Funciones submodulares y algunas aplicaciones

Arley Ramses Gómez Rios

Matemático

Universidad Nacional de Colombia

Facultad de Ciencias

Departamento de Matemáticas

Bogotá D.C.

Junio de 2011

Funciones submodulares y algunas aplicaciones

Arley Ramses Gómez Rios
Matemático

Trabajo final para optar al título de
Magister en Ciencias - Matemáticas

Director:
Humberto Sarria Zapata, PhD
Profesor asociado, Universidad Nacional de Colombia

Universidad Nacional de Colombia
Facultad de Ciencias
Departamento de Matemáticas
Bogotá D.C.
Junio de 2011

Título en Español

Funciones submodulares y algunas aplicaciones.

Title in English

Submodular functions and some applications.

Resumen: En este trabajo se hace una investigación sobre las funciones submodulares, tomando como base el estudio realizado por Satoru Fujishije [7]. Se estudian los conceptos de polimatroide y sistema submodular exhibiendo conjuntamente algunos ejemplos. Estos conceptos son aplicados en la construcción del Algoritmo Greedy, el cual permite solucionar cierto tipo de problemas de optimización lineal.

Abstract: On this paper a research about the submodular functions is made, the basis of the work comes from the studies made by Satoru Fujishije [7]. The polymatroid and submodular system concepts are studied exhibiting together some examples. These concepts are applied to the construction of the Greedy Algorithm, which can solve certain type of problems of linear optimization.

Palabras Clave: Funciónes submodulares, matroides, polimatroides, sistemas submodulares.

Keywords: Submodular functions, matroids, polymatroids, submodular systems.

Agradecimientos

Al profesor Humberto Sarria Zapata, por permitirme trabajar con él y por su valiosa colaboración y asesoría durante el desarrollo del trabajo, contribuyendo a mi formación académica y personal.

A la Universidad Nacional de Colombia, no sólo por mi formación académica, sino por los beneficios económicos y profesionales que he podido obtener gracias a ella, haciendo que esta experiencia fuese más agradable.

Muy especial a mis padres, Ismael Gómez y Livia Ríos, quienes me han apoyado incondicionalmente en cada paso que doy.

A mis hermanos y a mi novia, por su gran apoyo y colaboración.

A todas las personas: amigos, compañeros y familiares que de una u otra forma aportaron al desarrollo de este trabajo.

Índice general

Agradecimientos **II**

Introducción **2**

1. Definiciones, Notación, Teoremas Básicos y Preliminares **4**

1.1. Funciones Submodulares 4

1.2. Conjuntos 7

1.3. Grafos 8

1.4. Análisis Convexo y Desigualdades lineales 11

1.5. Matroides 12

2. Polimatroides **15**

2.1. Principales conceptos y resultados. 15

2.2. Ejemplos de Polimatroides: 25

2.3. Relación con matroides: 28

3. Sistemas Submodulares **33**

3.1. Principales conceptos y resultados 33

3.2. Reducciones y contracciones 40

3.3. Ejemplos de sistemas submodulares: 43

4. Algoritmo Greedy **50**

4.1. Retículos distributivos y posets. 51

4.2. Algoritmo greedy 57

Aportes del Autor **76**

Conclusiones **78**

Trabajo Futuro **79**

Bibliografía **80**

Introducción

Sea E un conjunto finito, denotemos por 2^E al conjunto de todos los subconjuntos de E. Se dice que una función $f : 2^E \longrightarrow \mathbb{R}$ es una *función submodular* en 2^E, si para cada par $X, Y \subseteq E$ se tiene:

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y).$$

Este tipo de funciones aparecen con frecuencia en el análisis de diferentes estructuras como grafos, redes, matroides, etc. En los últimos años, las funciones submodulares han tomado gran importancia en varios campos del análisis combinatorio. Fueron estudiadas inicialmente al rededor del año 1950 por H. Whitney y W.T. Tutte en relación con la Teoría de matroides, después, G. Choquet las analizó respecto a la Teoría de capacidad [2], y O. Ore las estudió dentro de la Teoría de grafos. Al rededor de 1960, J. Edmons desarrolló importantes resultados en sus trabajos relacionados con matroides y polimatroides. Después de esto, no sólo Edmons ha ido agregando importantes resultados teóricos al respecto [5], también encontramos autores como W. H. Cunningham, A. Frank, M. Iri, E. L. Lawler, y L. Lovász [4], [6] y [8], entre otros. Por otro lado, se han hecho estudios relacionados con Teoría de juegos de función característica [13], optimización [7], y aplicaciones a matroides y polimatroides, entre otros.

El objetivo principal de este trabajo es realizar un estudio detallado y recopilar información relacionada con las funciones submodulares y sus aplicaciones. Aquí se toma como base el estudio realizado por Satoru Fujishije en los capítulos I y II de [7], exhibiendo un trabajo más minucioso, escrito de manera clara y realizando numerosas pruebas que allí no aparecen.

En el capítulo 1, se presentan algunos resultados básicos relacionados con las funciones submodulares así como algunos preliminares, definiciones y teoremas previos al trabajo, estos están relacionados con la Teoría de matroides, la Teoría de grafos, la Teoría de redes y algunos resultados del Análisis convexo.

