Así, $\Lambda_{D_{\text{rec}}}$ se forma de la concatenación de los vectores de funciones con nivel de anidación l , $\Lambda_{D_{\text{rec}}}^{(l)}$, $r \in \{0, \dots, n\} \uparrow$.

Ejemplo 25. Continuando con el Ejemplo 22, las funciones PWL en $\Lambda_{D_{\text{eje}}}$ se ordenan de la siguiente forma:

- $\Lambda_{D_{\text{eje}}}^{(0)} : \gamma^{(0)}(1)$.
- $\Lambda_{D_{\text{eje}}}^{(1)} : \gamma^{(1)} \left(\pi_1^{(0)} \right), \gamma^{(1)} \left(\pi_1^{(1)} \right), \gamma^{(1)} \left(\pi_1^{(2)} \right), \gamma^{(1)} \left(\pi_2^{(0)} \right), \gamma^{(1)} \left(\pi_2^{(1)} \right)$.
- $\Lambda_{D_{\text{eje}}}^{(2)} : \gamma^{(2)} \left(\pi_1^{(0)}, \pi_1^{(0)} \right), \gamma^{(2)} \left(\pi_1^{(0)}, \pi_2^{(1)} \right), \gamma^{(2)} \left(\pi_1^{(1)}, \pi_2^{(0)} \right), \gamma^{(2)} \left(\pi_1^{(1)}, \pi_2^{(1)} \right), \gamma^{(2)} \left(\pi_1^{(2)}, \pi_2^{(0)} \right), \gamma^{(2)} \left(\pi_1^{(2)}, \pi_2^{(1)} \right)$.

Lema 15. El número de vértices en $V_{D_{\text{rec}}}$ es igual al número de funciones en $\Lambda_{D_{\text{rec}}}$.

Prueba. Sea $r \in \{1, \dots, n\}$, de (3.19), los vértices en $V_{D_{\text{rec}}}^{(r)}$ son de la forma

$$\mathbf{x} = q_{k_1''} \delta e_{k_1''} + \dots + q_{k_r''} \delta e_{k_r''}, \quad (3.21)$$

donde $k_1'' \in \{1, \dots, n-r+1\}$ y $k_j'' \in \{k_{j-1}'' + 1, \dots, n-r+j\}$, $\forall j \in \{2, \dots, r\}$, $q_{k_i''} \in \{1, \dots, m_{k_i''}\}$, con $i \in \{1, \dots, r\}$. Sea $l \in \{1, \dots, n\}$, de (3.20), las funciones en $\Lambda_{D_{\text{rec}}}^{(l)}$ son de la forma

$$\gamma^{(l)} \left(\pi_{k_1'}^{(j_{k_1'} \delta)}, \dots, \pi_{k_l'}^{(j_{k_l'} \delta)} \right), \quad (3.22)$$

donde $k_1' \in \{1, \dots, n-l+1\}$, $k_h' \in \{k_{h-1}', \dots, n-l+h\}$, $\forall h \in \{2, \dots, l\}$, $j_{k_i'} \in \{0, \dots, m_{k_i'} - 1\}$, con $i \in \{1, \dots, n\}$. De (3.21) y (3.22), se sigue que, si $r = l$, $V_{D_{\text{rec}}}^{(r)}$ tiene el mismo número de elementos que $\Lambda_{D_{\text{rec}}}^{(l)}$, por lo que $V_{D_{\text{rec}}}$ tiene el mismo número de elementos que $\Lambda_{D_{\text{rec}}}$. ■

Como se mencionó anteriormente, las funciones en $\Lambda_{D_{\text{rec}}}$ son también funciones HL-CPWL, por lo que estas están definidas en una partición, la cual se precisa en el siguiente lema [22].

Lema 16. *Las funciones HL-CPWL en $\Lambda_{D_{\text{rec}}}$ están definidas en la partición $H_{D_{\text{rec}}}$ y ninguna otra componente fue añadida a la partición.*

Prueba. Esta prueba se realiza siguiendo la prueba del Lema 12. Las funciones $\gamma^{(k)}(\pi_{r_1}, \dots, \pi_{r_k})$ cambian su expresión al cruzar los hiperplanos

$$\begin{cases} \mathbf{x} \in \mathbb{R}^n : \pi_{r_q}(\mathbf{x}) = x_{r_q} = 0 \\ \mathbf{x} \in \mathbb{R}^n : \pi_{r_i, r_j}^-(\mathbf{x}) = x_{r_i} - x_{r_j} = 0 \end{cases},$$

con $r_q \in \{1, \dots, n\}$, $r_i \in \{1, \dots, n-1\}$ y $r_j \in \{r_i + 1, \dots, n\}$. Aplicando la transformación T_p (3.18) a estos hiperplanos tenemos

$$\begin{cases} \mathbf{z} \in \mathbb{R}^n : \pi_{r_q}^{(\delta p_{r_q})}(\mathbf{z}) = 0 \\ \mathbf{z} \in \mathbb{R}^n : \pi_{r_i, r_j}^{-(\delta \lambda_{r_i}, 0)}(\mathbf{z}) = 0 \\ \mathbf{z} \in \mathbb{R}^n : \pi_{r_i, r_j}^{-(0, \delta \lambda_{r_j})}(\mathbf{z}) = 0 \end{cases},$$

con $r_q \in \{1, \dots, n\}$, $p_{r_q} \in \{0, \dots, m_{r_q} - 1\}$, $r_i \in \{1, \dots, n-1\}$, $r_i \in \{r_i + 1, \dots, n\}$, $\lambda_{r_i} \in \{0, \dots, m_{r_i} - 1\}$ y $\lambda_{r_j} \in \{0, \dots, m_{r_j} - 1\}$. Los cuales están definidos en la partición $H_{D_{\text{rec}}}$, por lo que las funciones $\gamma^{(l)}\left(\pi_{r_1}^{(j_{r_1} \delta)}, \dots, \pi_{r_l}^{(j_{r_l} \delta)}\right)$, $j_{r_q} \in \{0, \dots, m_{r_q} - 1\}$, están definidas en $H_{D_{\text{rec}}}$. ■

Por el Lema 15, $\Lambda_{D_{\text{rec}}}$ tiene el mismo número de funciones que $V_{D_{\text{rec}}}$ tiene de vértices. Así, sea $A_{D_{\text{rec}}} \in \mathbb{R}^{m \times m}$ la matriz cuyas filas son obtenidas de la evaluación de las funciones $\Lambda_{D_{\text{rec}}}$ en los vértices de $V_{D_{\text{rec}}}$, es decir,

$$A_{D_{\text{rec}}} = \begin{bmatrix} \Lambda_{D_{\text{rec}1}}(V_{D_{\text{rec}1}}) & \dots & \Lambda_{D_{\text{rec}m}}(V_{D_{\text{rec}1}}) \\ \vdots & & \vdots \\ \Lambda_{D_{\text{rec}1}}(V_{D_{\text{rec}m}}) & \dots & \Lambda_{D_{\text{rec}m}}(V_{D_{\text{rec}m}}) \end{bmatrix}, \quad (3.23)$$

donde $\Lambda_{D_{\text{rec}i}}$ es la i -ésima función en $\Lambda_{D_{\text{rec}}}$, $i \in \{1, \dots, m\}$, y $V_{D_{\text{rec}j}}$ es el j -ésimo vértice en $V_{D_{\text{rec}}}$, $j \in \{1, \dots, m\}$. Las ventajas numéricas de se mencionaron en la construcción del conjunto $V_{D_{\text{rec}}}$ y el vector $\Lambda_{D_{\text{rec}}}$ se ven en el siguiente lema [22].

Lema 17. *$A_{D_{\text{rec}}}$ es no singular.*

Prueba. Para probar este lema veremos que la matriz es triangular inferior y sus elementos en la diagonal principal son diferentes de 0. Sea

$$\mathbf{x} = q_{k_1''} \delta \mathbf{e}_{k_1''} + \dots + q_{k_r''} \delta \mathbf{e}_{k_r''},$$

donde $k_1'' \in \{1, \dots, n-r+1\}$ y $k_j'' \in \{k_{j-1}'' + 1, \dots, n-r+j\}$, $\forall j \in \{2, \dots, r\}$, $q_{k_i''} \in \{1, \dots, m_{k_i''}\}$, con $i \in \{1, \dots, r\}$, un vértice en $V_{D_{\text{rec}}}^{(r)}$, $r \in \{1, \dots, n\}$. Sea

$$g(\mathbf{x}) = \gamma^{(l)} \left(\pi_{k'_1}^{(j_{k'_1} \delta)}, \dots, \pi_{k'_l}^{(j_{k'_l} \delta)} \right),$$

donde $k'_1 \in \{1, \dots, n-l+1\}$, $k'_h \in \{k'_{h-1}, \dots, n-l+h\}$, $\forall h \in \{2, \dots, l\}$, $j_{k'_i} \in \{0, \dots, m_{k'_i} - 1\}$, con $i \in \{1, \dots, n\}$, una función definida en $\Lambda_{D_{\text{rec}}}^{(l)}$, $l \in \{1, \dots, n\}$. Así, $g(\mathbf{x})$ es un elemento en la diagonal principal de $A_{D_{\text{rec}}}$ sí y solo sí

- $r = l$
- $k'_v = k''_v, \forall v \in \{1, \dots, r\}$
- $q_{k''_v} = j_{k'_v} + 1, \forall v \in \{1, \dots, r\}$

Para organizar la prueba los elementos de $A_{D_{\text{rec}}}$ serán analizados, primero, por la clase r de los vértices donde la función g es evaluada. Después, es organizada por el nivel de anidación l para cada clase fija de vértices. Los elementos $r > l$ no serán considerados, pues se encuentran por debajo de la diagonal principal de $A_{D_{\text{rec}}}$.

Vértices de clase cero ($r = 0$): estos vértices solo poseen un punto, el origen. Al evaluarse en el bloque diagonal, es decir $l = 0$, este se evalúa únicamente en $\gamma^{(0)}(1) = 1$. Ahora veamos los elementos sobre la diagonal ($l > r$), tenemos que $\pi_{k'_v}^{(j_{k'_v} \delta)}(\mathbf{0}) = -j_{k'_v} \delta \leq 0$, $\forall v \in \{1, \dots, l\}$ y $\forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Vértices de clase uno ($r = 1$): sea $l = r$, tenemos los siguientes casos:

- Suponga $k'_1 \neq k''_1$, tenemos que $\pi_{k'_1}^{(j_{k'_1} \delta)}(\mathbf{x}) = -j_{k'_1} \delta \leq 0$, $\forall j_{k'_1} \in \{0, \dots, m_{k'_1} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.
- Suponga $k'_1 = k''_1$ y $q_{k''_1} - 1 = j_{k'_1}$, es decir, que es un elemento en la diagonal de $A_{D_{\text{rec}}}$, tenemos que $\pi_{k'_1}^{(j_{k'_1} \delta)}(\mathbf{x}) = q_{k''_1} \delta - j_{k'_1} \delta = \delta$, por (3.5) tenemos que $g(\mathbf{x}) = \delta$.
- Suponga $k'_1 = k''_1$ y $q_{k''_1} \leq j_{k'_1}$, es decir, elementos sobre la diagonal principal de $A_{D_{\text{rec}}}$, tenemos que $\pi_{k'_1}^{(j_{k'_1} \delta)}(\mathbf{x}) = q_{k''_1} \delta - j_{k'_1} \delta \leq 0$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Sea $l > r = 1$, de donde se desprenden dos casos:

- Suponga $k'_v \neq k''_1, \forall v \in \{1, \dots, l\}$, tenemos que $\pi_{k'_v}^{(j_{k'_v} \delta)}(\mathbf{x}) = -j_{k'_v} \delta \leq 0$, $\forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.
- Suponga que existe un $v^* \in \{1, \dots, l\}$ tal que $k'_{v^*} \neq k''_1$, tenemos que $\pi_{k'_{v^*}}^{(j_{k'_{v^*}} \delta)}(\mathbf{x}) = -j_{k'_{v^*}} \delta \leq 0$, $\forall j_{k'_{v^*}} \in \{0, \dots, m_{k'_{v^*}} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Vértices de clase ($r > 1$): sea $l = r$, tenemos los siguientes casos:

- Suponga que existe un $v \in \{1, \dots, r\}$ tal que $k'_v \neq k''_v$, tenemos que $\pi_{k'_v}^{(j_{k'_v} \delta)}(\mathbf{x}) = -j_{k'_v} \delta \leq 0$, $\forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.
- Suponga $k'_v = k''_v$ y $q_{k''_v} - 1 = j_{k'_v}$, $\forall v \in \{1, \dots, r\}$, es decir, que es un elemento en la diagonal de $A_{D_{\text{rec}}}$, tenemos que $\pi_{k'_v}^{(j_{k'_v} \delta)}(\mathbf{x}) = q_{k''_v} \delta - j_{k'_v} \delta = \delta$, por (3.5) tenemos que $g(\mathbf{x}) = \delta$.
- Suponga $k'_v = k''_v$ y $q_{k''_v} \leq j_{k'_v}$, $\forall v \in \{1, \dots, r\}$, es decir, elementos sobre la diagonal principal de $A_{D_{\text{rec}}}$, tenemos que $\pi_{k'_v}^{(j_{k'_v} \delta)}(\mathbf{x}) = q_{k''_v} \delta - j_{k'_v} \delta \leq 0$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Sea $l > r$, es decir, elementos sobre la diagonal principal de $A_{D_{\text{rec}}}$, existe un $v^* \in \{1, \dots, l\}$ tal que $k'_{v^*} \neq k''_{v^*}$, $\forall v \in \{1, \dots, r\}$, tenemos que $\pi_{k'_{v^*}}^{(j_{k'_{v^*}} \delta)}(\mathbf{x}) = -j_{k'_{v^*}} \delta \leq 0$, $\forall j_{k'_{v^*}} \in \{0, \dots, m_{k'_{v^*}} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Por lo anterior la matriz A es diagonal inferior y por tanto es no singular. ■

Ejemplo 26. Continuando con el Ejemplo 22, para el cual $V_{D_{\text{eje}}} =$

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

y $\Lambda_{D_{\text{eje}}} =$

$$\left[\begin{array}{cccccccccc} \gamma^{(0)}(1) & \gamma^{(1)}(\pi_1^{(0)}) & \gamma^{(1)}(\pi_1^{(1)}) & \gamma^{(1)}(\pi_1^{(2)}) & \gamma^{(1)}(\pi_2^{(0)}) & \gamma^{(1)}(\pi_2^{(1)}) & \gamma^{(2)}(\pi_1^{(0)}, \pi_1^{(0)}) \\ \gamma^{(2)}(\pi_1^{(0)}, \pi_2^{(1)}) & \gamma^{(2)}(\pi_1^{(1)}, \pi_2^{(0)}) & \gamma^{(2)}(\pi_1^{(1)}, \pi_2^{(1)}) & \gamma^{(2)}(\pi_1^{(2)}, \pi_2^{(0)}) & \gamma^{(2)}(\pi_1^{(2)}, \pi_2^{(0)}) & \gamma^{(2)}(\pi_1^{(2)}, \pi_2^{(1)}) \end{array} \right]^T,$$

tenemos que

$$A_{D_{\text{eje}}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 2 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 2 & 1 & 2 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 3 & 2 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 3 & 2 & 1 & 2 & 1 & 2 & 1 & 2 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Sea $\mathbf{c}_{D_{\text{rec}}}^{(i)} \in \mathbb{R}^{q_i}$, con $q_i = \left(\sum_{k_1=1}^{n-r+1} m_{k_1} \sum_{k_2=k_1+1}^{n-r+2} m_{k_2} \cdots \sum_{k_r=k_{r-1}+1}^n m_{k_r} \right)$, $i \in \{0, \dots, n\}$ y $q_0 = 1$, un vector de coeficientes $\mathbf{c}_{D_{\text{rec}}} \in \mathbb{R}^{\sum_{i=0}^n q_i}$, es tal que $\mathbf{c}_{D_{\text{rec}}} = \left[\mathbf{c}_{D_{\text{rec}}}^{(0)} \dots \mathbf{c}_{D_{\text{rec}}}^{(n)} \right]$. Así, funciones de la forma

$$g(\mathbf{x}) = \mathbf{c}_{D_{\text{rec}}}^T \mathbf{\Lambda}_{D_{\text{rec}}}, \quad (3.24)$$

$g : D_{\text{rec}} \rightarrow \mathbb{R}$ son funciones HL-CPWL definidas sobre la partición $H_{D_{\text{rec}}}$. Pero además toda función continua PWL definida en D_{rec} puede representarse como un elemento de $PWL_{H_{D_{\text{rec}}}}[D_{\text{rec}}]$ mediante (3.24) como se establece en el siguiente teorema [22].

Teorema 6. *Las funciones en el vector $\mathbf{\Lambda}_{D_{\text{rec}}}$ generan el espacio de funciones $PWL_{H_{D_{\text{rec}}}}[D_{\text{rec}}]$*

Prueba. Sea $f \in PWL_{H_{D_{\text{rec}}}}[D_{\text{rec}}]$ y el vector $\mathbf{\Lambda}_{D_{\text{rec}}}$ de funciones HL-CPWL definidas sobre la partición $H_{D_{\text{rec}}}$, la prueba se basa en encontrar una elección de parámetros tales que f pueda ser únicamente representada por (3.24). Puesto que los cambios en la función ocurren en la intersección de sus fronteras, f queda definida por los valores que adquiere en sus vértices [22]. Se pretende encontrar un vector $\mathbf{c}_{D_{\text{rec}}}$ tal que $\mathbf{b}_{D_{\text{rec}}} = A_{D_{\text{rec}}} \mathbf{c}_{D_{\text{rec}}}$, donde $A_{D_{\text{rec}}}$ es la matriz (3.23) y $\mathbf{b}_{D_{\text{rec}}}$ es el vector que resulta de evaluar f en los vértices de $V_{D_{\text{rec}}}$. Por Lema 17, $A_{[-1,1]^n}$ es no singular, de modo que existe un único $\mathbf{c}_{D_{\text{rec}}}$ tal que $\mathbf{c}_{D_{\text{rec}}} = A_{D_{\text{rec}}}^{-1} \mathbf{b}_{D_{\text{rec}}}$, el cual satisface (3.24). ■

3.4 Propiedades computacionales.

Sea f una función en el espacio de funciones PWL definido ya sea en la subsección 3.2 o 3.3. De Lema 13 y el Lema 17 tenemos que $A_{[-1,1]^n}$ y $A_{D_{\text{rec}}}$, respectivamente, son triangulares inferiores, esto nos permite calcular los vectores de coeficientes en (3.15) y (3.24) sin invertir las matrices $A_{[-1,1]^n}$ y $A_{D_{\text{rec}}}$. Sean $a_{i,j}$ el i, j -ésimo elemento ya sea de la matriz $A_{[-1,1]^n}$ o de la matriz $A_{D_{\text{rec}}}$, c_k el k -ésimo elemento en el vector de coeficientes ya sea en (3.15) o en (3.24), respectivamente, y sea b_l el l -ésimo elemento del vector que resulta de evaluar a f en los respectivos vértices de su partición. Tenemos que cada uno de estos coeficientes pueden ser obtenidos de forma recursiva por

$$\begin{aligned} c_1 &= \frac{1}{a_{1,1}} b_1, \\ &\vdots \\ c_k &= \frac{1}{a_{i,i}} \left(b_i - \sum_{p=1}^{i-1} a_{i,p} c_p \right), \end{aligned} \quad (3.25)$$

más aún, los valores en la diagonal principal de estas matrices satisfacen que $a_{1,1} = 1$ y $a_{i,i} = \delta$.

Como se muestra en [22], las representaciones HL-CPWL con partición simplicial presentadas en este capítulo poseen el número mínimo y exacto de parámetros requeridos para representar una función PWL f en los espacios de funciones presentados. A continuación hacemos una comparación de los parámetros requeridos para representar estas funciones

en el caso de la representación estándar y la representación CPWL. Esta comparación se hace de la siguiente forma. Suponga (3.16) con $m_i = k$, $i \in \{1, \dots, n\}$. En este caso D_{rec} tiene k^n hipercubos, cada hipercubo tiene $n!$ símlices y el número de vértices es $(k+1)^n$. Se sigue que:

- Representación estándar PWL: por cada símplex es necesario almacenar $(n+1)$ parámetros que definen al vector gradiente y la constante (2.4) de cada región, así, almacenamos $n!(n+1)$ parámetros por hipercubo. Entonces, el número total de parámetros es $k^n n!(n+1) = k^n (n+1)!$.
- Representación HL-CPWL: por el Lema 15 y el Teorema 6 el número total de parámetros en esta representación es $(k+1)^n$.
- Representación CPWL: El número de parámetros es igual al número de hiperplanos de la partición (3.17) más $(n+1)$ del término afín en la expresión CPWL lo que da un total de parámetros de $kn + \binom{n}{2}(2k-1) + n + 1$. Cabe señalar nuevamente que este tipo de representación solo puede usarse para representar funciones PWL con variación consistente.