En el capítulo 2, se presenta el concepto de polimatroide, el cuál generaliza de cierta manera el concepto de matroide y se hace a partir de la submodularidad de la función rango del matroide, tal observación se presenta en detalle y se generalizan las nociones de conjunto independiente, base y función de clausura entre otras. Se exponen además algunos ejemplos de polimatroides.

Otro concepto importante y que es mucho más general que los conceptos de matroide y polimatroide es el de sistema submodular, tal concepto se estudia detalladamente en el capítulo 3, no sólo se presentan definiciones y resultados análogos a los de polimatroides, sino que se estudian algunos otros aspectos como reducciones y contracciones de sistemas submodulares. También se exhiben algunos ejemplos.

En el capítulo 4, se considera un problema de optimización lineal en el que interviene un sistema submodular y se presenta el Algoritmo Greedy, que permite solucionar el problema. Aquí es necesario aplicar diversos resultados previos presentados a lo largo del trabajo.

Capítulo 1

Definiciones, Notación, Teoremas Básicos y Preliminares

Este capítulo está dedicado a introducir las definiciones y notaciones básicas usadas en el trabajo, así como los teoremas previos utilizados. Estos están relacionados con algunas nociones elementales de funciones submodulares, teoría de matroides, teoría de grafos, teoría de redes y algunos resultados del análisis convexo. Sólo se hará la prueba de los teoremas más relevantes, las demás pruebas se encontrarán fácilmente utilizando la bibliografía.

1.1. Funciones Submodulares

Definición 1.1.1. *Sea E un conjunto finito, decimos que una función $f : 2^E \longrightarrow \mathbb{R}$ es una función submodular en 2^E, si para todos $X, Y \subseteq E$ se tiene la desigualdad:*

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y).$$

Más adelante introduciremos de manera más general el concepto de función submodular. Tenemos los siguientes resultados básicos sobre funciones submodulares:

Teorema 1.1.1. *Sea $f : 2^E \longrightarrow \mathbb{R}$ donde E es un conjunto finito. Las siguientes afirmaciones son equivalentes:*

i) f *es submodular en* 2^E.

ii) Para todo $A \subseteq E$ *y para todos* $i, j \in E - A$, $i \neq j$ *se tiene:*
$f(A \cup \{j\}) - f(A) \geq f(A \cup \{i, j\}) - f(A \cup \{i\})$

iii) Para todos $A, B \subseteq E$ *con* $A \subseteq B$ *y* $j \in E - B$ *se tiene:*
$f(A \cup \{j\}) - f(A) \geq f(B \cup \{j\}) - f(B)$

iv) Para todos $A, B, C \subseteq E$ *con* $A \subseteq B$ *y* $B \cap C = \varnothing$ *se tiene:*
$f(A \cup C) - f(A) \geq f(B \cup C) - f(B)$

Prueba: $(i \Rightarrow ii)$ Como f es submodular,

$$f(A \cup \{i\}) + f(A \cup \{j\}) \geq f(A \cup \{i, j\}) + f(A)$$

en otras palabras,

$$f(A \cup \{j\}) - f(A) \geq f(A \cup \{i, j\}) - f(A \cup \{i\}).$$

$(ii \Rightarrow iii)$ Si $A = B$ la desigualdad es obvia. Suponga que $A \neq B$. y sea $B - A = \{b_1, b_2, ..., b_n\}$ donde todos los b_m son distintos, tenemos:

$$\begin{aligned} f(A \cup \{j\}) - f(A) &\geq f(A \cup \{b_1, j\}) - f(A \cup \{b_1\}) \\ &\geq f(A \cup \{b_1, b_2, j\}) - f(A \cup \{b_1, b_2\}) \geq \cdots \\ &\geq f(A \cup \{b_1, b_2, ..., b_n, j\}) - f(A \cup \{b_1, b_2, ..., b_n\}) \\ &= f(B \cup \{j\}) - f(B) \end{aligned}$$

$(iii \Rightarrow iv)$ Si $C = \varnothing$ la desigualdad es obvia. Suponga que $C \neq \varnothing$ y sea $C = \{c_1, c_2, ...c_n\}$ donde todos los c_m son distintos.
Tenemos entonces las siguientes desigualdades.

$$f(A \cup \{c_1\}) - f(A) \geq f(B \cup \{c_1\}) - f(B)$$

$$f(A \cup \{c_1\} \cup \{c_2\}) - f(A \cup \{c_1\}) \geq f(B \cup \{c_1\} \cup \{c_2\}) - f(B \cup \{c_1\})$$

$$f(A \cup \{c_1, c_2, c_3\}) - f(A \cup \{c_1, c_2\}) \geq f(B \cup \{c_1, c_2, c_3\}) - f(B \cup \{c_1, c_2\})$$
$$\vdots$$
$$f(A \cup C) - f(A \cup \{c_1, c_2, ..., c_{n-1}\}) \geq f(B \cup C) - f(B \cup \{c_1, c_2, ..., c_{n-1}\})$$

Sumando todas estas desigualdades obtenemos:

$$f(A \cup C) - f(A) \geq f(B \cup C) - f(B)$$

$(iv \Rightarrow i)$ Sean $X, Y \subseteq E$. Si tomamos $A = X \cap Y$, $B = X$ y $C = Y - X$ entonces $A \subseteq B \subseteq E$ y $B \cap C = \varnothing$ con lo cual:

$$f\big((X \cap Y) \cup (Y - X)\big) - f(X \cap Y) \geq f\big(X \cup (Y - X)\big) - f(X)$$

en otras palabras

$$f(Y) - f(X \cap Y) \geq f(X \cup Y) - f(X),$$

es decir, f es submodular

$\square$

Teorema 1.1.2. *Sea $E = \{1, 2, ..., n\}$ y sea $g : \mathbb{R} \longrightarrow \mathbb{R}$ una función cóncava, es decir, para todos $x, y \in \mathbb{R}$ y $0 \leq t \leq 1$, se tiene $g\big(tx + (1-t)y\big) \geq tg(x) + (1-t)g(y)$. La función $\rho : 2^E \longrightarrow \mathbb{R}$ definida como $\rho(X) = g(|X|)$ $(X \subseteq E)$ es una función submodular en 2^E.*

Prueba: Como g es cóncava, entonces para todo $x \in \mathbb{R}$ y $0 \leq t \leq 1$

tenemos

$$g\big(tx + (1-t)(x+2))\big) \geq tg(x) + (1-t)g(x+2),$$

en particular para $t = 1/2$, es decir,

$$g(x+1) \geq \tfrac{1}{2}\big(g(x) + g(x+2)\big)$$

con lo cual, $g(x+1) + g(x+1) \geq g(x) + g(x+2)$ o equivalentemente,

$$g(x+1) - g(x) \geq g(x+2) - g(x+1)$$

luego en particular para $x = |X|$ con $X \subseteq E$ tenemos

$$g(|X|+1) - g(|X|) \geq g(|X|+2) - g(|X|+1).$$

así, para $i, j \in E - X$, con $i \neq j$ se tiene

$$g(|X \cup \{j\}|) - g(|X|) \geq g(|X \cup \{i,j\}|) - g(|X \cup \{i\}|).$$

o equivalentemente:

$$\rho(X \cup \{j\}) - \rho(X) \geq \rho(X \cup \{i,j\}) - \rho(X \cup \{i\}).$$

Por la proposición anterior, podemos concluir que ρ es submodular.

□

A lo largo del trabajo encontraremos más ejemplos de funciones submodulares.

1.2. Conjuntos

Para cada conjunto X, su cardinalidad es denotada por $|X|$, si X es un subconjunto de Y escribimos $X \subseteq Y$ y si X es un sobconjunto propio de Y escribimos $X \subset Y$.

Una familia de subconjuntos disyuntos $B_i, (i \in I)$ de un conjunto E se llama una *partición* de E, si $\bigcup_{i \in I} B_i = E$.

Dados dos conjuntos A y B, el conjunto

$$\{(1,a) \mid a \in A\} \cup \{(2,b) \mid b \in B\} \tag{1.1}$$

se llama la *suma directa* de A y B y es denotada como $A \oplus B$

Para un poset $\mathcal{P} = (A, \preceq)$ un subconjunto B de A es llamado un *ideal* de $\mathcal{P}$, si dado $y \in B$, $x \preceq y$ implica $x \in B$. Si para cada $x, y \in A$ existen $\sup\{x,y\}$ e $\inf\{x,y\}$, entonces el poset $\mathcal{P} = (A, \preceq)$ es llamado un *retículo*, en tal caso, escribimos $\sup\{x,y\} = x \vee y$ e $\inf\{x,y\} = x \wedge y$. Las operaciones binarias $\vee$ y $\wedge$ satisfacen:

(i) $\forall x \in A : x \vee x = x,\ x \wedge x = x.$

(ii) $\forall x, y \in A : x \vee y = y \vee x,\ x \wedge y = y \wedge x$

(iii) $\forall x, y, z \in A : x \vee (y \vee z) = (x \vee y) \vee z, \quad x \wedge (y \wedge z) = (x \wedge y) \wedge z$

(iv) $\forall x, y \in A : x \wedge (y \vee z) = x, \quad x \vee (x \wedge y) = x$

Recíprocamente, si un conjunto A es dotado con dos operaciones binarias $\vee$ y $\wedge$ las cuales satisfacen (i)$\sim$(iv), entonces definiendo la relación binaria $\preceq$ como

$$x \preceq y \Leftrightarrow x \vee y = y \text{ (o } x \wedge y = x) \tag{1.2}$$

obtenemos un retículo $(A, \preceq)$ cuyas operaciones de retículo son $\wedge$ y $\vee$. Un retículo $(A, \preceq)$ con operaciones $\wedge$ y $\vee$ también se expresa como $(A, \vee, \wedge)$.

Un retículo $\mathcal{L} = (A, \vee, \wedge)$ es un *retículo distributivo*, si satisface además

(v) $\forall x, y, z \in A : x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ y

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z).$$

Obsérvese que para un conjunto finito E, un conjunto $\mathcal{D} \subseteq 2^E$ cerrado para uniones e intersecciones es un *retículo distributivo* cuyas operaciones de retículo son la unión y la intersección de conjuntos.

Si $\mathcal{D}$ es el conjunto de ideales de un poset finito $(E, \preceq)$ entonces D es un retículo distributivo con la unión y la intersección siendo sus operaciones de retículo.