Finalmente, desde el punto de vista de aplicabilidad, comparado con los métodos *look-up table*, note que ambos métodos almacenan el número de parámetros dados por los valores de la función f en sus vértices. Además, los métodos *look-up table* necesitan almacenar $(n+1)$ índices por símplex para identificar estos valores, lo que da a un total de $k^n (n+1)!$ de índices, a esto hay que sumar los pasos del algoritmo para interpolar valores intermedios de forma adecuada. Esto contrasta con la representación HL-CPWL que necesita almacenar y evaluar $(k+1)^n$ funciones $\gamma^{(i)}(\pi_1, \dots, \pi_i)$.

Capítulo 4

Resultados de este trabajo.

4.1 Generalización de la representación HL-CPWL con partición simplicial.

En esta sección se abordará un esquema que permite construir representaciones HL-CPWL de funciones continuas PWL con particiones basadas en mallas simpliciales más generales que las mostradas en el Capítulo 3. Para ello se seguirán los siguientes pasos: primero, se realizará una descripción del dominio de las funciones a representar, donde ahora los hiperplanos estarán determinados por un vector normal a ellos y un vector de desplazamiento, también, se verá que el número de hiperplanos en la partición es el necesario para garantizar que cada intersección sea con degeneración mínima; el segundo paso consiste en construir un vector de funciones HL-CPWL haciendo uso de la función $\gamma^{(k)}$ (3.5) y los hiperplanos en la partición descrita en esta sección. Además, se probará que las funciones que componen este vector generan el conjunto de funciones PWL con la partición y el dominio dados; seguido de esto, se describirá un algoritmo que permita construir representaciones HL-CPWL de las funciones PWL de que se están tratando; por último se realizará un análisis de las propiedades computacionales que este esquema de construcción posee.

4.1.1 Descripción del dominio.

Sea $\Xi = \{\xi_i : i = 1, \dots, n\}$ una base de \mathbb{R}^n , donde la j -ésima componente de ξ_i se denota por ξ_{ij} , sea $M \in \mathbb{R}^{n \times n}$ una matriz de cambio de base de la base canónica a la base Ξ , de modo que

$$M = \begin{bmatrix} \xi_{11} & \cdots & \xi_{n1} \\ \vdots & \ddots & \vdots \\ \xi_{1n} & \cdots & \xi_{nn} \end{bmatrix}.$$

Consideremos al espacio de funciones $PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$ con partición $H_{D_{\text{gen}}}$, que se precisa más abajo, definida en

$$D_{\text{gen}} := \{M^T \mathbf{x} + \mathbf{p} : \mathbf{x} = (x_1 \ \dots \ x_n) \in \mathbb{R}^n, 0 \leq x_i \leq m_i, i \in \{1, \dots, n\}\}, \quad (4.1)$$

donde $m_i \in \mathbb{Z}^+$, $\mathbf{p} \in \mathbb{R}^n$. Además, tenemos que

$$M^T \mathbf{x} = \sum_{i=1}^n x_i \boldsymbol{\xi}_i, \quad (4.2)$$

por lo que todo punto $\mathbf{x}' \in D_{\text{gen}}$ es de la forma:

$$\mathbf{x}' = \sum_{i=1}^n x_i \boldsymbol{\xi}_i + \mathbf{p}, \quad (4.3)$$

con $0 \leq x_i \leq m_i, i \in \{1, \dots, n\}$.

Ejemplo 27. Dados los vectores en \mathbb{R}^2 :

$$\boldsymbol{\xi}_1 = (1 \ -1)^T, \quad \boldsymbol{\xi}_2 = (1 \ 2)^T, \quad \mathbf{p} = (-3 \ -1)^T,$$

y los enteros $m_1 = 2, m_2 = 3$, en la Figura 4.1a podemos apreciar, en comparación con el esquema de partición ortogonal y uniformemente distribuida visto en el capítulo anterior, que los vectores $\boldsymbol{\xi}_1$ y $\boldsymbol{\xi}_2$ nos dan la dirección en la que se orienta al dominio de las funciones PWL a tratar. Además, vemos que el vector \mathbf{p} desplaza fuera del origen a D_{gen} .

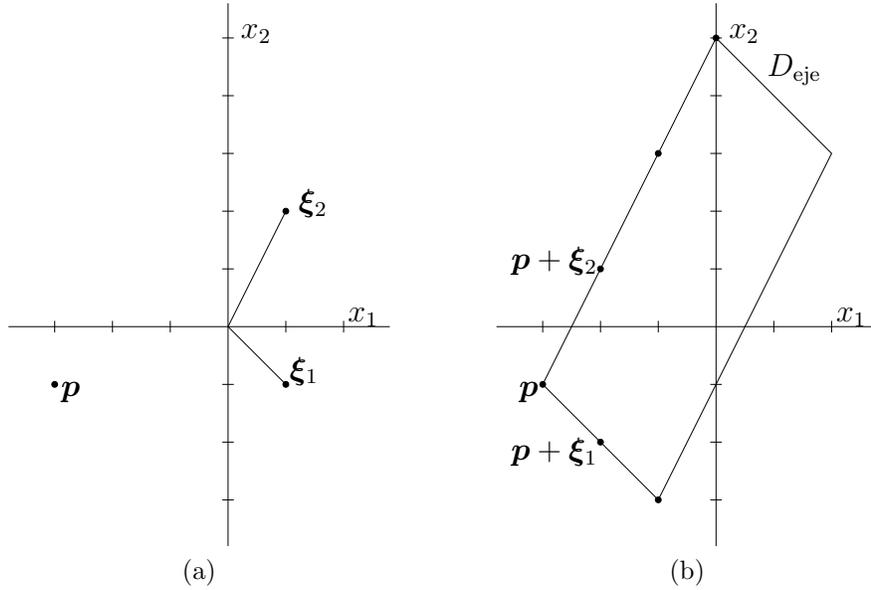


Figura 4.1: (a) Vectores $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2$ y \mathbf{p} del Ejemplo 27. (b) Conjunto $D_{\text{eje}} \subset \mathbb{R}^2$ del Ejemplo 27.

Tenemos entonces que

$$M_{D_{\text{eje}}} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix},$$

por lo que

$$D_{\text{eje}} = \left\{ M_{D_{\text{eje}}}^T \mathbf{x} + \mathbf{p} : \mathbf{x} = (x_1 \ x_2), 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 3 \right\},$$

de modo que todo punto $\mathbf{x}' \in D_{\text{eje}}$ (ver Figura 4.1b) es de la forma

$$\mathbf{x}' = x_1 \boldsymbol{\xi}_1 + x_2 \boldsymbol{\xi}_2 + \mathbf{p}, \quad 0 \leq x_1 \leq 2, \quad 0 \leq x_2 \leq 3.$$

A continuación, definiremos a la partición $H_{D_{\text{gen}}}$ de tal forma que sus intersecciones entre hiperplanos tengan la propiedad de degeneración mínima y sean del mismo orden que la dimensión del espacio de salida.

Sea $S_i, i \in \{1, \dots, n\}$, el subespacio de dimensión $(n-1)$ con vector normal $\boldsymbol{\alpha}_i$, generado por los vectores $\boldsymbol{\xi}_j, j \in \{1, \dots, n\}, i \neq j$, más aún, los vectores $\boldsymbol{\alpha}_i$ pueden ser elegidos tales que

$$\begin{cases} \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_i > 0, & i \in \{1, \dots, n\}, \\ \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_j = 0, & i \in \{1, \dots, n\}, \quad j \in \{1, \dots, n\}, \quad i \neq j. \end{cases} \quad (4.4)$$

Los hiperplanos S_i se intersectan en $\mathbf{0}$, de acuerdo a (2.15), para formar una intersección de degeneración mínima de n -ésimo orden se necesitan $N_{n-1} = \binom{n}{2}$ hiperplanos más intersectándose en $\mathbf{0}$.

Sea $S_{i,j}, i \in \{1, \dots, n-1\} \uparrow, j \in \{i+1, \dots, n\} \uparrow$, el hiperplano generado por el vector $\boldsymbol{\xi}_i + \boldsymbol{\xi}_j$ y los vectores $\boldsymbol{\xi}_k, k \in \{1, \dots, n\}, k \neq i, j$. Sea

$$\zeta = \sum_{h=1}^n \boldsymbol{\xi}_h, \quad (4.5)$$

tenemos que $\sigma \zeta \in S_{i,j}, \sigma \in \mathbb{R}$. Así, la intersección de cada una de las N_{n-1} intersecciones con degeneración mínima $S_{i,j}^{(1)}$ es una intersección $\hat{S}_{\mu(n-1)}^{(n-1)}$, esta intersección es una variedad lineal de dimensión 1, por lo que sus puntos \mathbf{z} son de la forma

$$\mathbf{z} = \sigma \zeta.$$

Sea $\boldsymbol{\omega}_{i,j}$ un vector normal a $\hat{S}_{i,j}$, tenemos

$$\begin{cases} \boldsymbol{\omega}_{i,j}^T \boldsymbol{\xi}_i \neq 0, \boldsymbol{\omega}_{i,j}^T \boldsymbol{\xi}_j \neq 0, & i \in \{1, \dots, n-1\}, \quad j \in \{i+1, \dots, n\}, \\ \boldsymbol{\omega}_{i,j}^T (\boldsymbol{\xi}_i + \boldsymbol{\xi}_j) = 0, & i \in \{1, \dots, n-1\}, \quad j \in \{i+1, \dots, n\}, \\ \boldsymbol{\omega}_{i,j}^T \boldsymbol{\xi}_k = 0, & i \in \{1, \dots, n-1\}, \quad j \in \{i+1, \dots, n\}, \\ & k \in \{1, \dots, n\}, \quad k \neq i \neq j. \end{cases} \quad (4.6)$$

Por (2.17), todo vector $\boldsymbol{\omega}_{i,j}$, $i \in \{1, \dots, n-1\}$, $j \in \{i+1, \dots, n\}$, es de la forma

$$\boldsymbol{\omega}_{i,j} = \rho_{i,j} (a_i \boldsymbol{\alpha}_i - a_j \boldsymbol{\alpha}_j), \quad (4.7)$$

donde los coeficientes a_k , $k \in \{1, \dots, n\}$ y $\rho_{i,j}$ son escogidos como en (2.21).

Lema 18. Sean $\boldsymbol{\omega}_{i,j}$ un vector como en (4.6) y $\boldsymbol{\alpha}_i$ y $\boldsymbol{\alpha}_j$ dos vectores como en (4.4), los cuales cumplen con la relación (4.7), con coeficientes (2.21), entonces $a_i \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_i = a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_j > 0$, para todo $i, j \in \{1, \dots, n\}$.

Prueba. Por (4.6) se tiene que

$$\boldsymbol{\omega}_{i,j}^T (\boldsymbol{\xi}_i + \boldsymbol{\xi}_j) = 0. \quad (4.8)$$

De (4.7) y (4.8) se sigue que

$$(a_i \boldsymbol{\alpha}_i^T - a_j \boldsymbol{\alpha}_j^T) (\boldsymbol{\xi}_i + \boldsymbol{\xi}_j) = 0. \quad (4.9)$$

Usando (4.4) y (4.9) tenemos

$$a_i \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_i = a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_j. \quad (4.10)$$

Por (4.4) tenemos que $a_k \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_i > 0$, $k \in \{1, \dots, n\}$, por lo que $a_i > 0$ sí y solo sí $a_j > 0$. Por (2.21) tenemos que $a_1 > 0$, entonces $a_k > 0$. Llamaremos a este valor común por

$$\delta := a_i \boldsymbol{\alpha}_i^T \boldsymbol{\xi}_i > 0, \quad \forall i \in \{1, \dots, n\}. \quad (4.11)$$

■

Para representar las funciones que caracterizan a los hiperplanos utilizados en la partición que se describe en esta sección de manera compacta, usaremos la siguiente notación. Dados los enteros, i , j , q , k_i , k_j y k_q ,

$$\begin{aligned} \text{a)} \quad & \pi_{\boldsymbol{\alpha}_q}^{(k_q), \mathbf{p}}(\mathbf{x}) = \boldsymbol{\alpha}_q^T (\mathbf{x} - k_q \boldsymbol{\xi}_q - \mathbf{p}), \\ \text{b)} \quad & \pi_{\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j}^{-(k_i, k_j), \mathbf{p}}(\mathbf{x}) = \boldsymbol{\alpha}_i^T (\mathbf{x} - k_i \boldsymbol{\xi}_i - \mathbf{p}) - \boldsymbol{\alpha}_j^T (\mathbf{x} - k_j \boldsymbol{\xi}_j - \mathbf{p}). \end{aligned} \quad (4.12)$$

Por lo anterior, la partición $H_{D_{\text{gen}}}$ está dada por los hiperplanos

$$\left\{ \begin{aligned} & \mathbf{x} \in \mathbb{R}^n : \pi_{\boldsymbol{\alpha}_q}^{(k_q), \mathbf{p}}(\mathbf{x}) = 0 \\ & \mathbf{x} \in \mathbb{R}^n : \pi_{\boldsymbol{\alpha}_i, \boldsymbol{\alpha}_j}^{-(k_i, k_j), \mathbf{p}}(\mathbf{x}) = 0 \end{aligned} \right\}, \quad (4.13)$$

donde $q \in \{1, \dots, n\}$, $k_q \in \{0, \dots, m_q - 1\}$, $k_i \in \{0, \dots, m_i - 1\}$, $k_j \in \{0, \dots, m_j - 1\}$, $i \in \{1, \dots, n-1\}$ y $j \in \{i+1, \dots, n\}$.

Ejemplo 28. Continuando con el Ejemplo 27, los hiperplanos generados por los vectores $\boldsymbol{\xi}_1$, $\boldsymbol{\xi}_2$ y $\boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$ son S_1 , S_2 , y $S_{1,2}$, respectivamente, señalados en la Figura 4.2a, estos a su vez tienen vectores normales

$$\boldsymbol{\alpha}_1 = (2 \quad -1)^T, \quad \boldsymbol{\alpha}_2 = (1 \quad 1)^T, \quad \boldsymbol{\omega}_{1,2} = (1 \quad -2)^T,$$

para los cuales, los coeficientes que cumplen con la relación (4.7) son

$$\omega_{1,2} = \rho_{i,j} (a_1 \alpha_1 - a_2 \alpha_2), \quad a_1 = 1, \quad a_2 = 1, \quad \rho_{1,2} = 1.$$

Por lo que la partición $H_{D_{\text{eje}}}$ (ver Figura 4.2b) se forma por los hiperplanos

$$\begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1}^{(0), \mathbf{p}}(\mathbf{x}) = 2x_1 - x_2 + 5 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1}^{(1), \mathbf{p}}(\mathbf{x}) = 2x_1 - x_2 + 2 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_2 \alpha_2}^{(0), \mathbf{p}}(\mathbf{x}) = x_1 + x_2 + 4 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_2 \alpha_2}^{(1), \mathbf{p}}(\mathbf{x}) = x_1 + x_2 + 1 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_2 \alpha_2}^{(2), \mathbf{p}}(\mathbf{x}) = x_1 + x_2 - 2 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1, a_2 \alpha_2}^{-, (0,0), \mathbf{p}}(\mathbf{x}) = x_1 - 2x_2 + 1 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1, a_2 \alpha_2}^{1, (0,1), \mathbf{p}}(\mathbf{x}) = x_1 - 2x_2 + 4 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1, a_2 \alpha_2}^{-, (0,2), \mathbf{p}}(\mathbf{x}) = x_1 - 2x_2 + 7 = 0 \end{cases}, \\ \begin{cases} \mathbf{x} \in \mathbb{R}^2 : \pi_{a_1 \alpha_1, a_2 \alpha_2}^{-, (1,0), \mathbf{p}}(\mathbf{x}) = x_1 - 2x_2 - 2 = 0 \end{cases}.$$

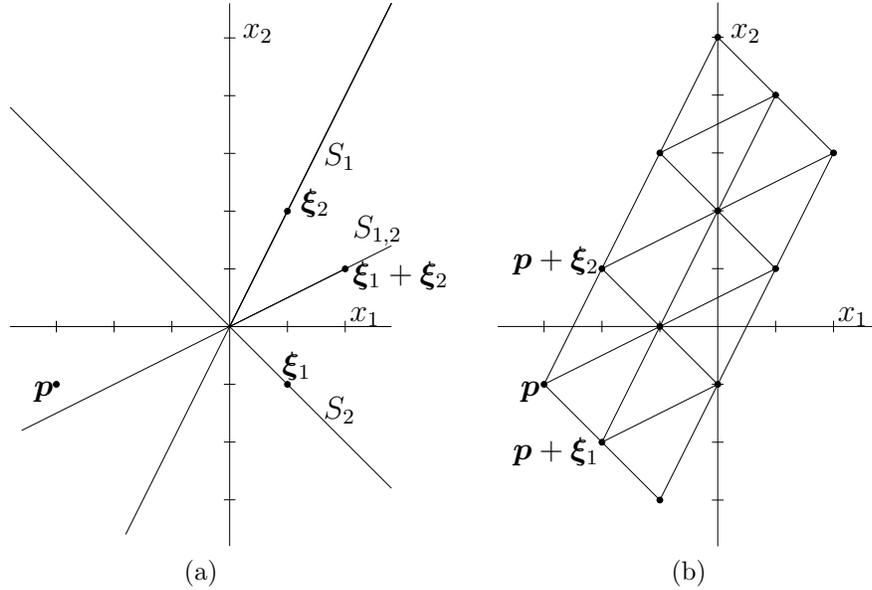


Figura 4.2: (a) Hiperplanos S_1 , S_2 y $S_{1,2}$ generados por los vectores ξ_1 , ξ_2 y $\xi_1 + \xi_2$, respectivamente. (b) Partición $H_{D_{\text{eje}}}$ de D_{eje} .

Por otra parte, dados los s mplices

$$\blacktriangle (\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n}) + \mathbf{u}, \quad \mathbf{u} \in \mathbb{R}^n, \tag{4.14}$$

donde

$$\mathbf{v}_{r_k} = \sum_{i=0}^k \boldsymbol{\xi}_{r_i} \quad \forall k \in \{0, \dots, n\},$$

con $r_1 \in \{1, \dots, n - k + 1\}$, $r_i \in \{r_{i-1} + 1, \dots, n - k + i\}$, $i \in \{2, \dots, n\}$ y $\mathbf{v}_{r_0} = \boldsymbol{\xi}_{r_0} = \mathbf{0}$, tenemos el siguiente lema.

Lema 19. *Los hiperplanos de la partición $H_{D_{gen}}$ y la frontera ∂D_{gen} de D_{gen} subdividen a D_{gen} en sımplices (4.14).*

Prueba. Para probar este lema primero veamos que la particion $H_{D'_{gen}}$ dada por (4.14) con $k_q, k_i, k_j = 0$, $q \in \{1, \dots, n\}$, $i \in \{1, \dots, n - 1\}$ y $j \in \{i + 1, \dots, n\}$ y $\mathbf{p} = \mathbf{0}$, esta dada por los hiperplanos

$$\begin{aligned} a) & \left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_q \alpha_q}^{(0), \mathbf{0}}(\mathbf{x}) = 0 \right\}, \\ b) & \left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_i \alpha_i, a_j \alpha_j}^{(0,0), \mathbf{0}}(\mathbf{x}) = 0 \right\} \end{aligned} \quad (4.15)$$

y la frontera $\partial D'_{gen}$ de $D'_{gen} = \{M^T \mathbf{x} : \mathbf{x} = (x_1 \dots x_n), 0 \leq x_i \leq 1, i = 1, \dots, n\}$ particionan a D'_{gen} en sımplices (4.14) de la forma $\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n})$.