1.3. Grafos

Sean V y A conjuntos finitos y suponga que existen dos funciones ∂^+,$\partial^- : A \longrightarrow V$, entonces $G = (V, A : \partial^+, \partial^-)$ es llamado un *grafo*. Decimos que V es el conjunto de vértices y que A es el conjunto de arcos. Para cada arco $a \in A$, $\partial^+ a$ es llamado el *vértice inicial* (o la cola) de a y $\partial^- a$ es llamado el *vértice terminal* (o la cabeza) de a. Si no hay posibilidad de confusión, también denotamos el grafo por $G = (V, A)$. Un arco a también puede ser expresado por $(\partial^+ a, \partial^- a)$. Un grafo $G = (V, A)$ con conjunto de arcos $A = \{(u, v) \mid u, v \in V, u \neq v\}$ se denomina un *grafo dirigido completo*.

Un grafo $H = (W, B; \partial_H^+, \partial_H^-)$ tal que $W \subseteq V$, $B \subseteq A$ y ∂_H^+ y ∂_H^- son las restricciones de ∂^+ y ∂^- a B, es llamado un *subgrafo* de G.

Para cada vértice $v \in V$ definimos:

$$\delta^+ v := \{a \in A \mid \partial^+ a = v\} \text{ y } \delta^- v := \{a \in A \mid \partial^- a = v\} \tag{1.3}$$

y para cada subconjunto U de V definimos:

$$\Delta^+ U := \{a \in A \mid \partial^+ a \in U, \partial^- a \in V - U\} \tag{1.4}$$

$$\Delta^- U := \{a \in A \mid \partial^- a \in U, \partial^+ a \in V - U\} \tag{1.5}$$

Un *camino dirigido* en G es una sucesión alternante $(v_0, a_1, v_1, a_2, ..., V_{k-1}, a_k, v_k)$ de vértices v_i y arcos a_i $(i = 0, 1, ..., k)$ tales que $\partial^+ a_i = v_{i-1}$ y $\partial^- a_i = v_i$ para cada $i = 1, 2, ..., k$. Decimos aquí que el camino va de v_0 a v_k.

Un grafo $G = (V, A)$ es llamado *fuertemente conexo*, si para cada dos vértices $u, v \in V$ existe un camino dirigido de u a v. Un subgrafo maximal de G fuertemente conexo es llamado una *componente fuertemente conexa* de G. G se descompone en sus componentes fuertemente conexas $H_i = (V_i, A_i)(i \in I)$ donde $\{V_i \mid i \in I\}$ es una partición de V pero $\{A_i \mid i \in I\}$ no necesariamente es una partición de A, aunque los conjuntos A_i son disyuntos. Para dos componentes fuertemente conexas H_{i_1} y H_{i_2} escribimos $H_{i_1} \preceq H_{i_2}$ si y sólo si existe un camino dirigido de un vértice de H_{i_2} a un vértice de H_{i_1}. Nótese que si $H_{i_1} \preceq H_{i_2}$, entonces existe un camino dirigido de cada vértice de H_{i_2} a cada vértice de H_{i_1}. Puede verse fácilmente que la relación binaria $\preceq$ en el conjunto de componentes fuertemente conexas de G es un orden parcial.

Una *red dos-terminal* es un grafo $G = (V, A)$ que tiene dos vértices distinguidos s^+ y s^- en V y una función de capacidad no negativa $c : A \longrightarrow R_+$ en el conjunto de arcos A. Esta red es denotada como $\mathcal{N} = (G = (V, A), s^+, s^-, c)$, los vértices s^+ y s^- son llamados respectivamente la *entrada* y la *salida* de la red.

Una función $\varphi : A \longrightarrow \mathbb{R}$ es un *flujo factible* de N si φ satisface:

$$\forall a \in A : 0 \leq \varphi(a) \leq c(a) \text{ y } \forall v \in V - \{s^+, s^-\} : \partial\varphi(v) = 0$$

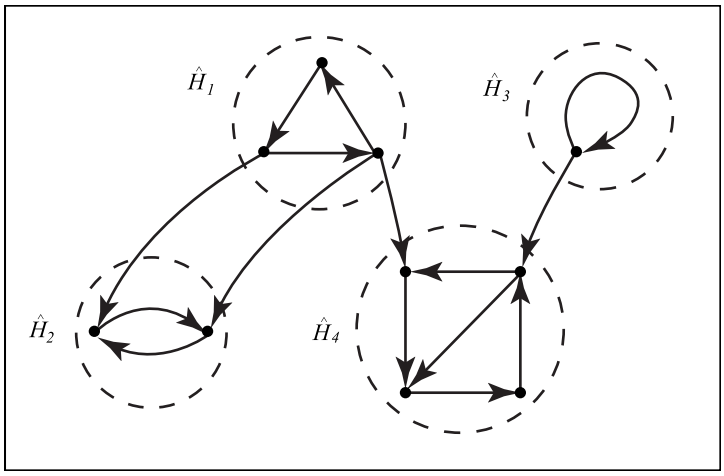

Figura 1.1: Descomposición de un grafo en componentes fuertemente conexas

donde $\partial\varphi : V \longrightarrow \mathbb{R}$ es la *función frontera* de φ definida como:

$$\partial\varphi(v) := \sum_{a\in\delta^+ v} \varphi(a) - \sum_{a\in\delta^- v} \varphi(a) \tag{1.6}$$

Suponga que tenemos un grafo $G = (V, A)$ y funciones de capacidad superior e inferior $\underline{c}, \overline{c} : A \longrightarrow \mathbb{R}$ con $\underline{c} \leq \overline{c}$. Denotemos esta red por $\mathcal{N}_0 = (G = (V, A), \underline{c}, \overline{c})$. Una *circulación factible* en $\mathcal{N}_0$ es una función $\varphi : A \longrightarrow \mathbb{R}$ tal que

$$\forall a \in A : \underline{c} \leq \varphi(a) \leq \overline{c}(a) \text{ y } \forall v \in V : \partial\varphi(v) = 0$$

en donde $\partial\varphi$ se define como en (1.6). Tenemos el siguiente resultado:

Teorema 1.3.1. *(Hoffman) Existe una circulación factible en la red $\mathcal{N}_0$ sii para cada $U \subseteq V$ tenemos:*

$$\sum_{a\in\Delta^+ U} \overline{c}(a) \geq \sum_{a\in\Delta^- U} \underline{c}(a) \tag{1.7}$$

Prueba: ver [12]

1.4. Análisis Convexo y Desigualdades lineales

Sea E un conjunto finito, para cada $x, y \in \mathbb{R}^E$ y $\lambda, \mu \in \mathbb{R}$ no negativos tales que $\lambda + \mu = 1$, el vector $\lambda x + \mu y \in \mathbb{R}^E$ es llamado una *combinación convexa* de x y y. Un conjunto $A \subseteq \mathbb{R}^E$ es llamado *convexo*, si todas las combinaciones convexas de cualesquiera dos elementos de A, están en A. Un ejemplo típico de un conjunto convexo es el conjunto solución del sistema de ecuaciones y desigualdades lineales:

$$\sum_{e \in E} a_i^{(1)}(e)x(e) = b_i^{(1)} \qquad (i \in I) \tag{1.8}$$

$$\sum_{e \in E} a_j^{(2)}(e)x(e) \leq b_j^{(2)} \qquad (j \in J) \tag{1.9}$$

donde $a_i^{(1)}(e), a_j^{(2)}(e), x(e), b_i^{(1)}, b_j^{(2)}$ $(i \in I, j \in J, e \in E)$ son constantes reales. Tal conjunto convexo es llamado un *conjunto convexo poliedral* o *poliedro convexo*.

Un conjunto $C \subseteq R^E$ es llamado un *cono convexo*, si todas las combinaciones de la forma $\lambda x + \mu y$ donde $\lambda, \mu \geq 0$ y $x, y \in C$, pertenecen a C (Una combinación de este tipo se llama *combinación no negativa* de x y y)

Un *conjunto afín* en $\mathbb{R}^E$ es una traslación de un subespacio del espacio vectorial $\mathbb{R}^E$. La *dimensión* de un conjunto afín A, es la dimensión del subespacio trasladado para obtener A. Una *linea* es un conjunto afín uno-dimensional. Un *rayo* es la traslación de un conjunto de la forma $\{\lambda x \mid \lambda \geq 0\}$ para algún vector no nulo $x \in \mathbb{R}^E$

Un conjunto convexo poliedral A se dice *acotado*, si y sólo si, no contiene ningún rayo. Un punto de un conjunto convexo poliedral A es llamado un *punto extremo* de A, si este no puede ser expresado como una combinación convexa de otros dos puntos de A. Un conjunto convexo poliedral se dice *puntiagudo*, si este tiene al menos un punto extremo. Para un poliedro convexo A descrito por las fórmulas (1.8) y (1.9) el *cono característico* de A es el conjunto solución de

$$\sum_{e \in E} a_i^{(1)}(e)x(e) = 0 \qquad (i \in I) \tag{1.10}$$

$$\sum_{e \in E} a_j^{(2)}(e)x(e) \leq 0 \qquad (j \in J) \tag{1.11}$$

el cono característico de A es denotado como $C(A)$.

Lema 1.4.1. *El poliedro convexo A descrito por las fórmulas (1.8) y (1.9) es puntiagudo, si y sólo si, A no contiene ninguna linea, o equivalentemente el sistema*

$$\sum_{e \in E} a_i^{(1)}(e)x(e) = 0 \qquad (i \in I) \tag{1.12}$$

$$\sum_{e \in E} a_j^{(2)}(e)x(e) = 0 \qquad (j \in J) \tag{1.13}$$

tiene única solución $x = 0$.

Prueba: ver [11]

1.5. Matroides

Sea E un conjunto finito, suponga que una familia $\mathcal{I} \subseteq 2^E$ satisface:

(I_0) $\varnothing \in \mathcal{I}$

(I_1) $I_1 \subseteq I_2 \in \mathcal{I} \Rightarrow I_1 \in \mathcal{I}$

(I_2) $I_1, I_2 \in \mathcal{I}, |I_1| < |I_2| \Rightarrow \exists e \in I_2 - I_1 : I_1 \cup \{e\} \in \mathcal{I}$

entonces el par $(E, \mathcal{I})$ es llamado un *matroide*. Cada $I \in \mathcal{I}$ es llamado un *conjunto independiente* del matroide $(E, \mathcal{I})$ e $\mathcal{I}$ se llama la *familia de conjuntos independientes* del matroide $(E, \mathcal{I})$.

Un conjunto independiente el cual es maximal en $\mathcal{I}$ respecto a la inclusión se llama una *base* de $(E, \mathcal{I})$.

Teorema 1.5.1. *La familia $\mathcal{B}$ de bases de un matroide $(E, \mathcal{I})$ satisface:*

(B_0) $\mathcal{B} \neq \varnothing$

(B_1) $\forall B_1, B_2 \in \mathcal{B}$, $\forall e_1 \in B_1 - B_2$, $\exists e_2 \in B_2 - B_1$:
$(B_1 - \{e_1\}) \cup \{e_2\} \in \mathcal{B}$.

Prueba: ver [10]

Un subconjunto de E el cual no es independiente es llamado un *conjunto dependiente*. Un conjunto dependiente el cual es minimal respecto a la inclusión es llamado un *circuito*.

Teorema 1.5.2. *la familia $\mathcal{C}$ de circuitos de un matroide $(E, \mathcal{I})$ satisface:*

(C_0) $\mathcal{C} \neq \varnothing$

(C_1) $\forall\, C_1, C_2 \in \mathcal{C}$, $C_1 \subseteq C_2 \Rightarrow C_1 = C_2$

(C_2) $\forall\, C_1, C_2 \in \mathcal{C}$ *con* $C_1 \neq C_2$, $\forall e \in C_1 \cap C_2$, $\exists C \in \mathcal{C}$:
$C \subseteq (C_1 \cup C_2) - \{e\}$.

Prueba: ver [10]

Definimos la *función rango* $\rho : 2^E \longrightarrow \mathbb{Z}$ del matroide $(E, \mathcal{I})$ para cada $X \subseteq E$ por

$$\rho(X) := \text{máx}\{\, |\, I|\ \mid\ I \subseteq X, I \in \mathcal{I}\} \tag{1.14}$$

Teorema 1.5.3. *La función rango ρ definida en (1.14) satisface:*

(ρ_0) $\forall X \subseteq E : 0 \leq \rho(X) \leq |X|$

(ρ_1) $X \subseteq Y \subseteq E \Rightarrow \rho(X) \leq \rho(Y)$

(ρ_2) $\forall X, Y \subseteq E : \rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$

(ρ_3) $\forall X, Y \subseteq E : X \subseteq Y \quad y \quad |X| + 1 = |Y|$
$\Rightarrow \rho(Y) = \rho(X) \ \ o \ \ \rho(Y) = \rho(X) + 1$

Prueba: ver [10]

Obsérvese que la condición ***(ρ_2)*** indica que la función rango de un matroide es una función submodular en 2^E. La propiedad ***(ρ_3)*** se llama la *propiedad de incremento unitario* de la función rango ρ.

Definimos además la *función clausura* $cl : 2^E \longrightarrow 2^E$ del matroide $(E, \mathcal{I})$ por:

$$cl(X) := \{e \in E \mid \rho(X \cup \{e\}) = \rho(X)\} \tag{1.15}$$

para cada $X \subseteq E$.

Teorema 1.5.4. *La función clausura cl satisface:*

(cl_0) $\forall X \subseteq E : X \subseteq cl(X)$

(cl_1) $\forall X, Y \subseteq E : X \subseteq cl(Y) \Rightarrow cl(X) \subseteq cl(Y)$

(cl_2) $\forall X \subseteq E, \forall e \in E : e' \in cl(X \cup \{e\}) - cl(X)$
$\Rightarrow e \in cl(X \cup \{e'\}) - cl(X)$

Prueba: Ver [10]

Capítulo 2

Polimatroides

En 1970 J. Edmonds [5] relacionó la teoría de matroides con ciertos conjuntos poliedrales a través de la función rango del matroide e introdujo la noción de polimatroide, en este capítulo estudiaremos este concepto, presentaremos algunos ejemplos e ilustraremos la relación existente con la teoría de matroides.

2.1. Principales conceptos y resultados.

Definición 2.1.1. *Sea E un conjunto finito y $\rho : 2^E \longrightarrow \mathbb{R}$, suponga que la función ρ satisface:*

(ρ_0) $\rho(\varnothing) = 0$
(ρ_1) *Si* $X \subseteq Y \subseteq E$ *entonces* $\rho(X) \leq \rho(Y)$
(ρ_2) *Para todos* $X, Y \subseteq E$, $\rho(X) + \rho(Y) \geq \rho(X \cup Y) + \rho(X \cap Y)$

entonces el par (E, ρ) es llamado un polimatroide y ρ la función rango del polimatroide.

Las condiciones *(ρ_1)* y *(ρ_2)* indican que ρ es una función submodular monótona no decreciente. Cabe destacar que ésta no necesariamente tiene la propiedad de incremento unitario como la función rango de un matroide. Decimos que el polimatroide (E, ρ) es *matroidal*, cuando ρ es la función rango de un matroide.

Definición 2.1.2. *Dado un polimatroide (E, ρ), definimos:*

$$P_{(+)}(\rho) := \{x \in \mathbb{R}^E \mid x \geq 0, \forall X \subseteq E : x(X) \leq \rho(X)\}$$

donde para cada $X \subseteq E$ y $x \in \mathbb{R}^E$,

$$x(X) = \sum_{e \in X} x(e) \quad y \quad x(\varnothing) = 0$$

$P_{(+)}(\rho)$ se llama el poliedro independiente asociado con el polimatroide (E, ρ). También definimos:

$$B(\rho) := \{x \in P_{(+)}(\rho) \mid x(E) = \rho(E)\},$$

este conjunto se llama el poliedro base asociado con el polimatroide (E, ρ).

Cada $x \in P_{(+)}(\rho)$ se denomina un *vector independiente* y cada $x \in B(\rho)$ una *base* del polimatroide (E, ρ). Se mostrará mas adelante que el poliedro base es siempre no vacío.