Veremos que cada arista de los sımplices (combinacion convexa de dos vertices) resulta de la interseccion de un conjunto de hiperplanos en la particion $H_{D'_{gen}}$ y $\partial D'_{gen}$.

Sea $\mathbf{z}' \in D'_{gen}$, de (4.2) \mathbf{z}' puede ser expresado por $\mathbf{z}' = M\mathbf{z}$, donde cada i -esimo elemento de \mathbf{z} es $0 \leq z_i \leq 1$, con $i \in \{1, \dots, n\}$. Por lo que, si existe una i -esima componente de \mathbf{z} tal que $z_i = 1$, tenemos que $\mathbf{z}' \in \partial D'_{gen}$. Ademas, de (4.2) se sigue que

$$M^T \mathbf{z} = \sum_{i=1}^n z_i \boldsymbol{\xi}_i.$$

Sea $\mathbf{z}' \in D'_{gen}$ y un vector $a_j \boldsymbol{\alpha}_j$, con $j \in \{1, \dots, n\}$, supongamos que $a_j \boldsymbol{\alpha}_j^T \mathbf{z}' = \delta$, por (4.16) tenemos que

$$a_j \boldsymbol{\alpha}_j^T \left(\sum_{i=1}^n z_i \boldsymbol{\xi}_i \right) = z_i a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_1 + \dots + z_n a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_n = \delta. \quad (4.16)$$

Ademas, por (4.4), tenemos que $a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_{(i)} = 0$, $\forall i \in \{1, \dots, n\}$, $i \neq j$, por lo que (4.16) se reduce a

$$z_j a_j \boldsymbol{\alpha}_j^T \boldsymbol{\xi}_j = \delta. \quad (4.17)$$

Por (4.11) y (4.17), tenemos $z_j \delta = \delta$, por lo que $z_j = 1$. Se sigue que \mathbf{z} esta en $\partial D'_{gen}$. Ası, para todo $\mathbf{x} \in D'_{gen}$ tal que $a_j \boldsymbol{\alpha}_j^T \mathbf{x} = \delta$, $\mathbf{x} \in \partial D'_{gen}$.

Sea \mathbf{y} un punto en una arista de un sımplice $\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n})$, puesto que los puntos en los bordes pueden ser vistos como una combinacion convexa entre dos vertices del sımplice, tenemos

$$\mathbf{y} = b \left(\sum_{i=0}^{k_1} \boldsymbol{\xi}_{r_i} \right) + (1-b) \left(\sum_{i=0}^{k_2} \boldsymbol{\xi}_{r_i} \right),$$

con $b \in [0, 1]$, $k_1, k_2 \in \{1, \dots, n\}$, $k_1 \neq k_2$, $r_1 \in \{1, \dots, n - k_1 + 1\}$, $r_i \in \{r_{i-1} + 1, \dots, n - k_1 + i\}$, $i \in \{2, \dots, n\}$, $r_j \in \{r_{j-1} + 1, \dots, n - k_2 + j\}$, $j \in \{2, \dots, n\}$. Supongamos sin pérdida de generalidad que $k_2 > k_1$, así

$$\mathbf{y} = \left(\sum_{i=0}^{k_1} \boldsymbol{\xi}_{r_i} \right) + (1-b) \left(\sum_{i=k_1+1}^{k_2} \boldsymbol{\xi}_{r_i} \right).$$

Además, dado un $a_j \boldsymbol{\alpha}_j$, $j \in \{1, \dots, n\}$, tenemos por (4.4) y por (4.11) que $a_j \boldsymbol{\alpha}_j^T \mathbf{y} \in \{0, (1-b)\delta, \delta\}$. De aquí se desprenden los siguientes casos.

- Supongamos que $a_j \boldsymbol{\alpha}_j^T \mathbf{y} = \delta$ para algún $j \in \{1, \dots, n\}$, entonces \mathbf{y} está en $\partial D'$. Además, supongamos $a_h \boldsymbol{\alpha}_h^T \mathbf{y} = \delta$ para algún $h \in \{1, \dots, n\}$, $j \neq h$, tenemos $a_h \boldsymbol{\alpha}_h^T \mathbf{y} - a_j \boldsymbol{\alpha}_j^T \mathbf{y} = 0$, por lo que \mathbf{y} está en el hiperplano $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_h \boldsymbol{\alpha}_h, a_j \boldsymbol{\alpha}_j}^{-(0,0),0}(\mathbf{x}) = 0 \right\}$, el cual está en la partición $H_{D'_{\text{gen}}}$.
- Supongamos que $a_j \boldsymbol{\alpha}_j^T \mathbf{y} = 0$ para algún $j \in \{1, \dots, n\}$, entonces \mathbf{y} está en $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_j \boldsymbol{\alpha}_j}^{(0),0}(\mathbf{x}) = 0 \right\}$, el cual está en la partición $H_{D'_{\text{gen}}}$. Además, supongamos $a_h \boldsymbol{\alpha}_h^T \mathbf{y} = \delta$ para algún $h \in \{1, \dots, n\}$, $j \neq h$, tenemos $a_h \boldsymbol{\alpha}_h^T \mathbf{y} - a_j \boldsymbol{\alpha}_j^T \mathbf{y} = 0$, por lo que \mathbf{y} está en el hiperplano $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_h \boldsymbol{\alpha}_h, a_j \boldsymbol{\alpha}_j}^{-(0,0),0}(\mathbf{x}) = 0 \right\}$, el cual está en la partición $H_{D'_{\text{gen}}}$.
- Supongamos que $a_j \boldsymbol{\alpha}_j^T \mathbf{y} = (1-b)\delta$ para un único $j \in \{1, \dots, n\}$, entonces $a_h \boldsymbol{\alpha}_h^T \mathbf{y} \in \{0, \delta\}$, $h \in \{1, \dots, n\}$, $j \neq h$, por lo que se cumplen alguno de los dos casos anteriores. Además, si $a_j \boldsymbol{\alpha}_j^T \mathbf{y} = (1-b)\delta$ para algún $j \in \{1, \dots, n\}$ y $a_h \boldsymbol{\alpha}_h^T \mathbf{y} = (1-b)\delta$, $h \in \{1, \dots, n\}$, $j \neq h$, entonces $a_h \boldsymbol{\alpha}_h^T \mathbf{y} - a_j \boldsymbol{\alpha}_j^T \mathbf{y} = 0$, por lo que \mathbf{y} está en el hiperplano $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_h \boldsymbol{\alpha}_h, a_j \boldsymbol{\alpha}_j}^{-(0,0),0}(\mathbf{x}) = 0 \right\}$, el cual está en la partición $H_{D'_{\text{gen}}}$.

Por lo que la partición $H_{D'_{\text{gen}}}$ y la frontera $\partial D'_{\text{gen}}$ de D'_{gen} particionan a D'_{gen} en símlices (4.14) de la forma $\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n})$.

Ahora, veamos que los hiperplanos $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_q \boldsymbol{\alpha}_q}^{(k_q), \mathbf{p}}(\mathbf{x}) = 0 \right\}$ y $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_i \boldsymbol{\alpha}_i, a_j \boldsymbol{\alpha}_j}^{-(k_i, k_j), \mathbf{p}}(\mathbf{x}) = 0 \right\}$ con $q \in \{1, \dots, n\}$, $i \in \{1, \dots, n-1\}$ y $j \in \{i+1, \dots, n\}$, $k_q \in \{0, \dots, m_q - 1\}$, $k_i \in \{0, \dots, m_i - 1\}$, $k_j \in \{0, \dots, m_j - 1\}$ y la frontera ∂D_{gen} de D_{gen} particionan a D_{gen} en símlices (4.14).

Sea $\mathbf{u} = \mathbf{p} + \sum_{i=1}^n k_i \boldsymbol{\xi}_i$, con $k_i \in \{0, \dots, m_i - 1\}$, y la transformación $T_{\mathbf{u}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$,

$$T_{\mathbf{u}}(\mathbf{x}) = \mathbf{x} + \mathbf{u} = \mathbf{z}, \quad (4.18)$$

la cual traslada los símlices $\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n})$ en D'_{gen} a $D_{\text{gen}} = D'_{\text{gen}} + \mathbf{u}$.

Puesto que los hiperplanos a) y b) de (4.15) particionan a D' en símlices $\blacktriangle(\mathbf{v}_{r_0}, \dots, \mathbf{v}_{r_n})$ en D'_{gen} , moviendo estos hiperplanos, estos subdividen a $D_{\text{gen}} = D'_{\text{gen}} + \mathbf{u}$ en símlices (4.14). Resta ver que estos hiperplanos trasladados están en la partición $H_{D_{\text{gen}}}$.

- Sea \mathbf{x} un punto en un hiperplano a), por (4.18) tenemos que

$$a_q \boldsymbol{\alpha}_q^T (\mathbf{z} - \mathbf{u}) = a_q \boldsymbol{\alpha}_q^T \left(\mathbf{z} - \mathbf{p} - \sum_{i=1}^n k_i \boldsymbol{\xi}_i \right) = 0,$$

de (4.4) tenemos que $a_q \boldsymbol{\alpha}_q^T (\mathbf{z} - k_q \boldsymbol{\xi}_q - \mathbf{p}) = 0$, por lo que $\mathbf{z} \in \left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_q \boldsymbol{\alpha}_q}^{(k_q), \mathbf{p}}(\mathbf{x}) = 0 \right\}$ con $q \in \{1, \dots, n\}$, $k_q \in \{0, \dots, m_q - 1\}$, el cual es un hiperplano en la partición $H_{D_{\text{gen}}}$.

- Sea \mathbf{x} un punto en un hiperplano b), por (4.18) tenemos que

$$\begin{aligned} a_i \boldsymbol{\alpha}_i^T (\mathbf{z} - \mathbf{u}) - a_j \boldsymbol{\alpha}_j^T (\mathbf{z} - \mathbf{u}) = \\ a_i \boldsymbol{\alpha}_i^T \left(\mathbf{z} - \mathbf{p} - \sum_{q=1}^n k_q \boldsymbol{\xi}_q \right) - a_j \boldsymbol{\alpha}_j^T \left(\mathbf{z} - \mathbf{p} - \sum_{q=1}^n k_q \boldsymbol{\xi}_q \right) = 0, \end{aligned}$$

de (4.4) tenemos que $a_i \boldsymbol{\alpha}_i^T (\mathbf{z} - k_i \boldsymbol{\xi}_i - \mathbf{p}) - a_j \boldsymbol{\alpha}_j^T (\mathbf{z} - k_j \boldsymbol{\xi}_j - \mathbf{p}) = 0$, por lo que $\mathbf{z} \in \left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_i \boldsymbol{\alpha}_i, a_j \boldsymbol{\alpha}_j}^{-(k_i, k_j), \mathbf{p}}(\mathbf{x}) = 0 \right\}$ con $k_i \in \{0, \dots, m_i - 1\}$, $k_j \in \{0, \dots, m_j - 1\}$, $i \in \{1, \dots, n - 1\}$ y $j \in \{i + 1, \dots, n\}$, el cual es un hiperplano en la partición $H_{D_{\text{gen}}}$.

Por lo que la partición $H_{D_{\text{gen}}}$ y la frontera ∂D_{gen} de D_{gen} particionan a D_{gen} en simplices de la forma (4.14). ■

Ejemplo 29. Continuando con el Ejemplo 28, los simplices resultantes de esta partici3n son de la forma $\mathbf{u} + \blacktriangle(\mathbf{0}, \boldsymbol{\xi}_1, \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2)$ y $\mathbf{u} + \blacktriangle(\mathbf{0}, \boldsymbol{\xi}_2, \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2)$, con $\mathbf{u} = \mathbf{p} + \sum_{q=1}^n k_q \boldsymbol{\xi}_q$, $k_1 \in \{0, 1\}$ y $k_2 \in \{0, 1, 2\}$ (ver Figura 4.2b).

Por razones numericas que seran vistas en la siguiente subsecci3n, los vertices de los simplices que particionan a D_{gen} seran ordenados en un vector $V_{D_{\text{gen}}}$ de la siguiente manera: primero, los vertices son separados por clases, cada clase esta dada por el numero de vectores base usados en su construcci3n; tambien, los vertices pertenecientes a una misma clase $r \in \{0, \dots, n\}$ son ordenados en un vector $V_{D_{\text{gen}}}^{(r)}$ como es descrito a continuaci3n.

- Vertices de clase 0 ($V_{D_{\text{gen}}}^{(0)}$): vertices sin vectores $\boldsymbol{\xi}$ en su construcci3n, $\mathbf{x} = \mathbf{p}$.
- Vertices de clase 1 ($V_{D_{\text{gen}}}^{(1)}$): $\sum_{k_1=1}^n m_{k_1}$ vertices vector $\boldsymbol{\xi}$ en su construcci3n:

$$\mathbf{x} = q_{k_1} \boldsymbol{\xi}_{k_1} + \mathbf{p},$$

donde $k_1 \in \{1, \dots, n - 1\} \uparrow$. Ademas, para cada k_1 fijo, tenemos $q_{k_1} \in \{1, \dots, m_{k_1}\} \uparrow$.

- Vertices de clase $2 \leq r \leq n$ ($V_{D_{\text{gen}}}^{(r)}$): $\left(\sum_{k_1=1}^{n-r+1} m_{k_1} \sum_{k_2=k_1+1}^{n-r+2} m_{k_2} \cdots \sum_{k_r=k_{r-1}+1}^n m_{k_r} \right)$ vertices con r vectores $\boldsymbol{\xi}$ en su construcci3n:

$$\mathbf{x} = \mathbf{p} + \sum_{p=1}^r q_{k_p} \boldsymbol{\xi}_{k_p}, \quad (4.19)$$

donde las tuplas $\{k_1, \dots, k_r\}$, $k_1 \in \{1, \dots, n - r + 1\}$ y $k_j \in \{k_{j-1}, \dots, n - r + j\}$, $\forall j \in \{2, \dots, r\}$ es A ordenado. Además, para cada tupla fija de índices $\{k_1, \dots, k_r\}$, $q_{k_i} \in \{1, \dots, m_{k_i}\}$, $i \in \{1, \dots, r\}$, es B ordenado.

Así, $V_{D_{\text{gen}}}$ se forma de la concatenación de los vectores de vértices de clase r , $V_{D_{\text{gen}}}^{(r)}$, $\{0, \dots, n\} \uparrow$.

Ejemplo 30. Continuando con el Ejemplo 29, los vértices de la partición $H_{D_{\text{eje}}}$ en D_{eje} se ordenan de la siguiente manera:

- $V_{D_{\text{eje}}}^{(0)}: (-3 \ -1)^T$,
- $V_{D_{\text{eje}}}^{(1)}: (-2 \ 2)^T, (-1 \ -3)^T, (-2 \ 1)^T, (-1 \ 3)^T, (0 \ 5)^T$,
- $V_{D_{\text{eje}}}^{(2)}: (-1 \ 0)^T, (0 \ 2)^T, (1 \ 4)^T, (0 \ -1)^T, (1 \ 1)^T, (2 \ 3)^T$.

4.1.2 Conjunto generador del espacio $PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$.

Sea $\mathfrak{S}_{H_{D_{\text{gen}}}} = \left\{ 1, \pi_{a_k \alpha_k}^{(j_k) \cdot \mathbf{p}}, k \in \{1, \dots, n\}, j_k \in \{0, \dots, m_k - 1\} \right\}$ con $\pi_{a_k \alpha_k}^{(j_k) \cdot \mathbf{p}}$ como en a) de (4.12) y la función $\gamma^{(k)}(\cdot, \dots, \cdot)$ definida en (3.5). Un vector $\boldsymbol{\Lambda}_{D_{\text{gen}}}$ de funciones HL-CPWL es definido de la siguiente manera: primero, las funciones son separadas por su nivel de anidación $l \in \{0, \dots, n\}$ en vectores $\boldsymbol{\Lambda}_{D_{\text{gen}}}^{(l)}$, además, funciones con un mismo nivel de anidación son ordenadas como se muestra a continuación:

- Funciones con nivel de anidación igual a 0 ($\boldsymbol{\Lambda}_{D_{\text{gen}}}^{(0)}$): la función constante $\gamma^{(0)}(1) = 1$.
- Funciones con nivel de anidación igual a 1 ($\boldsymbol{\Lambda}_{D_{\text{gen}}}^{(1)}$): $\sum_{k_1=1}^n m_{k_1}$ funciones de la forma:

$$\gamma^{(1)} \left(\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) \cdot \mathbf{p}} \right),$$

con $k_1 \in \{1, \dots, n\} \uparrow$. Además, para cada k_1 fijo, tenemos $j_{k_1} \in \{0, \dots, m_{k_1} - 1\} \uparrow$.

- Funciones con nivel de anidación igual a $2 \leq l \leq n$ ($\boldsymbol{\Lambda}_{D_{\text{gen}}}^{(l)}$): $\left(\sum_{k_1=1}^{n-l+1} m_{k_1} \sum_{k_2=k_1+1}^{n-l+2} m_{k_2} \dots \sum_{k_l=k_{l-1}+1}^n m_{k_l} \right)$ funciones de la forma:

$$\gamma^{(l)} \left(\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) \cdot \mathbf{p}}, \dots, \pi_{a_{k_l} \alpha_{k_l}}^{(j_{k_l}) \cdot \mathbf{p}} \right), \quad (4.20)$$

donde las tuplas de índices $\{k_1, \dots, k_l\}$, $k_1 \in \{1, \dots, n - l + 1\}$, $k_j \in \{k_{j-1}, \dots, n - l + j\}$, $\forall j \in \{2, \dots, l\}$ es A ordenada. Además, cada tupla fija de índices $\{k_1, \dots, k_l\}$, $j_{k_i} \in \{0, \dots, m_{k_i} - 1\}$, con $i \in \{1, \dots, n\}$, es B ordenada.

Así, $\Lambda_{D_{\text{gen}}}$ se forma de la concatenación de los vectores de funciones HL-CPWL con nivel de anidación l , $\Lambda_{D_{\text{gen}}}^{(l)}$, $\{0, \dots, n\}$.

Ejemplo 31. Continuando con el Ejemplo 27, los vértices de la partición H_D en D se ordenan de la siguiente manera:

- $\Lambda_{D_{\text{eje}}}^{(0)} : \gamma^{(0)}(1)$.
- $\Lambda_{D_{\text{eje}}}^{(1)} : \gamma^{(1)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}\right), \gamma^{(1)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}\right), \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right), \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right), \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right)$.
- $\Lambda_{D_{\text{eje}}}^{(2)} : \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right), \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right), \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right), \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right), \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right), \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right)$.

Como se vio anteriormente, los elementos en $\Lambda_{D_{\text{gen}}}$ constan de funciones de valor absoluto anidadas, más aún, estas son funciones HL-CPWL, por lo que estas se encuentran definidas dentro de una partición lineal, la cual se precisa en el siguiente lema.

Lema 20. *Las funciones HL-CPWL en $\Lambda_{D_{\text{gen}}}$ están definidas en la partición $H_{D_{\text{gen}}}$ y ningún otro hiperplano fuera de $H_{D_{\text{gen}}}$ fue añadido a la partición.*

Prueba. Dado el vector $\Lambda_{D_{\text{gen}}}$ de funciones de la forma (4.20), puesto que una función PWL se define por sus funciones locales afines (2.4) y las cuales cambian cuando se cruza un hiperplano de su partición, para probar este Lema, veamos que los cambios de valor de las funciones en $\Lambda_{D_{\text{gen}}}$ ocurren solamente al cruzar un hiperplano de la partición $H_{D_{\text{gen}}}$. La prueba la realizamos por inducción en el nivel de anidación de las funciones en $\Lambda_{D_{\text{gen}}}$.

Sea $\gamma^{(1)}\left(\pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}\right)$, con $k_1 \in \{1, \dots, n\}$ y $j_{r_1} \in \{0, \dots, m_{r_1} - 1\}$ una función en $\Lambda_D^{(1)}$, reduciendo (3.2) tenemos que $\gamma^{(1)}\left(\pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}\right) = \frac{1}{2}\left(\pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}} + \left|\pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}\right|\right)$.