Estos conceptos se pueden representar gráficamente en los casos en que E tiene dos o tres elementos (Ver figura 2.1). Además, los polimatroides generalizan de cierta manera el concepto de matroide, como se verá más adelante en la sección 2.3. De aquí, los nombres de vector independiente y base.

Definición 2.1.3. *Para cada $x \in P_{(+)}(\rho)$ definimos:*

$$D(x) := \{X \subseteq E \mid x(X) = \rho(X)\}$$

Proposición 2.1.1. *$D(x)$, como se definió antes, es cerrado respecto a la unión y a la intersección.*

Prueba: Si $x \in \mathbb{R}^E$, entonces se observa fácilmente que x es modular, es decir, $x(X)+x(Y) = x(X \cup Y)+x(X \cap Y)$ para cualesquiera $X, Y \subseteq E$, luego si $X, Y \in D(x)$ con $x \in P_{(+)}(\rho)$ tenemos:

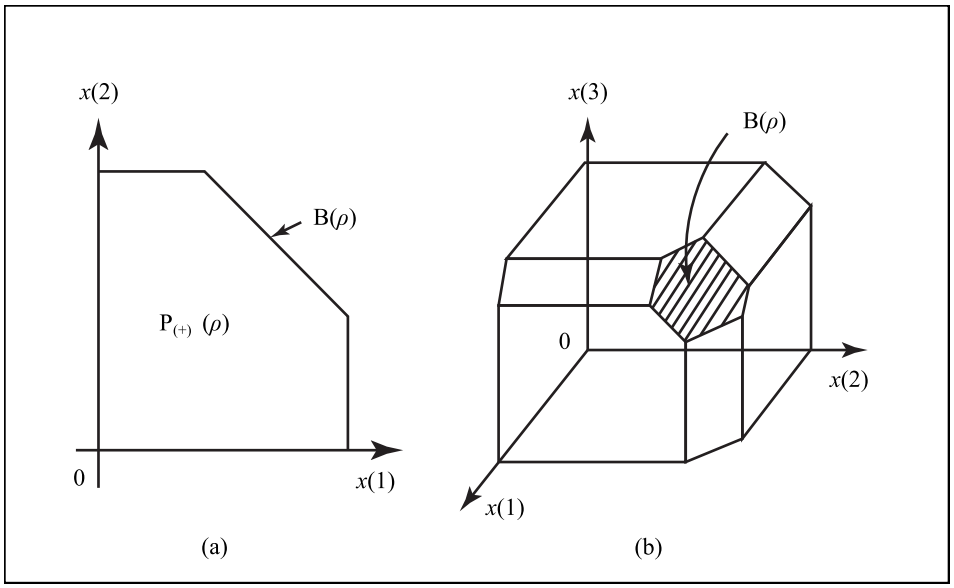

Figura 2.1: Polimatroides, casos dos y tres elementos

$$\begin{aligned} 0 &= \rho(X) - x(X) + \rho(Y) - x(Y) \geq \rho(X \cap Y) + \rho(X \cup Y) - x(X) - x(Y) \\ &= \rho(X \cap Y) - x(X \cap Y) + \rho(X \cup Y) - \rho(X \cup Y) \geq 0 \end{aligned}$$

concluyéndose así que $\rho(X \cap Y) = x(X \cap Y)$ y $\rho(X \cup Y) = \rho(X \cup Y)$, es decir, $X \cap Y \in D(x)$ y $X \cup Y \in D(x)$.

$\square$

Observamos entonces que $D(x)$ es un retículo distributivo con la unión y la intersección como sus operaciones de retículo.

Definición 2.1.4. *Al único elemento maximal de $D(x)$ lo llamamos la saturación de x y lo denotamos por $sat(x)$, es decir,*

$$sat(x) := \bigcup \{X \subseteq E \mid x(X) = \rho(X)\} \tag{2.1}$$

A partir de esta definición, obtenemos una función $sat : P_{(+)}(\rho) \longrightarrow 2^E$, la cual es llamada la *función de saturación* del polimatroide (E, ρ) y tenemos además el siguiente resultado:

Proposición 2.1.2. *Dada $x \in P_{(+)}(\rho)$, $sat(x)$ es el conjunto de las componentes saturadas de x, más precisamente,*

$$sat(x) = \{e \in E \mid \forall\, \alpha > 0, x + \alpha\chi_e \notin P_{(+)}(\rho)\} \tag{2.2}$$

donde χ_e es la función característica de $\{e\}$.

Prueba: Si $e \in sat(x)$ y $\alpha > 0$, tenemos:

$(x + \alpha\chi_e)(sat(x)) = x(sat(x)) + \alpha\chi_e(sat(x)) = \rho(sat(x)) + \alpha > \rho(sat(x))$,

luego $x + \alpha\chi_e \notin P_{(+)}(\rho)$.

Recíprocamente, si $e \in E$ es tal que para todo $n \in \mathbb{Z}_+ = \{1, 2, 3, ...\}$ existe $Y_n \subseteq E$ tal que $(x + \frac{1}{n}\chi_e)(Y_n) > \rho(Y_n)$, entonces $\frac{1}{n}\chi_e(Y_n) > 0$ para todo $n \in \mathbb{Z}_+$, con lo cual $e \in Y_n$ para todo n.