Dado un $\mathbf{x} \in \mathbb{R}^n$, tenemos

$$\gamma^{(1)}\left(\pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}\right)(\mathbf{x}) = \begin{cases} \pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}(\mathbf{x}) & \text{si } \pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}(\mathbf{x}) \geq 0, \\ 0 & \text{si } \pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}(\mathbf{x}) < 0. \end{cases}$$

Por lo que hay un cambio al cruzar $\left\{\mathbf{x} \in \mathbb{R}^n : \pi_{a_{r_1}\alpha_{r_1}}^{(j_{r_1}),\mathbf{p}}(\mathbf{x}) = 0\right\}$, el cual es un hiperplano en la partición $H_{D_{\text{gen}}}$.

Supongamos que las funciones HL-CPWL en $\Lambda_{D_{\text{gen}}}^{(r)}$ están definidas en la partición $H_{D_{\text{gen}}}$, es decir que las funciones de la forma

$$\gamma^{(r)}\left(\pi_{a_{k_1}\alpha_{k_1}}^{(j_{k_1}),\mathbf{p}}, \dots, \pi_{a_{k_r}\alpha_{k_r}}^{(j_{k_r}),\mathbf{p}}\right),$$

con $k_1 \in \{1, \dots, n - r + 1\}$, $k_h \in \{k_{h-1}, \dots, n - r + h\}$, $\forall h \in \{2, \dots, r\}$ y $j_{k_i} \in \{0, \dots, m_{k_i} - 1\}$, con $i \in \{1, \dots, n\}$ están definidas en la partición $H_{D_{\text{gen}}}$. Sea

$$\gamma^{(r+1)} \left(\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right),$$

con $k_1 \in \{1, \dots, n - (r + 1) + 1\}$, $k_j \in \{k_{j-1}, \dots, n - (r + 1) + j\}$, $\forall j \in \{2, \dots, (r + 1)\}$ y $j_{k_i} \in \{0, \dots, m_{k_i} - 1\}$, con $i \in \{1, \dots, n\}$ una función en $\Lambda_D^{(r+1)}$. Por (3.5) tenemos que

$$\gamma^{(r+1)} \left(\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right) = \begin{cases} \text{a) } \pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}} & \text{si } 0 \leq \pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}} \leq \gamma^{(r)} \left(\pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right), \\ \text{b) } \gamma^{(r)} \left(\pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right) & \text{si } 0 \leq \gamma^{(r)} \left(\pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right) \leq \pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}}, \\ \text{c) } 0 & \text{si } \pi_{a_{k_q} \alpha_{k_q}}^{(j_{k_q}) : \mathbf{p}} \leq 0 \text{ para algún } q \in \{1, \dots, k + 1\}, \end{cases}$$

de donde se desprenden tres casos.

- (i) Supongamos que ocurre b), por hipótesis $\gamma^{(r+1)} \left(\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right)$ está definida en la partición $H_{D_{\text{gen}}}$
- (ii) Supongamos que ocurre c), tenemos que los cambios ocurren cuando se cruza el hiperplano $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_{k_q} \alpha_{k_q}}^{(j_{k_q}) : \mathbf{p}}(\mathbf{x}) = 0 \right\}$, $q \in \{1, \dots, r + 1\}$ el cual es un hiperplano en la partición $H_{D_{\text{gen}}}$.
- (iii) Supongamos que ocurre a) y b) tenemos que

$$\pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}} = \gamma^{(r)} \left(\pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right), \quad (4.21)$$

por (3.5) tenemos que

$$\gamma^{(r)} \left(\pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \dots, \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}} \right) = \begin{cases} \pi_{a_{k_2} \alpha_{k_2}}^{(j_{k_2}) : \mathbf{p}}, \\ \vdots \\ \pi_{a_{k_{r+1}} \alpha_{k_{r+1}}}^{(j_{k_{r+1}}) : \mathbf{p}}, \end{cases} \quad (4.22)$$

en la región $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_{k_q} \alpha_{k_q}}^{(j_{k_q}) : \mathbf{p}}(\mathbf{x}) \geq 0 \right\}$, $\forall q \in \{2, \dots, r + 1\}$, las cuales son funciones afines, por (4.21) y (4.22) los cambios de expresión ocurren en $\left\{ \mathbf{x} \in \mathbb{R}^n : \pi_{a_{k_1} \alpha_{k_1}}^{(j_{k_1}) : \mathbf{p}}(\mathbf{x}) - \pi_{a_{k_q} \alpha_{k_q}}^{(j_{k_q}) : \mathbf{p}}(\mathbf{x}) = 0 \right\}$, es decir, en $\left\{ \mathbf{x} \in \mathbb{R}^n : a_{k_1} \alpha_{k_1}^T (\mathbf{x} - j_{k_1} \boldsymbol{\xi}_{k_1} - \mathbf{p}) - a_{k_q} \alpha_{k_q}^T (\mathbf{x} - j_{k_q} \boldsymbol{\xi}_{k_q} - \mathbf{p}) = \pi_{a_{k_1} \alpha_{k_1}, a_{k_q} \alpha_{k_q}}^{(j_{k_1}, j_{k_q}) : \mathbf{p}}(\mathbf{x}) = 0 \right\}$, el cual es un hiperplano en la partición $H_{D_{\text{gen}}}$.

Por lo anterior, las funciones PWL en $\Lambda_{D_{\text{gen}}}$ están definidas en la partición $H_{D_{\text{gen}}}$ y ningún otro hiperplano fue añadido a la partición. \blacksquare

Sea $A_{D_{\text{gen}}} \in \mathbb{R}^{h \times l}$ la matriz cuyas filas son obtenidas de la evaluación de las funciones $\Lambda_{D_{\text{gen}}}$ en los vértices de $V_{D_{\text{gen}}}$, es decir,

$$A_{D_{\text{gen}}} = \begin{bmatrix} \Lambda_{D_{\text{gen}1}}(V_{D_{\text{gen}1}}) & \cdots & \Lambda_{D_{\text{gen}h}}(V_{D_{\text{gen}1}}) \\ \vdots & & \vdots \\ \Lambda_{D_{\text{gen}1}}(V_{D_{\text{gen}l}}) & \cdots & \Lambda_{D_{\text{gen}h}}(V_{D_{\text{gen}l}}) \end{bmatrix}, \quad (4.23)$$

donde $\Lambda_{D_{\text{gen}i}}$ es la i -ésima función en $\Lambda_{D_{\text{gen}}}$, $i \in \{1, \dots, h\}$, y $V_{D_{\text{gen}j}}$ es el j -ésimo vértice en $V_{D_{\text{gen}}}$, $j \in \{1, \dots, l\}$.

Lema 21. $A_{D_{\text{gen}}}$ es una matriz cuadrada.

Prueba. Para esto probaremos que $V_{D_{\text{gen}}}$ tiene el mismo número de elementos que $\Lambda_{D_{\text{gen}}}$. Dado un $r \in \{0, \dots, n\}$, por (4.19), los vértices en $V_{D_{\text{gen}}}^{(r)}$ son de la forma

$$\mathbf{x} = q_{k_1''} \boldsymbol{\xi}_{k_1''} + \dots + q_{k_r''} \boldsymbol{\xi}_{k_r''} + \mathbf{p}, \quad (4.24)$$

donde $k_i'' \in \{k_{i-1}'' + 1, \dots, n - r + i\}$, $i \in \{1, \dots, r\}$ y $q_{k_j''} \in \{1, \dots, m_{k_j''}\}$, $j \in \{1, \dots, r\}$.

Por (4.20), las funciones en $\Lambda_{D_{\text{gen}}}^{(l)}$ son de la forma

$$\gamma^{(l)} \left(\begin{matrix} (j_{k_1}') \cdot \mathbf{p} \\ \pi_{a_{k_1}'} \boldsymbol{\alpha}_{k_1}' \\ \dots, \pi_{a_{k_l}'} \boldsymbol{\alpha}_{k_l}' \end{matrix} \right), \quad (4.25)$$

donde $k_i' \in \{k_{i-1}' + 1, \dots, n - l + i\}$, $i \in \{1, \dots, l\}$ y $j_{k_j}' \in \{0, \dots, m_{k_j}' - 1\}$, $j \in \{1, \dots, l\}$.

De (4.24) y (4.1.2), $V_{D_{\text{gen}}}^{(r)}$ tiene el mismo número de elementos que $\Lambda_{D_{\text{gen}}}^{(l)}$, si $l = r$. Se sigue que $V_{D_{\text{gen}}}$ tiene el mismo número de elementos que $\Lambda_{D_{\text{gen}}}$, así, $A_{D_{\text{gen}}}$ es una matriz cuadrada. \blacksquare

Las ventajas numéricas de la construcción ordenada de $V_{D_{\text{gen}}}$ y $\Lambda_{D_{\text{gen}}}$ pueden verse en la prueba del siguiente lema.

Lema 22. $A_{D_{\text{gen}}}$ es no singular.

Prueba. Para probar esto veremos que $A_{D_{\text{gen}}}$ es triangular inferior y que sus elementos en la diagonal principal son diferentes de 0. Sea \mathbf{x} un vértice en $V_D^{(r)}$ (como en (4.24)) de la forma

$$\mathbf{x} = q_{k_1''} \boldsymbol{\xi}_{k_1''} + \dots + q_{k_r''} \boldsymbol{\xi}_{k_r''} + \mathbf{p},$$

donde $k_i'' \in \{k_{i-1}'' + 1, \dots, n - r + i\}$, $i \in \{1, \dots, r\}$ y $q_{k_j''} \in \{1, \dots, m_{k_j''}\}$, $j \in \{1, \dots, r\}$. Y sea $g(\mathbf{x})$ una función en $\Lambda_D^{(l)}$ (como en (4.1.2)) de la forma

$$\gamma^{(l)} \left(\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}, \dots, \pi_{a_{k'_l}'} \alpha_{k'_l}^{\mathbf{p}} \right),$$

donde $k'_i \in \{k'_{i-1} + 1, \dots, n - l + i\}$, $i \in \{1, \dots, l\}$ y $j_{k'_j} \in \{0, \dots, m_{k'_j} - 1\}$, $j \in \{1, \dots, l\}$.

Dada la evaluación de la función $g(\mathbf{x})$, esta es un elemento en la diagonal de $A_{D_{\text{gen}}}$ sí y solo sí:

- $r = l$.
- $k'_v = k''_v$, $v \in \{1, \dots, l\}$.
- $q_{k''_v} = j_{k'_v} + 1$, $v \in \{1, \dots, l\}$.

Para simplificar la prueba, las entradas en $A_{D_{\text{gen}}}$ serán clasificadas, primero, por la clase r de los vértices en los que es evaluada la función g . Además, las funciones g se organizan por su nivel de anidación para cada clase de vértices fija.

Vértices de clase cero ($r = 0$): entonces $\mathbf{x} = \mathbf{p}$. Supongamos $l = 0$, tenemos que $\gamma^{(0)}(1) = 1$. Supongamos $l > r$, tenemos que $\pi_{a_{k'_v}'} \alpha_{k'_v}^{\mathbf{p}}(\mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T(\mathbf{p} - j_{k'_v} \boldsymbol{\xi}_{k'_v} - \mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T(-j_{k'_v} \boldsymbol{\xi}_{k'_v}) = -j_{k'_v} a_{k'_v} \alpha_{k'_v}^T \boldsymbol{\xi}_{k'_v}$, por (4.11) tenemos $\pi_{a_{k'_v}'} \alpha_{k'_v}^{\mathbf{p}}(\mathbf{p}) = -j_{k'_v} \delta \leq 0$, $\forall v \in \{1, \dots, l\}$ y $\forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Vértices de clase uno ($r = 1$): supongamos $l = r = 1$, tenemos los siguientes casos:

- Supongamos $k'_1 \neq k''_1$, tenemos $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = a_{k'_1} \alpha_{k'_1}^T(\mathbf{p} + q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1} - \mathbf{p}) = a_{k'_1} \alpha_{k'_1}^T(q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1})$, por (4.4), se sigue $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = a_{k'_1} \alpha_{k'_1}^T(-j_{k'_1} \boldsymbol{\xi}_{k'_1}) = -j_{k'_1} a_{k'_1} \alpha_{k'_1}^T \boldsymbol{\xi}_{k'_1}$, por (4.11) tenemos $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = -j_{k'_1} \delta \leq 0$, $\forall j_{k'_1} \in \{0, \dots, m_{k'_1} - 1\}$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.
- Supongamos $k'_1 = k''_1$ y $q_{k''_1} = j_{k'_1} + 1$, es decir, que $g(\mathbf{x})$ es un elemento en la diagonal de $A_{D_{\text{gen}}}$, tenemos $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = a_{k'_1} \alpha_{k'_1}^T(\mathbf{p} + q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1} - \mathbf{p}) = a_{k'_1} \alpha_{k'_1}^T(q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1}) = a_{k'_1} \alpha_{k'_1}^T((j_{k'_1} + 1) \boldsymbol{\xi}_{k'_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1}) = a_{k'_1} \alpha_{k'_1}^T \boldsymbol{\xi}_{k'_1}$, por (4.11) se sigue $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = \delta$, por (3.5) tenemos que $g(\mathbf{x}) = \delta$.
- Supongamos $k'_1 = k''_1$ y $q_{k''_1} \leq j_{k'_1}$, es decir, que $g(\mathbf{x})$ es un elemento sobre la diagonal principal de $A_{D_{\text{gen}}}$, tenemos $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = a_{k'_1} \alpha_{k'_1}^T(\mathbf{p} + q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1} - \mathbf{p}) = a_{k'_1} \alpha_{k'_1}^T(q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_1} \boldsymbol{\xi}_{k'_1}) = (q_{k''_1} - j_{k'_1}) a_{k'_1} \alpha_{k'_1}^T \boldsymbol{\xi}_{k'_1}$, por (4.11) se sigue $\pi_{a_{k'_1}'} \alpha_{k'_1}^{\mathbf{p}}(\mathbf{x}) = (q_{k''_1} - j_{k'_1}) \delta \leq 0$, por (3.5) tenemos que $g(\mathbf{x}) = 0$.

Supongamos $l > r = 1$, de donde se desprenden dos casos:

- Supongamos $k'_v \neq k''_1, \forall v \in \{1, \dots, l\}$, tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (\mathbf{p} + q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_v} \boldsymbol{\xi}_{k'_v} - \mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T (q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_v} \boldsymbol{\xi}_{k'_v})$, por (4.4), se sigue $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (-j_{k'_v} \boldsymbol{\xi}_{k'_v}) = -j_{k'_v} a_{k'_v} \alpha_{k'_v}^T \boldsymbol{\xi}_{k'_v}$, por (4.11) tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = -j_{k'_v} \delta \leq 0, \forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) se sigue que $g(\mathbf{x}) = 0$.
- Supongamos que existe un $v^* \in \{1, \dots, l\}$ tal que $k'_{v^*} = k''_1$, puesto que $l > r$, existe un $v \in \{1, \dots, l\}$ tal que $k'_v \neq k''_1, v \neq v^*$, tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (\mathbf{p} + q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_v} \boldsymbol{\xi}_{k'_v} - \mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T (q_{k''_1} \boldsymbol{\xi}_{k''_1} - j_{k'_v} \boldsymbol{\xi}_{k'_v})$, por (4.4), se sigue $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (-j_{k'_v} \boldsymbol{\xi}_{k'_v}) = -j_{k'_v} a_{k'_v} \alpha_{k'_v}^T \boldsymbol{\xi}_{k'_v}$, por (4.11) tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = -j_{k'_v} \delta \leq 0, \forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) se sigue que $g(\mathbf{x}) = 0$.

Vértices de clase r ($n \geq r > 1$): supongamos $l = r > 1$, tenemos los siguientes casos:

- Supongamos que existe un $v \in \{1, \dots, l\}$ tal que $k'_v \neq k''_{v^*}, \forall v^* \in \{1, \dots, r\}$, tenemos que $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (\mathbf{p} + \sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_v} \boldsymbol{\xi}_{k'_v} - \mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T (\sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_v} \boldsymbol{\xi}_{k'_v})$, por (4.4), se sigue $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (-j_{k'_v} \boldsymbol{\xi}_{k'_v}) = -j_{k'_v} a_{k'_v} \alpha_{k'_v}^T \boldsymbol{\xi}_{k'_v}$, por (4.11) tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = -j_{k'_v} \delta \leq 0, \forall j_{k'_v} \in \{0, \dots, m_{k'_v} - 1\}$, por (3.5) se sigue que $g(\mathbf{x}) = 0$.
- Supongamos $k'_v = k''_{v^*}$ y $q_{k''_{v^*}} = j_{k'_v} + 1, \forall v \in \{1, \dots, l\}$, es decir, que $g(\mathbf{x})$ es un elemento en la diagonal de $A_{D_{\text{gen}}}$, tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_v} \alpha_{k'_v}^T (\mathbf{p} + \sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_v} \boldsymbol{\xi}_{k'_v} - \mathbf{p}) = a_{k'_v} \alpha_{k'_v}^T (\sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_v} \boldsymbol{\xi}_{k'_v}) = a_{k'_v} \alpha_{k'_v}^T ((j_{k'_v} + 1) \boldsymbol{\xi}_{k'_v} - j_{k'_v} \boldsymbol{\xi}_{k'_v}) = a_{k'_v} \alpha_{k'_v}^T \boldsymbol{\xi}_{k'_v}$, por (4.11) tenemos $\pi_{a_{k'_v} \alpha_{k'_v}}^{(j_{k'_v}) \cdot \mathbf{p}}(\mathbf{x}) = \delta$, por (3.5) se sigue que $g(\mathbf{x}) = \delta$.
- Supongamos $k'_v = k''_{v^*}, \forall v \in \{1, \dots, l\}$ y existe un $v^* \in \{1, \dots, l\}$ tal que $q_{k''_{v^*}} \leq j_{k'_{v^*}}$, es decir, que $g(\mathbf{x})$ es un elemento sobre la diagonal principal de $A_{D_{\text{gen}}}$, tenemos $\pi_{a_{k'_{v^*}} \alpha_{k'_{v^*}}}^{(j_{k'_{v^*}}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (\mathbf{p} + \sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}} - \mathbf{p}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (\sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}})$, por (4.4), se sigue $\pi_{a_{k'_{v^*}} \alpha_{k'_{v^*}}}^{(j_{k'_{v^*}}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (q_{k''_{v^*}} \boldsymbol{\xi}_{k''_{v^*}} - j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}}) = (q_{k''_{v^*}} - j_{k'_{v^*}}) a_{k'_{v^*}} \alpha_{k'_{v^*}}^T \boldsymbol{\xi}_{k'_{v^*}}$, por (4.11) tenemos $\pi_{a_{k'_{v^*}} \alpha_{k'_{v^*}}}^{(j_{k'_{v^*}}) \cdot \mathbf{p}}(\mathbf{x}) = (q_{k''_{v^*}} - j_{k'_{v^*}}) \delta \leq 0$, por (3.5) se sigue que $g(\mathbf{x}) = 0$.

Sea $l > r$, existe un $v^* \in \{1, \dots, l\}$ tal que $k'_{v^*} \neq k''_{v^*}, \forall v \in \{1, \dots, r\}$, tenemos $\pi_{a_{k'_{v^*}} \alpha_{k'_{v^*}}}^{(j_{k'_{v^*}}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (\mathbf{p} + \sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}} - \mathbf{p}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (\sum_{i=1}^r q_{k''_i} \boldsymbol{\xi}_{k''_i} - j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}})$, por (4.4), se sigue $\pi_{a_{k'_{v^*}} \alpha_{k'_{v^*}}}^{(j_{k'_{v^*}}) \cdot \mathbf{p}}(\mathbf{x}) = a_{k'_{v^*}} \alpha_{k'_{v^*}}^T (-j_{k'_{v^*}} \boldsymbol{\xi}_{k'_{v^*}}) = -j_{k'_{v^*}} a_{k'_{v^*}} \alpha_{k'_{v^*}}^T \boldsymbol{\xi}_{k'_{v^*}}$, por

(4.11) tenemos $\pi_{a_{k'v^*}}^{(j_{k'v^*})_{\mathbf{p}}}(\mathbf{x}) = -j_{k'v^*} \delta \leq 0, \forall j_{k'v^*} \in \{0, \dots, m_{k'v^*} - 1\}$, por (3.5) se sigue que $g(\mathbf{x}) = 0$.

Por lo que que la matriz $A_{D_{\text{gen}}}$ es diagonal inferior y sus elementos en la diagonal principal son diferentes de 0, por tanto es no singular. ■

Ejemplo 32. Continuando con el Ejemplo 31, para el cual $V_{D_{\text{eje}}} =$

$$\{(-3 \ -1), (-2 \ 2), (-1 \ -3), (-2 \ 1), (-1 \ 3), (0 \ 5), (-1 \ 0), (0 \ 2), (1 \ 4), (0 \ -1), (1 \ 1), (2 \ 3)\}$$

y $\Lambda_{D_{\text{eje}}} =$

$$\begin{bmatrix} \gamma^{(0)}(1) & \gamma^{(1)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}\right) & \gamma^{(1)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}\right) & \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right) & \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right) & \gamma^{(1)}\left(\pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right) \\ \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(0),\mathbf{p}}, \pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(0),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right) \\ \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(1),\mathbf{p}}\right) & \gamma^{(2)}\left(\pi_{a_1\alpha_1}^{(1),\mathbf{p}}, \pi_{a_2\alpha_2}^{(2),\mathbf{p}}\right) & & & & \end{bmatrix}^T,$$

tenemos que

$$A_{D_{\text{eje}}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 6 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 6 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 9 & 6 & 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 3 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 6 & 3 & 0 & 3 & 3 & 0 & 0 & 0 & 0 \\ 1 & 3 & 0 & 9 & 6 & 3 & 3 & 3 & 3 & 0 & 0 & 0 \\ 1 & 6 & 3 & 3 & 0 & 0 & 3 & 0 & 0 & 3 & 0 & 0 \\ 1 & 6 & 3 & 6 & 3 & 0 & 6 & 3 & 0 & 3 & 3 & 0 \\ 1 & 6 & 3 & 9 & 6 & 3 & 6 & 6 & 3 & 3 & 3 & 3 \end{bmatrix}.$$

Sea $\mathbf{c}_{D_{\text{gen}}}^{(i)} \in \mathbb{R}^{q_i}$, con $q_i = \left(\sum_{k_1=1}^{n-r+1} m_{k_1} \sum_{k_2=k_1+1}^{n-r+2} m_{k_2} \dots \sum_{k_r=k_{r-1}+1}^n m_{k_r}\right)$, $i \in \{0, \dots, n\}$ y $q_0 = 1$, un vector de coeficientes $\mathbf{c}_{D_{\text{gen}}} \in \mathbb{R}^{\sum_{i=0}^n q_i}$, $\mathbf{c}_{D_{\text{gen}}} = \left[\mathbf{c}_{D_{\text{gen}}}^{(0)} \dots \mathbf{c}_{D_{\text{gen}}}^{(n)}\right]$. Así, funciones de la forma

$$g(\mathbf{x}) = \mathbf{c}_{D_{\text{gen}}}^T \Lambda_{D_{\text{gen}}}, \quad (4.26)$$

$g : D_{\text{gen}} \rightarrow \mathbb{R}$ son funciones HL-CPWL definidas sobre la partición $H_{D_{\text{gen}}}$. Además, esta es una representación general para las funciones PWL en $PWL_{H_{D_{\text{gen}}}}[D_{\text{gen}}]$ como se establece en el siguiente teorema.

Teorema 7. Las funciones en $\Lambda_{D_{\text{gen}}}$ generan el espacio $PWL_{H_{D_{\text{gen}}}}[D_{\text{gen}}]$.

Prueba. Dada una función $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$, se determinará un vector de coeficientes $\mathbf{c}_{D_{\text{gen}}}$ tal que f pueda ser expresada como una combinación lineal de funciones en $\mathbf{\Lambda}_{D_{\text{gen}}}$ como sigue:

$$f(\mathbf{x}) = \mathbf{c}_{D_{\text{gen}}}^T \mathbf{\Lambda}_{D_{\text{gen}}}, \quad \forall \mathbf{x} \in D_{\text{gen}}, \quad (4.27)$$

como se mencionó anteriormente, toda función PWL está únicamente determinada por su valor en los vértices de su partición y por el Lema 20, las funciones en $\mathbf{\Lambda}_{D_{\text{gen}}}$ están definidas en la partición $H_{D_{\text{gen}}}$. Así, probar esto es equivalente a probar (4.27) para todo $\mathbf{x} \in V_{D_{\text{gen}}}$. Sea $\mathbf{b}_{D_{\text{gen}}}$ un vector para el cual su i -ésima componente es igual a la evaluación de f en el i -ésimo vértice en $V_{D_{\text{gen}}}$ y sea $A_{D_{\text{gen}}}$ la matriz definida en (4.23), se sigue del Lema 22 que existe la matriz inversa $A_{D_{\text{gen}}}^{-1}$, en consecuencia, existe un único vector $\mathbf{c}_{D_{\text{gen}}}$ tal que

$$\mathbf{b}_{D_{\text{gen}}}^T = \mathbf{c}_{D_{\text{gen}}}^T A_{D_{\text{gen}}}^T. \quad (4.28)$$

Se sigue que (4.27) se verifica para esta elección de $\mathbf{c}_{D_{\text{gen}}}$. Así, $f(\mathbf{x}) \equiv \mathbf{c}_{D_{\text{gen}}}^T \mathbf{\Lambda}_{D_{\text{gen}}}$, $\forall \mathbf{x} \in D_{\text{gen}}$. ■

Ejemplo 33. Continuando con el Ejemplo 32, suponga que f es una función en el espacio vectorial lineal $PWL_{H_{D_{\text{eje}}}} [D_{\text{eje}}]$ la cual toma los valores dados en la Tabla 4.1.

\mathbf{x}	$f(\mathbf{x})$	\mathbf{x}	$f(\mathbf{x})$
$(-3 \ -1)^T$	1	$(-1 \ 0)^T$	0
$(-2 \ -2)^T$	4	$(0 \ 2)^T$	-1
$(-1 \ -3)^T$	0	$(1 \ 4)^T$	5
$(-2 \ 1)^T$	-4	$(0 \ -1)^T$	-5
$(-1 \ 3)^T$	5	$(1 \ 1)^T$	2
$(0 \ 5)^T$	-2	$(2 \ 3)^T$	-1

Tabla 4.1: Valores de f en los vértices $V_{D_{\text{eje}}}$ del Ejemplo 33.

Sea $\mathbf{b}_{D_{\text{gen}}}$ el vector cuya i -ésima componente es igual a la evaluación de f en el i -ésimo vértice de $V_{D_{\text{eje}}}$,

$$\mathbf{b}_{D_{\text{gen}}} = (1 \ 4 \ 0 \ -4 \ 5 \ -2 \ 0 \ -1 \ 5 \ -5 \ 2 \ -1)^T.$$

De (4.28), tenemos que

$$\mathbf{c}_{D_{\text{eje}}} = A_{D_{\text{eje}}}^{-1} \mathbf{b}_{D_{\text{gen}}} = \left(1 \ 1 \ -\frac{7}{3} \ -\frac{5}{3} \ \frac{14}{3} \ -\frac{16}{3} \ \frac{1}{3} \ -\frac{10}{3} \ \frac{13}{3} \ -\frac{1}{3} \ \frac{7}{3} \ \frac{1}{3} \right)^T.$$

Así, la representación HL-CPWL de f es $f(\mathbf{x}) = \mathbf{c}_{D_{\text{eje}}}^T \mathbf{\Lambda}_{D_{\text{eje}}}(\mathbf{x})$. Una gráfica de la función f se muestra en la Figura 4.3.

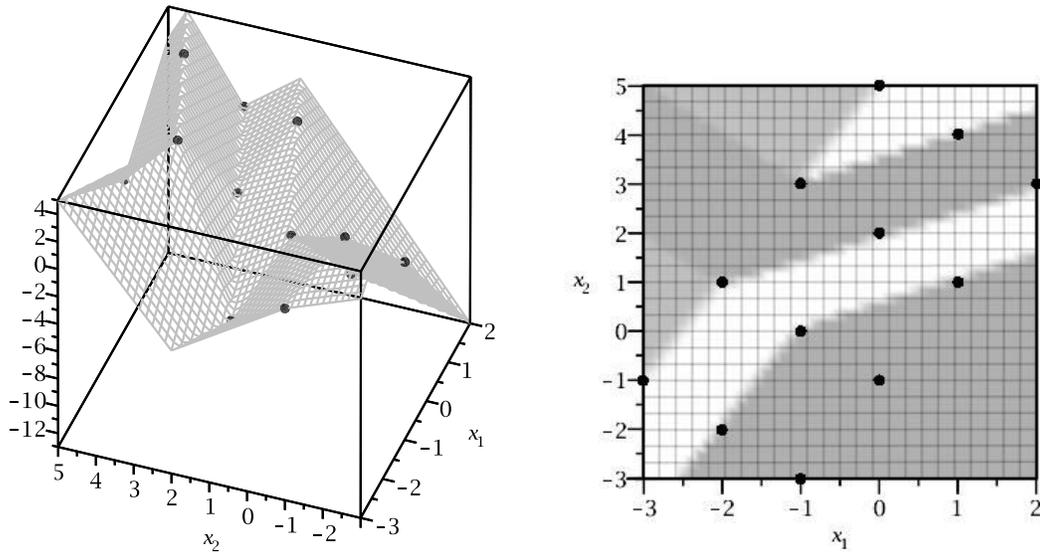


Figura 4.3: Función f del Ejemplo 33 (a) y vista sobre el plano x_1x_2 .

4.1.3 Propiedades computacionales.

En esta subsección se hace un análisis análogo al realizado en la Sección 3.4. Las propiedades de esta propuesta son similares a las que presentan las particiones simpliciales ortogonales pues comparten características similares, al ser una generalización de las mismas. Veamos que, del Lema 22, la matriz $A_{D_{\text{gen}}}$ es triangular inferior, por lo que los coeficientes $\mathbf{c}_{D_{\text{gen}}}$ pueden ser calculados sin la necesidad de invertir la matriz $A_{D_{\text{gen}}}$. Sean $a_{i,j}$ el i, j -ésimo elemento de la matriz $A_{D_{\text{gen}}}$, c_k el k -ésimo elemento del vector $\mathbf{c}_{D_{\text{gen}}}$ y b_l el l -ésimo elemento del vector $\mathbf{b}_{D_{\text{gen}}}$. Entonces, cada coeficiente puede ser obtenido de forma recursiva

$$\begin{aligned} c_1 &= \frac{1}{a_{1,1}} b_1, \\ &\vdots \\ c_i &= \frac{1}{a_{i,i}} \left(b_i - \sum_{p=1}^{i-1} a_{i,p} c_p \right), \end{aligned} \tag{4.29}$$

donde $a_{1,1} = 1$ y $a_{i,i} = \delta$, $i \in \{2, \dots, k\}$.

La representación $HL-CPWL$ presentada en este capítulo posee el número mínimo y exacto de parámetros requeridos para representar una función PWL $f \in PWL_{HD_{\text{gen}}}[D_{\text{gen}}]$. Además, puesto que las particiones simpliciales vistas pueden verse como funciones en el hipercubo $[0, 1]$ bajo un isomorfismo, la comparación de parámetros requeridos en esta representación con respecto a las mostradas en capítulos anteriores puede hacerse de la siguiente manera. Suponga (4.1) con $m_i = k$, $i \in \{1, \dots, n\}$. D_{gen} tiene k^n hipercubos bajo un isomorfismo, cada uno tiene $n!$ símplexes y el número de vértices es $(k+1)^n$, De aquí que:

- Representación estándar PWL: tiene un número total de parámetros $k^n (n+1)!$

- Representación CPWL: con un número total de parámetros $kn + \binom{n}{2}(2k - 1) + n + 1$. Además que esta solo puede representar funciones que cumplan con la propiedad de variación consistente.
- Representación HL-CPWL ortogonal: la cual tiene un número total de parámetros $(k + 1)^n$. Además que esta propuesta solo puede representar particiones que forman mallas ortogonales igualmente distribuidas.
- Representación HL-CPWL simplicial propuesta en este trabajo: por el Lema 21 y el Teorema 7, el número total de parámetros es $(k + 1)^n$.

Por esto, cabe señalar que: el método propuesto requiere una menor cantidad de parámetros para describir el comportamiento de un función lineal a tramos en comparación con la representación estándar; a pesar de tener un mayor número de parámetros que la representación CPWL, este no requiere verificar la propiedad de variación consistente; utiliza el mismo número de parámetros que la representación HL-CPWL, sin embargo, este puede representar funciones PWL sobre particiones más generales, más aún, si se substituyen los vectores ξ por vectores canónicos, entonces este esquema de construcción es el dado en [22], por lo que se considera a este último un caso particular del propuesto en este capítulo.

Comparando con los métodos look-up table, ambos métodos almacenan el número de parámetros dados por la función f en los vértices de su partición. Sin embargo, ahora es posible establecer esquemas que no sean rectangulares en lo que se refiere a la posición de los datos almacenados.

Por lo anterior, vemos que esta propuesta conserva las características computacionales del esquema de partición simplicial propuesto en [22], pero posee una mayor versatilidad respecto a las particiones simpliciales que puede representar.

4.1.4 Algoritmo para la construcción de una representación HL-CPWL de funciones $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$.

El algoritmo para la construcción una representación HL-CPWL de la forma (4.27) de una función $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$ consta de tres etapas principales las cuales se representan en la Figura 4.4.

Para este algoritmo, los datos de entrada son los siguientes: Dimensión del espacio n ; vectores ξ_1, \dots, ξ_2 ; vector de desplazamiento \mathbf{p} y los enteros m_1, \dots, m_n , mencionados en (4.1) y los valores de f en los vértices de su partición. La salida del algoritmo es una representación HL-CPWL de la función $f \in PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$ de la forma (4.27).

Obtención del arreglo de índices.

Este bloque del algoritmo tiene a su entrada la dimensión n del espacio de salida de las función f , y en su salida entrega un arreglo al cual llamaremos Ind de tuplas A ordenadas,

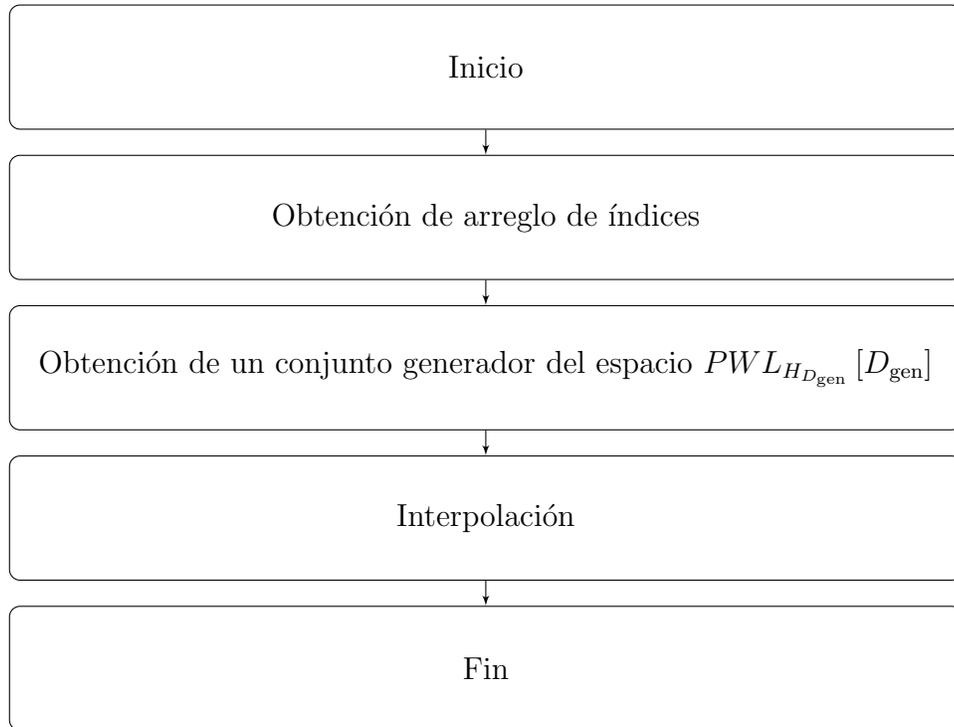


Figura 4.4: Algoritmo para la construcción de una representación HL-CPWL de funciones $f \in PWL_{HD_{gen}} [D_{gen}]$.

ENTRADA: Dimensión del espacio n .

SALIDA: Arreglo de arreglos de tuplas A ordenadas.

```

Inicio;
Ind := arreglo vacío;
for  $i = 1, \dots, n$  do
    |  $Ind = Ind$ , arreglos A ordenados de tuplas  $r_j \in \{1, \dots, n\}$ ,  $j \in \{1, \dots, i\}$ ;
end
return Ind;
Fin;
    
```

Tabla 4.2: Obtención del arreglo de índices.

es decir, que cada elemento $i \in Ind$ será una tupla de $k \in \{1, \dots, n\} \uparrow$ elementos y el orden de estas tuplas está dado por la Definición 14, como se muestra en la Tabla 4.2.

Ejemplo 34. Suponga que a la entrada de la obtención del arreglo de tuplas tiene $n = 3$, entonces la salida será un arreglo:

$$Ind = [1], [2], [3], [1, 2], [1, 3], [2, 3], [1, 2, 3].$$

En este ejemplo vemos que Ind se forma por siete tuplas, las primeras tres se forman por un elemento y son A ordenados; las siguientes tres tuplas se forman de dos elementos y son A ordenados y la última tupla se forma de tres elementos. Más adelante veremos que estas tuplas de índices están relacionadas con el nivel de anidación de las funciones (4.20) y con la clase de los vértices (4.19).

Para simplificar, diremos que Ind_i , es la i -ésima tupla en Ind y $Ind_{i,j}$ es el j -ésimo elemento en la tupla Ind_i .

Obtención de un conjunto generador del espacio $PWL_{HD_{\text{gen}}} [D_{\text{gen}}]$.

Este bloque del algoritmo tiene a su entrada la dimensión n del espacio de salida de las función PWL f ; los vectores ξ_1, \dots, ξ_n que dan orientación a la malla que particiona el espacio de salida de f , es decir, los vectores Ξ de (4.1); el vector de desplazamiento \mathbf{p} ; los enteros m_1, \dots, m_n que dan el número de divisiones en la dirección de los vectores Ξ de la malla; y el arreglo de índices Ind . A su salida entrega un arreglo representaciones HL-CPWL de funciones PWL. Este bloque se conduce como se muestra en la Tabla 4.3.

Note que la instrucción $dim(Ind)$ obtiene el número de elementos (arreglos) en Ind . Las variables que se declaran son:

- α , es un arreglo de dimensión n , es decir, que el número de elementos que tiene es igual a n , de vectores en el espacio de salida de f , este almacena los vectores que cumplen con la relación (4.4), así, α_i , $i \in \{1, \dots, n\}$, se refiere al i -ésimo vector en α . Para calcular α_i , recordemos que este es el vector normal al hiperplano generado por los vectores ξ_j , $j \in \{1, \dots, n\}$, $j \neq i$. Llamemos $\xi_{j,h}$, $h \in \{1, \dots, n\}$, al h -ésimo elemento de ξ_j , la función que caracteriza a este hiperplano puede ser obtenida, de acuerdo con [26], calculando el determinante

$$\begin{vmatrix} x_1 & \dots & x_n \\ \xi_{1,1} & \dots & \xi_{1,n} \\ \vdots & & \vdots \\ \xi_{i-1,1} & \dots & \xi_{i-1,n} \\ \xi_{i+1,1} & \dots & \xi_{i+1,n} \\ \vdots & & \vdots \\ \xi_{n,1} & \dots & \xi_{n,n} \end{vmatrix},$$

así, el l -ésimo elemento del vector α_i , $l \in \{1, \dots, n\}$, denotado por $\alpha_{i,l}$, se obtiene calculando el determinante

ENTRADA: Dimensión del espacio n , vectores ξ_1, \dots, ξ_n , vector de desplazamiento \mathbf{p} , enteros m_1, \dots, m_n , arreglo Ind .

SALIDA: Arreglo de funciones HL-CPWL.

```

Inicio;
 $\alpha$  := arreglo de vectores;
 $\omega$  := arreglo de vectores;
 $a$  := arreglo de coeficientes;
 $S$  := arreglo de arreglos de funciones;
Calcular el vector  $\alpha_1$ ;
 $\Lambda$  := arreglo vacío de funciones;
 $a_1 = 1$ ;
for  $j = 2, \dots, n$  do
    | Calcular los vectores  $\alpha_j$ ;
    | Calcular los vectores  $\omega_{j-1}$ ;
    | Calcular los coeficientes  $a_j$ ;
end
for  $i = 1, \dots, n$  do
    | for  $q_i = 1, \dots, m_i$  do
    | | Calcular los hiperplanos  $S_{i,q_i} = \pi_{a_i \alpha_i}^{(q_i-1), \mathbf{p}}$ ;
    | end
end
 $\Lambda = \Lambda, 1$ ;
for  $i = 1, \dots, \dim(Ind)$  do
    |  $\Lambda = \Lambda$ , arreglo de funciones
    |  $\gamma^{\dim(Ind_i)} \left( S_{Ind_{i,1}, q_{Ind_{i,1}}}, \dots, S_{Ind_{i, \dim(Ind_i)}, q_{Ind_{i, \dim(Ind_i)}}} \right)$  B ordenado para las
    | tuplas  $\{ q_{Ind_{i,1}}, \dots, q_{Ind_{i, \dim(Ind_i)}} \}$ ;
end
return  $\Lambda$ ;
Fin;

```

Tabla 4.3: Obtención de un conjunto generador del espacio $PWL_{H_{D_{\text{gen}}}} [D_{\text{gen}}]$.

$$\alpha_{i,l} = \begin{vmatrix} \xi_{1,1} & \cdots & \xi_{1,l-1} & \xi_{1,l+1} & \cdots & \xi_{1,n} \\ \vdots & & \vdots & & & \\ \xi_{i-1,1} & \cdots & \xi_{i-1,l-1} & \xi_{i-1,l+1} & \cdots & \xi_{i-1,n} \\ \xi_{i+1,1} & \cdots & \xi_{i+1,l-1} & \xi_{i+1,l+1} & \cdots & \xi_{i+1,n} \\ \vdots & & \vdots & & & \\ \xi_{n,1} & \cdots & \xi_{n,l-1} & \xi_{n,l+1} & \cdots & \xi_{n,n} \end{vmatrix} \left((-1)^{(l+1)} \right).$$

- ω , es un arreglo de dimensión $n - 1$ de vectores en el espacio de salida de f , este almacena los vectores que cumplen con la relación (4.6), así, ω_{i-1} , $i \in \{2, \dots, n\}$, se refiere al $(i - 1)$ -ésimo vector en ω . Sea $\omega_{i,j}$, $j \in \{1, \dots, n\}$, el j -ésimo elemento en ω_i , de forma similar a los elementos de los vectores en α , tenemos que este puede ser obtenido por el determinante

$$\omega_{i-1,j} = \begin{vmatrix} \xi_{1,1} + \xi_{i,1} & \cdots & \xi_{1,j-1} + \xi_{i,j-1} & \xi_{1,j+1} + \xi_{i,j+1} & \cdots & \xi_{1,n} + \xi_{i,n} \\ \vdots & & \vdots & & & \\ \xi_{i-1,1} & \cdots & \xi_{i-1,j-1} & \xi_{i-1,j+1} & \cdots & \xi_{i-1,n} \\ \xi_{i+1,1} & \cdots & \xi_{i+1,j-1} & \xi_{i+1,j+1} & \cdots & \xi_{i+1,n} \\ \vdots & & \vdots & & & \\ \xi_{n,1} & \cdots & \xi_{n,j-1} & \xi_{n,j+1} & \cdots & \xi_{n,n} \end{vmatrix} \left((-1)^{(j+1)} \right).$$

- a , es un arreglo de dimensión n que almacena los coeficientes que cumplen con la relación (4.11), así, a_i , $i \in \{1, \dots, n\}$, se refiere al i -ésimo elemento en a . Para calcular a_i , $i \in \{1, \dots, n\}$, primero obtenemos los coeficientes $c_{k,1}$ y $c_{k,k+1}$, $k \in \{1, \dots, n - 1\}$, tales que

$$\omega_k = c_{k,1}\alpha_1 + c_{k,k+1}\alpha_{k+1},$$

los cuales, por Lema 5, existen. Después, haciendo $a_1 = 1$, tenemos que

$$\rho_k = c_{k,1}, \quad a_{k+1} = -\frac{c_{k,k+1}}{\rho_k}.$$

- S , el cual es un arreglo de dimensión n de arreglos de dimensión m_i , $i \in \{1, \dots, n\}$, de funciones afines. Donde S_i indica el i -ésimo arreglo de S y S_{q_i} indica la q_i -ésima función en S_i .
- Λ es un arreglo de elementos, B ordenados, de un arreglo de funciones $\gamma^{(\dim(Ind_i))}$ $\left(S_{Ind_{i,1}, q_{Ind_{i,1}}}, \dots, S_{Ind_{i, \dim(Ind_i)}, q_{Ind_{i, \dim(Ind_i)}}} \right)$ para cada elemento de Ind fijo.

Ejemplo 35. Continuando con el Ejemplo 34, supongamos $i = 5$ en el algoritmo de la Tabla 4.3, es decir, fijamos el arreglo de índices $Ind_5 = [1, 3]$, por lo que $\dim(Ind_5) = 2$. Ahora supongamos $m_1 = 2$ y $m_3 = 3$. Entonces

$$\Lambda = \Lambda, \gamma^{(2)}(S_{1,1}, S_{2,1}), \gamma^{(2)}(S_{1,1}, S_{2,2}), \gamma^{(2)}(S_{1,1}, S_{2,3}), \gamma^{(2)}(S_{1,2}, S_{2,1}), \gamma^{(2)}(S_{1,2}, S_{2,2}), \gamma^{(2)}(S_{1,2}, S_{2,3}),$$

es decir, que Λ almacena los valores de su estado anterior y agrega funciones B ordenadas en cada iteración de i .

Interpolación.

Este bloque del algoritmo tiene a su entrada a n ; los vectores ξ_1, \dots, ξ_n ; al vector \mathbf{p} ; los enteros m_1, \dots, m_n ; el arreglo de índices Ind ; y el arreglo de funciones Λ . A su salida entrega una representación HL-CPWL de la función f . Este bloque se conduce como se muestra en la Tabla 4.4.

Las variables que se declaran son:

- V almacena los vértices de la partición de la función f , por lo que es un arreglo de elementos, B ordenados, de un arreglo de vectores $\mathbf{p} + \sum_{p=1}^{dim(Ind_i)} q_{Ind_i,p} \xi_{Ind_i,p}$ para cada elemento de Ind fijo. Al igual que en el Ejemplo 35, V almacena los valores de su estado anterior y agrega vectores B ordenados en cada iteración de i .
- A , la cual es una matriz, donde cada fila A_i se forma de evaluar el arreglo Λ en el vector V_i .
- \mathbf{b} , es un arreglo el cual almacena los valores de f al evaluarse en los elementos de V .
- c es un arreglo de coeficientes, donde cada k -ésimo elemento c_k se obtiene usando (4.29), con $a_{i,j}$ siendo el i, j -ésimo elemento de A y b_l el l -ésimo elemento de \mathbf{b} .
- F es la función en la salida del algoritmo.

ENTRADA: Dimensión del espacio n , vectores ξ_1, \dots, ξ_n , vector de desplazamiento \mathbf{p} , enteros m_1, \dots, m_n , arreglo Ind , arreglo de funciones Λ .

SALIDA: Representación HL-CPWL de f .

```

Inicio;
V := arreglo vacío de vectores;
A := matriz;
b := arreglo vacío de valores de  $f$ ;
c := arreglo de coeficientes;
F := función;
V = V, p;
for  $i = 1, \dots, \dim(Ind)$  do
    | V = V, arreglo de vectores  $\mathbf{p} + \sum_{p=1}^{\dim(Ind_i)} q_{Ind_i,p} \xi_{Ind_i,p}$  B ordenado para las
    | tuplas  $\{q_{Ind_i,1}, \dots, q_{Ind_i,\dim(Ind_i)}\}$ ;
end
for  $i = 1, \dots, \dim(V)$  do
    |  $A_i = \Lambda$  evaluada en  $V_i$ ;
end
for  $i = 1, \dots, \dim(V)$  do
    |  $\mathbf{b}_i =$  Valor de  $f$  en  $V_i$ ;
end
for  $i = 1, \dots, \dim(V)$  do
    | Calcular  $c_i$ ;
end
F =  $\sum_{i=1}^{\dim(V)} c_i \Lambda_i$ ;
return F;
Fin;

```

Tabla 4.4: Interpolación de la función.

4.2 Un método para extender la representación CPWL de una a dos dimensiones.

En esta sección se presenta una extensión del procedimiento para calcular los coeficientes usados en una representación CPWL de funciones PWL $f : \mathbb{R} \rightarrow \mathbb{R}$, de manera que este pueda ser aplicado para la construcción de representaciones con un dominio en \mathbb{R}^2 . Esta propuesta se basa en que una representación de una función PWL con un dominio en \mathbb{R}^2 puede ser aproximada al reconstruirla por una colección de proyecciones paralelas de funciones PWL con un dominio en \mathbb{R} . Los resultados de esta propuesta pueden ser consultados en [27], en donde se verifica su desempeño en tiempo computacional requerido para la construcción de una representación, contrastado con el de la representación simplicial HL-CPWL del Capítulo 3. Esta comparación revela que el procedimiento propuesto requiere de menor tiempo de cómputo, así como de menor espacio en memoria.

4.2.1 Proyección de representaciones CPWL.

Partiendo de la representación CPWL de funciones PWL dada en (2.23) y como se vio en el Capítulo 2, se sigue que toda función continua PWL $g : \mathbb{R} \rightarrow \mathbb{R}$ tiene una representación de la forma

$$g(\mathbf{x}) = a + bx + \sum_{i=1}^{\sigma} c_i |x - \beta_i|, \quad (4.30)$$

donde σ es el número de puntos de quiebre β_i , en el dominio de la función, los cuales particionan al dominio en $\sigma + 1$ intervalos consecutivos. Además $a, b, c_i \in \mathbb{R}$ pueden ser calculados por (2.37), (2.35) y (2.36), respectivamente, de la siguiente forma:

$$\begin{aligned} a &= g(0) - \sum_{i=1}^{\sigma} c_i |\beta_i|, \\ b &= \frac{J_1 + J_{\sigma+1}}{2}, \\ c_i &= \frac{J_{i+1} - J_i}{2}, \quad i \in \{1, \dots, \sigma\}, \end{aligned} \quad (4.31)$$

donde J_i denota la pendiente de g en el i -ésimo intervalo en el dominio de g .

Considere ahora una nube de puntos

$$\{(x_i \ z_j \ y_{i,j}) : i = 0, \dots, 1 - p, j = 0, \dots, 1 - q, y_{i,j} = f((x_i \ z_j))\},$$

en \mathbb{R}^3 . Para cada $i = 1, \dots, 1 - p$ consideramos los puntos $(z_j \ y_{i,j})$, $j = 0, \dots, 1$ en \mathbb{R}^2 y los interpolamos por una función PWL en la variable z , la cual denominaremos $f((i \ z))$, usando (4.30) y (4.31). En la Figura 4.5 se han representado estas funciones en el espacio tridimensional.

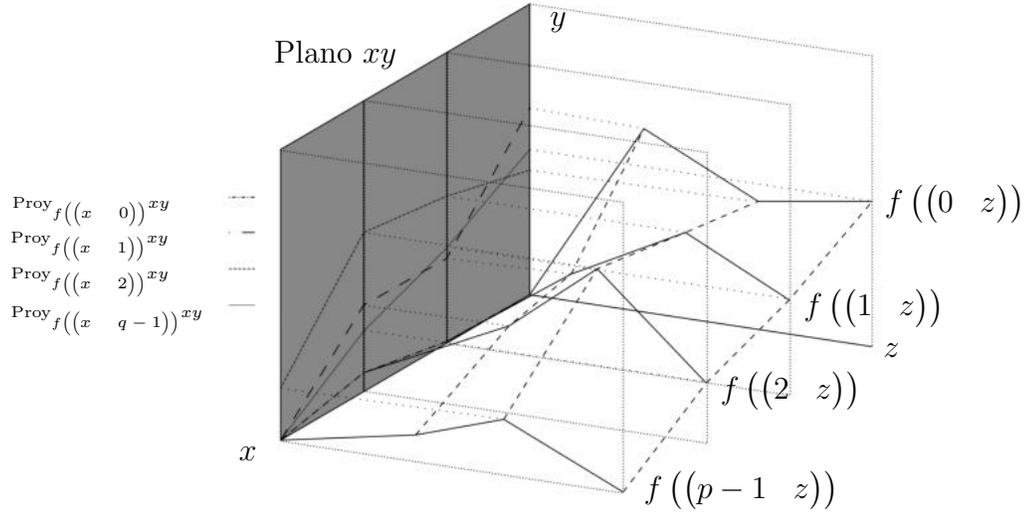


Figura 4.5: Proyecciones $\text{Proy}_f((x, i))_{xy}$ en el plano xy (Plano de color gris). Funciones $f((x, i))$ (línea espaciada en el espacio xyz). Funciones $f((x, i))$ (línea sólida en el espacio xyz).

Las funciones $f((i, z))$ se usarán ahora para definir una función de dos variables $f((x, z))$ que imite la forma canónica (4.30) de una representación PWL en una variable. Los lugares de las constantes a, b, c_i que figuran en esa expresión los ocuparán funciones PWL $A(z), B(z)$ y $C_i(z)$, en la variable z . Así,

$$f((x, z)) = A(z) + B(z)x + \sum_{i=1}^{p-1} C_i(z)|x - i|. \quad (4.32)$$

Donde, por analogía con (4.31),

$$\begin{aligned} B(z) &= \frac{\iota_1(z) + \iota_{p-1}(z)}{2}, \\ C_i(z) &= \frac{\iota_{i+1}(z) - \iota_i(z)}{2}, \quad i \in \{1, \dots, \sigma_p\}, \end{aligned} \quad (4.33)$$

y

$$\iota_i(z) = f((i, z)) - f((i-1, z)), \quad i \in \{1, \dots, p-1\}. \quad (4.34)$$

Por último, despejando $A(z)$ de (4.32) y haciendo $x = 0$, tenemos

$$A(z) = f((0, z)) - \sum_{i=1}^{p-1} C_i(z)|x_i| = f((0, z)) - \sum_{i=1}^{p-1} C_i(z)|f((i, 0))|, \quad (4.35)$$

con lo cual, sustituyendo (4.33) y (4.35) en (4.32) queda determinada $f((x, z))$.

Ejemplo 36. Considere a $D = \{(0, 0), (0, 1), (0, 2), (0, 3), (1, 0), (1, 1), (1, 2), (1, 3), (2, 0), (2, 1), (2, 2), (2, 3), (3, 0), (3, 1), (3, 2), (3, 3)\}$. Indexemos a estos puntos por su orden de aparición en la lista precedente $(x_i, z_i), i \in \{1, \dots, 16\}$. Estos puntos corresponden al dominio de una función $f((x, z))$ y una partición con 4 puntos de quiebre, equidistantes en cada eje. Consideremos a cada par ordenado (x_i, z_i) le corresponde a un $f((x_i, z_i)) = y_i$, como se muestra en la Tabla 4.5.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x	0	0	0	0	1	1	1	1	2	2	2	2	3	3	3	3
z	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
y_i	2	2	2	0	3	1	1	1	2	2	0	0	3	1	1	0

Tabla 4.5: Puntos en D y su valor correspondiente $y_i = f((x_i, z_i))$.

A partir de estos datos, las funciones $f((0, z)), f((1, z)), f((2, z))$ y $f((3, z))$, son obtenidas por (4.30) y (4.31) de la siguiente manera

$$\begin{aligned}
 f((0, z)) &= 4 - z - |z - 2|, \\
 f((1, z)) &= 2 - z + |z - 1|, \\
 f((2, z)) &= 1 - |z - 1| + |z - 2|, \\
 f((3, z)) &= 3 - \frac{3}{2}z + |z - 1| - \frac{1}{2}|z - 2|,
 \end{aligned} \tag{4.36}$$

donde $f((0, z))$ se ajusta a los puntos $\{(0, y_1), (1, y_2), (2, y_3), (3, y_4)\}$, $f((1, z))$ se ajusta a los puntos $\{(0, y_5), (1, y_6), (2, y_7), (3, y_8)\}$, $f((2, z))$ se ajusta a los puntos $\{(0, y_9), (1, y_{10}), (2, y_{11}), (3, y_{12})\}$ y $f((3, z))$ se ajusta a los puntos $\{(0, y_{13}), (1, y_{14}), (2, y_{15}), (3, y_{16})\}$.

Basados en el esquema de construcción propuesto, las funciones (4.36) pueden ser utilizadas para determinar por (4.34) a las funciones de las pendientes variables: $\iota_1(z)$, $\iota_2(z)$ y $\iota_3(z)$. Una vez que las funciones de las pendientes variables son obtenidas, el siguiente paso es determinar los parámetros $B(z)$, $C_i(z)$ y $A(z)$ usando (4.33) y (4.35). Finalmente, la función $f((x, z))$ que se ajusta a los puntos de la Tabla 4.5 es obtenida por (4.32), dando como resultado

$$\begin{aligned}
 f((x, z)) &= \left(\frac{1}{2} + z\right) - \left(\frac{5}{2} - \frac{3}{2}x\right)|z - 1| + \left(\frac{3}{2} - \frac{1}{4}x\right)|z - 2| - \frac{3}{4}xz + \left(\frac{1}{2} + \frac{1}{2}z\right)|x - 1| \\
 &+ \left(\frac{3}{2} - \frac{5}{4}z\right)|x - 2| - \frac{3}{2}|xz - x - z + 1| + 2|xz - x - 2z + 2| - \frac{5}{4}|xz - 2x - 2z + 4|.
 \end{aligned}$$

Capítulo 5

Sobre el ajuste a una nube de puntos.

En este capítulo buscamos evaluar el desempeño de nuestro método al usarlo para ajustar una nube de puntos en el espacio tridimensional por medio de funciones PWL siguiendo los pasos descritos en el Sección 4.1 y compararlo con el de funciones PWL definidas en mallas ortogonales como las vistas en la Sección 3.3.

En la sección 5.1 explicamos los patrones de muestreo usados para obtener las nubes de puntos a partir de la evaluación de algunas funciones definidas en \mathbb{R}^2 . En la Sección 5.2 se definen distintas particiones, tanto ortogonales a los ejes x_1 y x_2 como no ortogonales a estos, definidas sobre \mathbb{R}^2 . Estas particiones determinan las bases para los espacios de funciones PWL que nos permitirán encontrar las expresiones correspondientes de las funciones HL-CPWL que ajusten a la nube de puntos estudiada. El ajuste se realizará por mínimos cuadrados en la Sección 5.3 y posteriormente se evaluarán los errores obtenidos.

5.1 Nube de puntos.

Para realizar la comparación usaremos cuatro diferentes nubes de puntos obtenidas al muestrear, con dos patrones distintos, puntos sobre las gráficas de dos funciones diferentes. La primera de estas funciones será la función $\sin(\hat{x}_1)$ y la segunda función será $\sin\left(\frac{\hat{x}_1}{2} + \frac{\hat{x}_2}{2}\right)$, de esta forma el primer par de nubes de puntos se compone de datos de la forma $(\hat{x}_1, \hat{x}_2, \sin(\hat{x}_1))$ y el segundo par de nubes de puntos se compone de datos de la forma $(\hat{x}_1, \hat{x}_2, \sin\left(\frac{\hat{x}_1}{2} + \frac{\hat{x}_2}{2}\right))$.

Las nubes de puntos fueron tomadas evaluando estas funciones en el rectángulo $[0, 6] \times [0, 6]$, ordenando las tuplas de coordenadas (\hat{x}_1, \hat{x}_2) , en patrones uniformemente espaciados, de la siguiente manera.

- Muestreo a: los valores $\hat{x}_1 \in \left\{0, \frac{2}{3}, \frac{4}{3}, \dots, \frac{16}{3}, \frac{18}{3}\right\} \uparrow$ y $\hat{x}_2 \in \left\{0, \frac{2}{3}, \frac{4}{3}, \dots, \frac{16}{3}, \frac{18}{3}\right\} \uparrow$. Esta nube de puntos puede verse gráficamente en el plano x_1x_2 en la Figura 5.1a.
- Muestreo b: los valores $\hat{x}_1 \in \left\{0, \frac{3}{4}, \frac{6}{4}, \dots, \frac{21}{4}, \frac{24}{4}\right\} \uparrow$, además, si $\hat{x}_1 \in \left\{0, \frac{6}{4}, \frac{12}{4}, \frac{18}{4}, \frac{24}{4}\right\}$, entonces $\hat{x}_2 \in \left\{\frac{3}{8}, \frac{9}{8}, \dots, \frac{39}{8}, \frac{45}{8}\right\} \uparrow$, si $\hat{x}_1 \in \left\{\frac{3}{4}, \frac{9}{4}, \frac{15}{4}, \frac{21}{4}\right\}$, entonces $\hat{x}_2 \in \left\{0, \frac{3}{4}, \frac{6}{4}, \dots, \frac{21}{4}, \frac{24}{4}\right\} \uparrow$. Esta nube de puntos puede verse gráficamente en el plano x_1x_2 en la Figura 5.1b.

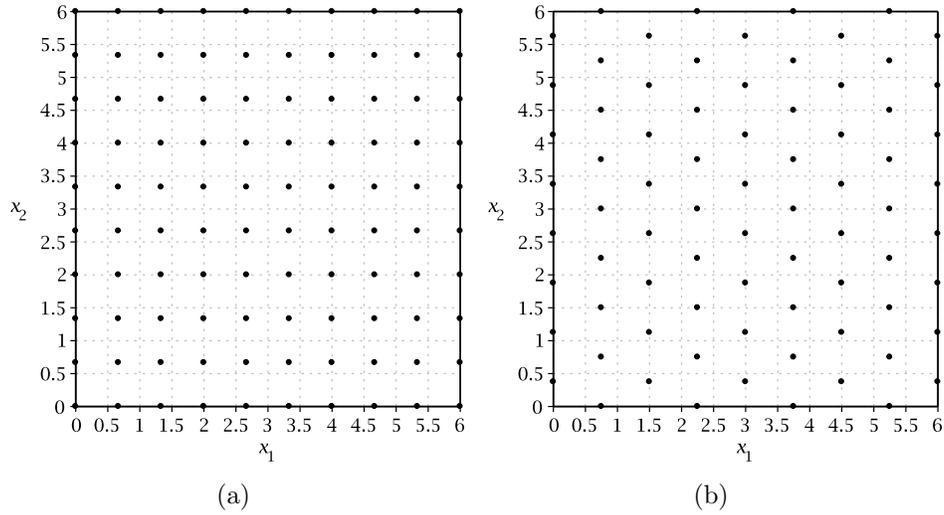


Figura 5.1: (a) Muestreo a. (b) Muestreo b. Puntos del plano x_1x_2 donde fueron evaluadas las funciones para obtener las nubes de puntos en el espacio de tres dimensiones. La línea sólida indica la frontera del rectángulo del que fueron tomados estos puntos.

5.2 Funciones continuas PWL a ajustar.

Como se mencionó anteriormente, el ajuste a las nubes de puntos se hará por funciones continuas PWL con diferentes particiones. Para dar estas particiones, de acuerdo con (4.1), primero daremos los valores de los vectores Ξ , \mathbf{p} y los parámetros m_1 y m_2 para cada función a ajustar.

5.2.1 Particiones del dominio.

Partición H_1 .

Para la primera partición, a la cual llamaremos H_1 , observemos que, en (4.1), si sustituimos los vectores Ξ por los vectores canónicos multiplicados por un escalar δ y el vector \mathbf{p} lo situamos en el origen de \mathbb{R}^n , entonces, (4.1) es de la forma (3.16), de donde se sigue que la propuesta de construcción dada en la Sección 4.1 del Capítulo 4 es una generalización de la dada en [22]. Así, los vectores Ξ de esta partición serán los vectores canónicos multiplicados por el escalar 2, $\mathbf{p} = \mathbf{0}$, con $m_1 = 3$ y $m_2 = 3$. Así, definimos a

$$D_1 = \{(2x_1 \ 2x_2) \in \mathbb{R}^2 : 0 \leq x_1, \leq 3, 0 \leq x_2 \leq 3\},$$

con partición H_1 formada por los hiperplanos:

$$\begin{aligned}
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_2 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 4 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_2 - 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 8 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_2 - 8 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 + 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 + 8 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 - 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 - 8 = 0\}.
\end{aligned}$$

Para una representación gráfica de D_1 y la partición H_1 ver Figura 5.2a.

Partición H_2 .

Para la segunda partición, a la cual llamaremos H_2 , se consideran en (4.1), los vectores y valores $\boldsymbol{\xi}_1 = (2 \ -2)^T$, $\boldsymbol{\xi}_2 = (2 \ 2)^T$, $\mathbf{p} = (-3 \ 3)^T$, $m_1 = 3$ y $m_2 = 3$. Así, definimos a

$$D_2 := \{M\mathbf{x} + \mathbf{p} : \mathbf{x} = (x_1 \ x_2) \in \mathbb{R}^2, 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 3\},$$

donde

$$M = \begin{bmatrix} 2 & 2 \\ -2 & 2 \end{bmatrix},$$

con partición H_2 formada por los hiperplanos:

$$\begin{aligned}
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 + 12 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 + 4 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 8 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 - 4 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 16 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : -4x_2 + 12 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : -4x_2 + 20 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : -4x_2 + 28 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : -4x_2 + 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : -4x_2 - 4 = 0\}.
\end{aligned}$$

Para una representación gráfica de D_2 y la partición H_2 ver Figura 5.2b.

Partición H_3 .

Para la tercera partición, a la cual llamaremos H_3 , se consideran en (4.1), los vectores y valores $\boldsymbol{\xi}_1 = (2 \ 2)^T$, $\boldsymbol{\xi}_2 = (-2 \ 2)^T$, $\mathbf{p} = (3 \ -3)^T$, $m_1 = 3$ y $m_2 = 3$. Así, definimos a

$$D_3 := \{M^T\mathbf{x} + \mathbf{p} : \mathbf{x} = (x_1 \ x_2) \in \mathbb{R}^2, 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 3\},$$

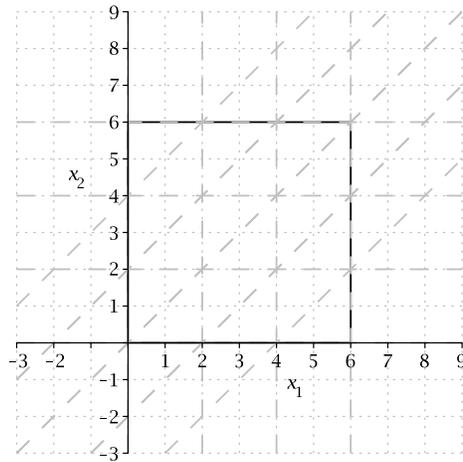
donde

$$M = \begin{bmatrix} 2 & -2 \\ 2 & 2 \end{bmatrix},$$

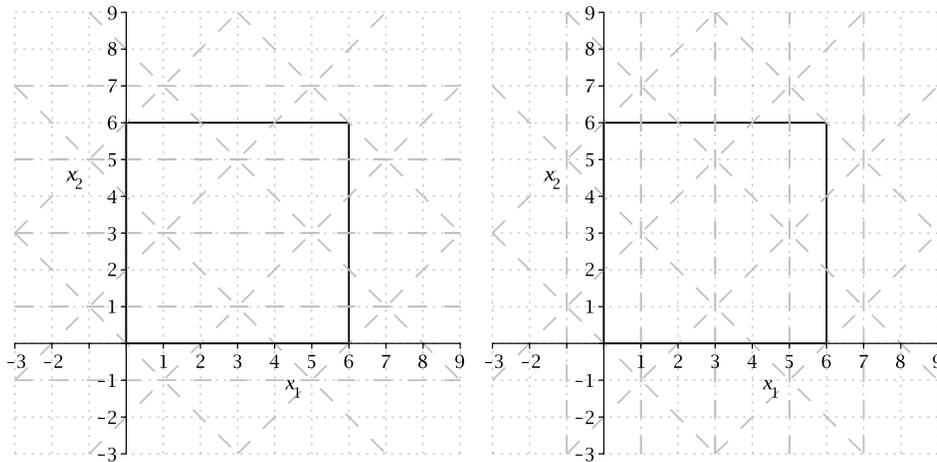
con partición H_3 formada por los hiperplanos:

$$\begin{aligned}
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : -2x_1 + 2x_2 + 12 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 8 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : -2x_1 + 2x_2 + 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 16 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : -2x_1 + 2x_2 - 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 4x_1 - 12 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 4x_1 - 4 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 4x_1 + 4 = 0\}, & \{\mathbf{x} \in \mathbb{R}^2 : 4x_1 - 20 = 0\}, \\
&\{\mathbf{x} \in \mathbb{R}^2 : 4x_1 - 28 = 0\}.
\end{aligned}$$

Para una representación gráfica de D_3 y la partición H_3 ver Figura 5.2c.



(a)



(b)

(c)

Figura 5.2: (a) Partición H_1 de D_1 . (b) Partición H_2 de D_2 . Partición H_3 de D_3 . Para los tres casos, tanto los hiperplanos de cada partición como la frontera de los conjuntos que particionan se encuentran señalados por la línea espaciada de color gris. Se señala también, con una línea de color negro, la frontera del conjunto de donde se toman las nubes de puntos.

5.2.2 Conjuntos generadores.

Conjunto generador Λ_{D_1} .

De acuerdo con la partición H_1 y la construcción descrita en la Sección 4.1.2, dado $\mathfrak{S}_{H_1} = \{1, \pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}(\mathbf{x}), \pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}(\mathbf{x}), \pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}(\mathbf{x}), \pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}(\mathbf{x}), \pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}(\mathbf{x}), \pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}(\mathbf{x})\}$, con

$$\begin{aligned} \pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}(\mathbf{x}) &= 2x_1, & \pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}(\mathbf{x}) &= 2x_2, \\ \pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}(\mathbf{x}) &= 2x_1 - 4, & \pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}(\mathbf{x}) &= 2x_2 - 4, \\ \pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}(\mathbf{x}) &= 2x_1 - 8, & \pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}(\mathbf{x}) &= 2x_2 - 8, \end{aligned}$$

los elementos de una función generadora del espacio $PWL_{H_1}[D_1]$ se ordenan en el vector $\Lambda_{D_1} =$

$$\begin{aligned} & \left[\gamma^{(0)}(1) \quad \gamma^{(1)}\left(\pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}\right) \quad \gamma^{(1)}\left(\pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}\right) \quad \gamma^{(1)}\left(\pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}\right) \quad \gamma^{(1)}\left(\pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}\right) \right. \\ & \gamma^{(1)}\left(\pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}\right) \quad \gamma^{(1)}\left(\pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}\right) \\ & \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(0),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}\right) \\ & \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(1),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(0),(0 \ 0)^T}\right) \\ & \left. \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(1),(0 \ 0)^T}\right) \quad \gamma^{(2)}\left(\pi_{(2 \ 0)^T}^{(2),(0 \ 0)^T}, \pi_{(0 \ 2)^T}^{(2),(0 \ 0)^T}\right) \right]^T. \end{aligned}$$

Conjunto generador Λ_{D_2} .

De acuerdo con la partición H_2 y la construcción descrita en la Sección 4.1.2, dado $\mathfrak{S}_{H_2} = \{1, \pi_{(2 \ -2)^T}^{(0),(-3 \ 3)^T}(\mathbf{x}), \pi_{(2 \ -2)^T}^{(1),(-3 \ 3)^T}(\mathbf{x}), \pi_{(2 \ -2)^T}^{(2),(-3 \ 3)^T}(\mathbf{x}), \pi_{(2 \ 2)^T}^{(0),(-3 \ 3)^T}(\mathbf{x}), \pi_{(2 \ 2)^T}^{(1),(-3 \ 3)^T}(\mathbf{x}), \pi_{(2 \ 2)^T}^{(2),(-3 \ 3)^T}(\mathbf{x})\}$, con

$$\begin{aligned} \pi_{(2 \ -2)^T}^{(0),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 - 2x_2 + 12, & \pi_{(2 \ 2)^T}^{(0),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 + 2x_2, \\ \pi_{(2 \ -2)^T}^{(1),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 - 2x_2 + 4, & \pi_{(2 \ 2)^T}^{(1),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 + 2x_2 - 8, \\ \pi_{(2 \ -2)^T}^{(2),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 - 2x_2 - 4, & \pi_{(2 \ 2)^T}^{(2),(-3 \ 3)^T}(\mathbf{x}) &= 2x_1 + 2x_2 - 16, \end{aligned}$$

los elementos de una función generadora del espacio $PWL_{H_2}[D_2]$ se ordenan en el vector $\Lambda'_{D_2} =$

$$\left[\begin{array}{l} \gamma^{(0)}(1) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \end{array} \right]^T.$$

Notemos en la Figura 5.2b que los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : -4x_2 + 28 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : -4x_2 - 4 = 0\}$ no se intersectan con la región de donde se tomaron las nubes de puntos. De igual forma, la intersección de los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 + 12 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 16 = 0\}$ y la intersección de los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 - 2x_2 - 4 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 = 0\}$, se encuentran fuera del conjunto de donde se tomaron las nubes de puntos. De la prueba del Lema 20 se sigue que la función $\gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right)$ está definida en el hiperplano $\{\mathbf{x} \in \mathbb{R}^2 : (2x_1 - 2x_2 + 12) - (2x_1 + 2x_2 - 16) = -4x_2 + 28 = 0\}$ y la función $\gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right)$ está definida en el hiperplano $\{\mathbf{x} \in \mathbb{R}^2 : (2x_1 - 2x_2 - 4) - (2x_1 + 2x_2) = -4x_2 - 4 = 0\}$. Por lo que ambas funciones pueden salir de Λ'_{D_2} , dando como resultado $\Lambda_{D_2} =$

$$\left[\begin{array}{l} \gamma^{(0)}(1) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(1)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}\right) \\ \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(-3 \ 3)^T}\right) \quad \gamma^{(2)}\left(\pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(-3 \ 3)^T}, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(-3 \ 3)^T}\right) \end{array} \right]^T.$$

Conjunto generador Λ_{D_3} .

De acuerdo con la partición H_3 y la construcción descrita en la Sección 4.1.2, dado un $\mathfrak{S}_{H_3} = \left\{ 1, \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(\mathfrak{3} \ -3)^T}(\mathbf{x}), \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(\mathfrak{3} \ -3)^T}(\mathbf{x}), \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(\mathfrak{3} \ -3)^T}(\mathbf{x}), \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(0),(\mathfrak{3} \ -3)^T}(\mathbf{x}), \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(1),(\mathfrak{3} \ -3)^T}(\mathbf{x}), \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(2),(\mathfrak{3} \ -3)^T}(\mathbf{x}) \right\}$, con

$$\begin{array}{ll} \pi_{\binom{2}{2} \binom{-3}{-2}}^{(0),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = 2x_1 + 2x_2 & \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(0),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = -2x_1 + 2x_2 + 12, \\ \pi_{\binom{2}{2} \binom{-3}{-2}}^{(1),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = 2x_1 + 2x_2 - 8, & \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(1),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = -2x_1 + 2x_2 + 4, \\ \pi_{\binom{2}{2} \binom{-3}{-2}}^{(2),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = 2x_1 + 2x_2 - 16, & \pi_{\binom{2}{-2} \binom{-3}{-2}}^{(2),(\mathfrak{3} \ -3)^T}(\mathbf{x}) = -2x_1 + 2x_2 - 4, \end{array}$$

los elementos de una función generadora del espacio $PWL_{H_3} [D_3]$ se ordenan en el vector $\Lambda'_{D_3} =$

$$\left[\begin{array}{l} \gamma^{(0)}(1) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \end{array} \right]^T.$$

Notemos en la Figura 5.2b que los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : 4x_1 + 4 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : 4x_1 - 28 = 0\}$ no se intersectan con la región de donde se tomaron las nubes de puntos. De igual forma, la intersección de los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : -2x_1 + 2x_2 - 4 = 0\}$ y la intersección de los hiperplanos $\{\mathbf{x} \in \mathbb{R}^2 : 2x_1 + 2x_2 - 16 = 0\}$ y $\{\mathbf{x} \in \mathbb{R}^2 : -2x_1 + 2x_2 + 12 = 0\}$, se encuentran fuera del conjunto de donde se tomaron las nubes de puntos. De la prueba del Lema 20 se sigue que la función $\gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right)$ está definida en el hiperplano $\{\mathbf{x} \in \mathbb{R}^2 : (2x_1 + 2x_2) - (-2x_1 + 2x_2 - 4) = 4x_1 + 4 = 0\}$ y la función $\gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right)$ está definida en el hiperplano $\{\mathbf{x} \in \mathbb{R}^2 : (2x_1 + 2x_2 - 16) - (-2x_1 + 2x_2 + 12) = 4x_1 - 28 = 0\}$. Por lo que ambas funciones pueden salir de Λ'_{D_3} , dando como resultado $\Lambda_{D_3} =$

$$\left[\begin{array}{l} \gamma^{(0)}(1) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(1)} \left(\pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(0),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(1),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(1),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \\ \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(0),(3) \quad -3} \right) \quad \gamma^{(2)} \left(\pi_{\binom{2}{2}}^{(2),(3) \quad -3}, \pi_{\binom{-2}{2}}^{(2),(3) \quad -3} \right) \end{array} \right]^T.$$

Observemos que Λ_{D_1} es un vector con 16 elementos, mientras que Λ_{D_2} y Λ_{D_3} tienen 14, esto se traduce computacionalmente en un menor uso de memoria para almacenarlos y en una menor cantidad de procesos al evaluarlos.

5.3 Ajuste por mínimos cuadrados.

Sean X_h , $h \in \{1, \dots, 4\}$, las cuatro nubes de puntos descritas en la Sección 5.1 donde cada elemento $\hat{\mathbf{x}}_i \in X_h$, con $i \in \{1, \dots, q_h\}$, siendo q_h el número de elementos en X_h , es de la forma $(\hat{x}_{i,1} \quad \hat{x}_{i,2} \quad \hat{x}_{i,3})$, y sea Λ_{D_j} , $j \in \{1, 2, 3\}$, una de las tres bases obtenidas en la

Sección 5.2.2. La solución del ajuste por mínimos cuadrados nos devuelve un vector $\mathbf{c}_{h,j}$, de la misma dimensión de $\mathbf{\Lambda}_{D_j}$, que minimiza el valor de $E_{h,j}$ donde

$$E_{h,j}^2 = \sum_{i=1}^{q_h} (\hat{x}_{i,3} - \mathbf{c}_{h,j}^T \mathbf{\Lambda}_{D_j} ((\hat{x}_{i,1} \quad \hat{x}_{i,2})))^2. \quad (5.1)$$

En la Tabla 5.1, donde las columnas nos indican las bases con las que se hace el ajuste a las nubes de puntos y las filas nos indican estas nubes de puntos, la intersección de fila y columna nos da los valores de $E_{h,j}$.

		$\mathbf{\Lambda}_{D_1}$	$\mathbf{\Lambda}_{D_2}$	$\mathbf{\Lambda}_{D_3}$
sen(x_1)	Muestreo a	1.6392293289631	3.6367643187529	1.2968163700078
	Muestreo b	1.2057495712198	3.0368647496623	1.2199720946749
sen($\frac{x_1}{2} + \frac{x_2}{2}$)	Muestreo a	1.1465415948070	1.0517311284773	1.0517311284773
	Muestreo b	0.8529014203304	0.7769915915593	0.7898330073039

Tabla 5.1: Bases que se ajustan a las nubes de puntos (columnas). Nubes de puntos obtenidas de las funciones $\text{sen}(x_1)$ y $\text{sen}(\frac{x_1}{2} + \frac{x_2}{2})$ por los muestreos a y b (filas). Valores de $E_{h,j}$ (intersección).

En la Tabla 5.1 vemos que el ajuste a la nube de puntos, obtenida evaluando la función $\text{sen}(x_1)$ con el muestreo a, usando la base $\mathbf{\Lambda}_{D_3}$ es numéricamente mejor que el ajuste por la base $\mathbf{\Lambda}_{D_1}$, el cual a su vez supera el ajuste por la base $\mathbf{\Lambda}_{D_2}$, mientras que en la nube de puntos obtenidas por el muestreo b, el ajuste por la base $\mathbf{\Lambda}_{D_1}$ es numéricamente mejor a los ajustes por las bases $\mathbf{\Lambda}_{D_2}$ y $\mathbf{\Lambda}_{D_3}$. Además, se puede ver también que para las nubes de puntos obtenidas de la función $\text{sen}(\frac{x_1}{2} + \frac{x_2}{2})$, tanto del muestreo a como el muestreo b, tenemos que los ajustes obtenidos de $\mathbf{\Lambda}_{D_2}$ y $\mathbf{\Lambda}_{D_3}$ son numéricamente mejores a los obtenidos por $\mathbf{\Lambda}_{D_1}$. Con esto podemos afirmar; primero, que se puede dar casos en los que con una partición simplicial, no ortogonal y usando las propuestas en este trabajo, mejore el ajuste a una nube de puntos en comparación con el esquema de construcción propuesto en [22]; y segundo, que el encontrar una malla que mejor ajuste a una nube de puntos requiere tomar en consideración al menos dos factores, los cuales son el comportamiento de la superficie de la cual son tomados los datos y el esquema por el cual son tomados sus datos.

Para ayudarnos a visualizar de que manera influye la elección de una partición al ajustarse a una nube de puntos obtenida de una superficie veamos la Figura 5.3 y la Figura 5.4. En ambas podemos corroborar gráficamente los resultados de la Tabla 5.1. En la Figura 5.3 vemos que las funciones continuas PWL obtenidas por el ajuste de la base de $\mathbf{\Lambda}_{D_2}$, las cuales fueron las que presentaron el ajuste más bajo para este caso, muestra un comportamiento diferente al que se esperaría de una función PWL que pretende ajustarse a una función $\text{sen}(x_1)$ o a una nube de puntos obtenida de esta. Lo mismo ocurre en los casos presentados en la Figura 5.4, donde, de acuerdo con la Tabla 5.1, las funciones continuas PWL obtenidas por $\mathbf{\Lambda}_{D_2}$, son las que presentan un menor ajuste. Lo que ambas casos comparten en común es que el vector $\boldsymbol{\omega}_{1,2}$, de sus respectivas particiones es ortogonal a la dirección del vector en el que ambas funciones presentan mayor cambio, esto puede

evaluarse, ya que el gradiente de las funciones de donde fueron tomados los datos, y su inverso aditivo son las direcciones de cambio más rápido, $\nabla \sin(x_1)$ es siempre un vector horizontal, y $\nabla \sin\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$ está siempre a 45° respecto al eje x_1 . Dicho de otra manera, el hiperplano generado por el $\omega_{1,2}$ es paralelo a la dirección de mayor cambio en sus respectivas funciones, por lo que el ajuste no refleja adecuadamente estos cambios. Lo cual nos permite establecer un punto de partida para determinar la partición simplicial de una función PWL que mejor se ajuste a una nube de puntos obtenida de una superficie.

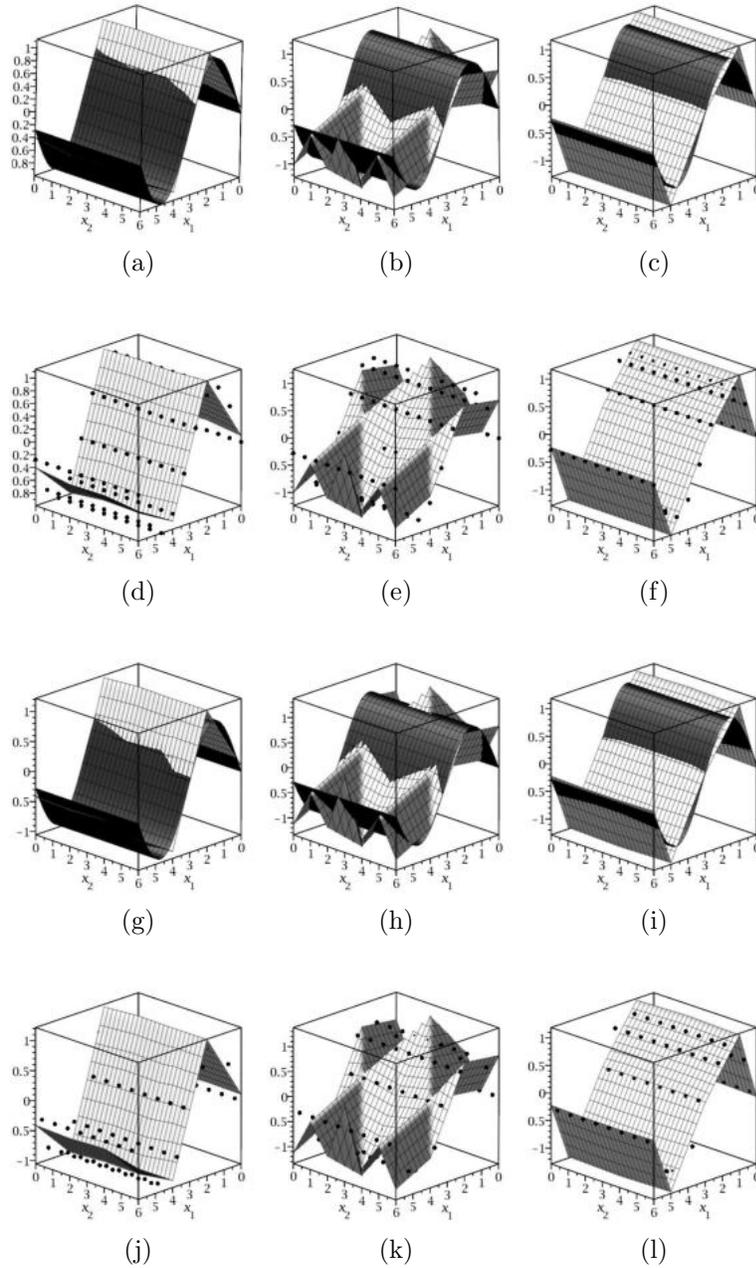


Figura 5.3: (a - c) Función $\sin(x_1)$ (color negro) y el ajuste a las nubes de puntos, muestreo a, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (d - f) Nubes de puntos, muestreo a obtenido de la función $\sin(x_1)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (g - i) Función $\sin(x_1)$ (color negro) y el ajuste a las nubes de puntos, muestreo b, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (j - l) Nubes de puntos, muestreo b obtenido de la función $\sin(x_1)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco).

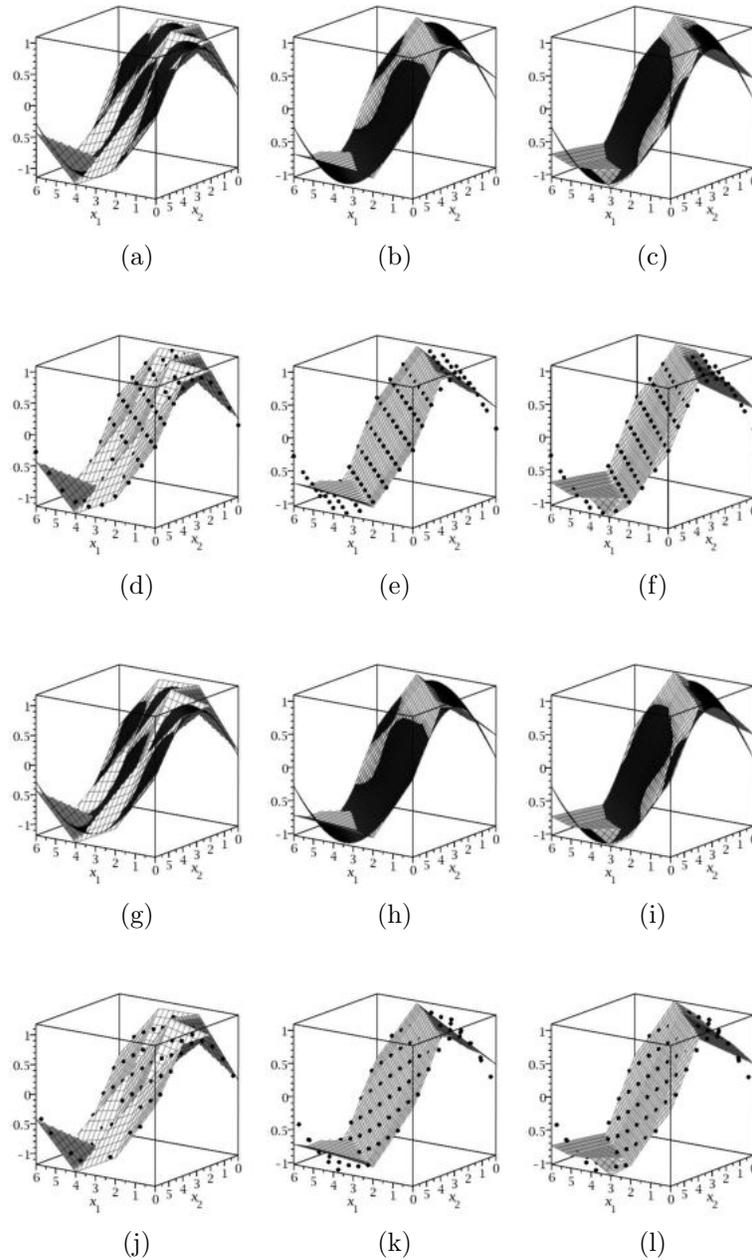


Figura 5.4: (a - c) Función $\sin\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$ (color negro) y el ajuste a las nubes de puntos, muestreo a, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (d - f) Nubes de puntos, muestreo a obtenido de la función $\sin\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (g - i) Función $\sin\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$ (color negro) y el ajuste a las nubes de puntos, muestreo b, por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco). (j - l) Nubes de puntos, muestreo b obtenido de la función $\sin\left(\frac{x_1}{2} + \frac{x_2}{2}\right)$, y ajuste a esta nube de puntos por Λ_{D_1} , Λ_{D_2} , Λ_{D_3} , respectivamente (color blanco).

Discusión.

Conclusiones.

En este trabajo se ha descrito un nuevo método para encontrar expresiones canónicas de funciones PWL con particiones simpliciales no ortogonales. Este esquema es una generalización de la construcción dada en [22], como se mencionó en la Sección 5.2.1, si sustituimos la base Ξ en (4.1), por los vectores canónicos, tenemos entonces la partición simplicial propuesta en [22]. Como se vio en la Sección 4.1.3, se mantuvieron las ventajas computacionales de la construcción simplicial ortogonal uniformemente espaciada, tales como el procedimiento de construcción de los vectores que forman una base para el espacio de funciones PWL con partición simplicial, así como la existencia de una solución para hallar el vector de coeficientes \mathbf{c} en (4.28) puesto que este esquema de construcción permite garantizar la existencia de la matriz inversa de A en (4.28). Además, se probó también que la matriz A es triangular inferior, por lo que el cálculo de coeficientes en \mathbf{c} puede ser obtenido por (4.29) el cual representa un menor esfuerzo computacional que el utilizado para el cálculo de la inversión de A . Más aún, puesto que la construcción aquí propuesta considera que cada intersección presenta a lo más la propiedad de degeneración mínima, esta tiene el número mínimo de parámetros necesarios para la representación de una función continua PWL, tal y como se establece en [20]. En la Sección 4.1.4, se brinda un algoritmo que sintetiza el método propuesto y permite la construcción de representaciones HL-CPWL de funciones PWL con partición simplicial, no necesariamente ortogonal.

Sobre el ajuste a una nube de puntos, se vio en la Sección 5.3 que, para algunos casos, se puede dar una función continua PWL con un esquema de partición simplicial la cual requiera de un menor número de parámetros de ajuste en comparación con la propuesta dada en [22], esto se consigue seleccionando una partición del dominio que tome en consideración características relacionadas al comportamiento de la superficie de la cual fueron extraídos los datos de la nube de puntos, así como el patrón (de haberlo) en que fueron tomados.

En la Sección 4.2 se vio que una extensión de las fórmulas (2.35), (2.36) y (2.37), usadas para la construcción de representaciones CPWL puede ser usada para calcular funciones con un dominio en \mathbb{R}^2 que interpolen datos en \mathbb{R}^3 que se encuentran alineados sobre una malla ortogonal espaciada en forma uniforme proyectada en el plano \mathbb{R}^2 , considerando cada una de las columnas como la proyección de una colección de representaciones CPWL de funciones PWL con un dominio en \mathbb{R} . Una comparación de esta propuesta contra la

descrita [22] revela que esta requiere un menor tiempo de cómputo y demanda menor uso de memoria [27]. Cabe señalar que esta función no es una función PWL.

En conclusión, fue posible establecer un algoritmo que permite construir representaciones canónicas de funciones PWL con una partición simplicial, no necesariamente ortogonal, que se ajusten a una nube de puntos, con el menor número posible de parámetros y haciendo el menor uso de recursos computacionales en comparación a los métodos existentes en la literatura.

Trabajos futuros.

Dada una nube de puntos en el espacio y su proyección a un plano, es un problema abierto el determinar una partición simplicial sobre ese plano (lo cual incluye determinar los vectores de la base, \mathbf{p} , los enteros m_1, m_2) que optimice el ajuste a la nube de puntos por medio de la función HL-CPWL cuya construcción se ha descrito en este trabajo. Además, este problema incluye la estimación de cotas que permita refinar o engrosar las mallas de la partición, con la finalidad de optimizar el tiempo de cómputo al obtener una función HL-CPWL que ajuste a la nube de puntos.

La propuesta dada en la Sección 4.2 para calcular los coeficientes usados en una representación CPWL con un dominio en \mathbb{R}^2 ajusta únicamente datos sobre una malla ortogonal uniformemente espaciada, es decir, datos que se encuentran en la intersección de los hiperplanos dados en una partición como la propuesta en [22], por lo que, combinar esta estrategia de construcción con la desarrollada en la Sección 4.1 para particiones simpliciales no necesariamente ortogonales, permitiría obtener una función la cual ajuste datos sobre mallas más flexibles para representaciones CPWL.

Referencias

- [1] L. Chua, S. M. Kang, Section-wise piecewise-linear functions: Canonical representation, properties, and applications, *Proceedings of the IEEE* 65 (6) (1977) 915–929. doi:10.1109/proc.1977.10589.
- [2] S. Kang, L. Chua, A global representation of multidimensional piecewise-linear functions with linear partitions, *IEEE Transactions on Circuits and Systems* 25 (11) (1978) 938–940. doi:10.1109/tcs.1978.1084401.
- [3] L. Chua, R. Ying, Canonical piecewise-linear analysis, *IEEE Transactions on Circuits and Systems* 30 (3) (1983) 125–140. doi:10.1109/tcs.1983.1085342.
- [4] L. Chua, A. C. Deng, Canonical piecewise-linear modeling, *IEEE Transactions on Circuits and Systems* 33 (5) (1986) 511–525. doi:10.1109/tcs.1986.1085952.
- [5] L. Chua, C. Desoer, E. Kuh, *Linear and Nonlinear Circuits*, Circuits and systems, McGraw-Hill, 1987.
- [6] A. Oliveri, M. Reimers, M. Storace, Automatic domain partitioning of piecewise-affine simplicial functions implementing model predictive controllers, *IEEE Transactions on Circuits and Systems II: Express Briefs* 62 (9) (2015) 886–890. doi:10.1109/TCSII.2015.2435971.
- [7] V. M. Jimenez-Fernandez, H. Vazquez-Leal, U. A. Filobello-Nino, M. Jimenez-Fernandez, L. J. Morales-Mendoza, M. Gonzalez-Lee, Exploring the use of two-dimensional piecewise-linear functions as an alternative model for representing and processing grayscale-images, *Journal of Applied Research and Technology* 14 (5) (2016) 311–318. doi:10.1016/j.jart.2016.09.001.
- [8] J. Zheng, C. Xu, D. Guo, A programmable cnn architecture and its hardware-software co-design approach for image processing and stimulating visual illusions, in: *2016 IEEE International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 2016, pp. 389–394. doi:10.1109/QRS-C.2016.59.
- [9] J. Zhang, K. S. Chin, M. Ławryńczuk, Multilinear model decomposition and predictive control of mimo two-block cascade systems, *Industrial & Engineering Chemistry Research* 56 (47) (2017) 14101–14114. doi:10.1021/acs.iecr.7b03265.

-
- [10] J. Zhang, K.-S. Chin, M. Ławryńczuk, Nonlinear model predictive control based on piecewise linear hammerstein models, *Nonlinear Dynamics* doi:10.1007/s11071-018-4105-5.
- [11] X. Gao, Y. Jiang, T. Chen, D. Huang, Optimizing scheduling of refinery operations based on piecewise linear models, *Computers Chemical Engineering* 75 (2015) 105–119. doi:10.1016/j.compchemeng.2015.01.022.
- [12] M. Y. Cheong, S. Werner, M. J. Bruno, J. L. Figueroa, J. E. Cousseau, R. Wichman, Adaptive piecewise linear predistorters for nonlinear power amplifiers with memory, *IEEE Transactions on Circuits and Systems I: Regular Papers* 59 (7) (2012) 1519–1532. doi:10.1109/tcsi.2011.2177007.
- [13] T. Poggi, M. Rubagotti, A. Bemporad, M. Storace, High-speed piecewise affine virtual sensors, *IEEE Transactions on Industrial Electronics* 59 (2) (2012) 1228–1237. doi:10.1109/tie.2011.2161064.
- [14] N. Rodriguez, P. Julian, E. Paolini, A simplicial piecewise linear approach for efficient hardware realization of neural networks : (invited presentation), in: 2019 53rd Annual Conference on Information Sciences and Systems (CISS), 2019, pp. 1–3. doi:10.1109/CISS.2019.8692902.
- [15] V. M. Tlapa-Carrera, V. M. Jimenez-Fernandez, H. Vazquez-Leal, U. A. Filobello-Nino, J. Garcia-Guzman, F. J. Castro-Gonzalez, Modeling of three-dimensional surfaces using high-level canonical piecewise-linear functions in cylindrical coordinates, *Computational and Applied Mathematics* 37 (4) (2018) 5500–5513. doi:10.1007/s40314-018-0647-x.
- [16] L. Chua, A. C. Deng, Canonical piecewise-linear representation, *IEEE Transactions on Circuits and Systems* 35 (1) (1988) 101–111. doi:10.1109/31.1705.
- [17] C. Kahlert, L. O. Chua, A generalized canonical piecewise-linear representation, *IEEE Transactions on Circuits and Systems* 37 (3) (1990) 373–383. doi:10.1109/31.52731.
- [18] J. N. Lin, H. Q. Xu, R. Unbehauen, A generalization of canonical piecewise-linear functions, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 41 (4) (1994) 345–347. doi:10.1109/81.285696.
- [19] P. Julian, The complete canonical piecewise-linear representation: functional form for minimal degenerate intersections, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 50 (3) (2003) 387–396. doi:10.1109/TCSI.2003.808904.
- [20] C. Kahlert, L. Chua, The complete canonical piecewise-linear representation. i. the geometry of the domain space, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 39 (3) (1992) 222–236. doi:10.1109/81.128016.

-
- [21] X. Gao, Y. Jiang, D. Huang, Z. Xiong, A novel high level canonical piecewise linear model based on the simplicial partition and its application, *ISA Transactions* 53 (5) (2014) 1420–1426. doi:10.1016/j.isatra.2013.12.027.
- [22] P. Julian, A. Desages, O. Agamennoni, High-level canonical piecewise linear representation using a simplicial partition, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 46 (4) (1999) 463–480. doi:10.1109/81.754847.
- [23] X. Gao, Y. Jiang, D. Huang, Z. Xiong, A novel high level canonical piecewise linear model based on the simplicial partition and its application, in: *2013 10th IEEE International Conference on Control and Automation (ICCA)*, 2013, pp. 1856–1861. doi:10.1109/ICCA.2013.6565070.
- [24] E. Chong, S. Zak, *An Introduction to Optimization*, Wiley Series in Discrete Mathematics and Optimization, Wiley, 2013.
- [25] Ji-Nan Lin, R. Unbehauen, Canonical piecewise-linear networks, *IEEE Transactions on Neural Networks* 6 (1) (1995) 43–50. doi:10.1109/72.363451.
- [26] D. C. Lay, J. J. McDonald, S. R. Lay, P. M. González, J. A. Pastor, C. G. Nagore, *Álgebra lineal y sus aplicaciones*, Pearson educación, 2016.
- [27] V. Jimenez-Fernandez, M. Jimenez-Fernandez, H. Vazquez-Leal, U. Filobello-Nino, C. Castañeda-Roldan, V. Tlapa-Carrera, A new methodology to extend the canonical piecewise-linear model from one to two dimensions, *National Academy Science Letters* doi:10.1007/s40009-020-00970-8.