Ahora, si definimos $\mathcal{C} = \{Y_1, Y_2, ...\}$, entonces $\mathcal{C}$ es un conjunto finito, por lo tanto existe $Y \in \mathcal{C}$ tal que

$$\rho(Y) - x(Y) = min\{\rho(Y_n) - x(Y_n) : n \in \mathbb{Z}_+\}$$

asi, dado $n \in \mathbb{Z}_+$, $(x + \frac{1}{n}\chi_e)(Y_n) > \rho(Y_n)$ con lo cual:

$$\frac{1}{n} = \frac{1}{n}\,\chi_e(Y_n) > \rho(Y_n) - x(Y_n) \geq \rho(Y) - x(Y) \geq 0$$

Esto implica que $\rho(Y) = x(Y)$, es decir, existe $Y \subseteq E$ con $e \in Y$ tal que $\rho(Y) = x(Y)$. En otras palabras $e \in sat(x)$.

□

La función saturación, se puede ver como una generalización de la función clausura de un matroide (ver sección 2.3).

Definición 2.1.5. *Para un vector independiente $x \in P_{(+)}(\rho)$ y $e \in sat(x)$, definimos el conjunto*

$$D(x, e) := \{X \subseteq E \mid e \in X,\ x(X) = \rho(X)\} \tag{2.3}$$

Observamos que para $X, Y \in D(x,e)$, $e \in X \cup Y$ y $e \in X \cap Y$, luego claramente $D(x,e)$ es un subretículo de $D(x)$, por lo tanto debe tener un único elemento minimal. Esto da lugar a la siguiente definición:

Definición 2.1.6. *Para $x \in P_{(+)}(\rho)$ y $e \in sat(x)$ la dependencia de x respecto a e, denotada por $dep(x,e)$, es el único elemento minimal de $D(x,e)$, es decir,*

$$dep(x,e) = \bigcap \{X \subseteq E \mid e \in X,\, x(X) = \rho(X)\}$$

Si definimos para $e \in E - sat(x)$, $dep(x,e) = \varnothing$, entonces obtenemos una función $dep : P_{(+)}(\rho) \times E \longrightarrow 2^E$, la cual se denomina la *función dependencia* del polimatroide (E, ρ). Esta función satisface la siguiente propiedad:

Proposición 2.1.3. *Si $x \in P_{(+)}(\rho)$ y $e \in sat(x)$ entonces*

$$dep(x,e) = \{e' \in E \mid \exists\, \alpha > 0 : x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho)\}$$

además, si $e' \in dep(x,e) - \{e\}$ tenemos $x(e') > 0$

Prueba: Sea $e' \in dep(x,e)$. Si para todo $\alpha > 0$, $x + \alpha(\chi_e - \chi_{e'}) \notin P_{(+)}(\rho)$, entonces en particular para todo $n \in \mathbb{Z}_+$ existe $Y_n \subseteq E$ tal que $(x + \frac{1}{n}(\chi_e - \chi_{e'}))(Y_n) > \rho(Y_n)$, es decir,

$$x(Y_n) + \frac{1}{n}(\chi_e(Y_n) - \chi_{e'}(Y_n)) > \rho(Y_n).$$

De aquí observamos que $(\chi_e(Y_n) - \chi_{e'}(Y_n)) > 0$ para todo n, luego $e \in Y_n$ y $e' \notin Y_n$ para todo $n \in \mathbb{Z}_+$. Por lo tanto se debe tener $x(Y_n) + \frac{1}{n} > \rho(Y_n)$.

Sea $\mathcal{C} := \{Y_1, Y_2, ...\}$, nótese que $\mathcal{C}$ es un conjunto finito, luego existe $Y \in \mathcal{C}$ tal que

$$\rho(Y) - x(Y) = min\{\rho(Y_n) - x(Y_n) : n \in \mathbb{Z}_+\}$$

entonces, para cualquier $n \in \mathbb{Z}_+$:

$$\frac{1}{n} > \rho(Y_n) - x(Y_n) \geq \rho(Y) - x(Y) \geq 0$$

esto implica que $x(Y) = \rho(Y)$, donde $e \in Y$ y $e' \notin Y$. Pero esto es absurdo ya que $e' \in dep(x, e)$ implica que e' debe pertenecer a cualquier conjunto Y que cumpla $x(Y) = \rho(Y)$ con $e \in Y$.

Recíprocamente, sea $e' \in E$ para el cual existe $\alpha > 0$ con

$$x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho),$$

entonces en particular

$$(x + \alpha(\chi_e - \chi_{e'}))(dep(x, e)) \leq \rho(dep(x, e)) = x(dep(x, e))$$

con lo cual

$$x(dep(x, e)) + \alpha\big(\chi_e(dep(x, e)) - \chi_{e'}(dep(x, e))\big) \leq x(dep(x, e))$$

así, $\chi_e(dep(x, e)) - \chi_{e'}(dep(x, e)) \leq 0$, es decir, $1 - \chi_{e'}(dep(x, e)) \leq 0$, luego $\chi_{e'}(dep(x, e)) = 1$, en otras palabras; $e' \in dep(x, e)$.

Finalmente, observamos que, si $e' \in dep(x, e) - \{e\}$ entonces $x(e') > 0$, ya que $x(e') = 0$ implica

$$\rho(dep(x, e)) = x(dep(x, e)) = x(dep(x, e) - \{e\}) \leq \rho(dep(x, e) - \{e\}).$$

Es decir, $\rho(dep(x, e)) = \rho(dep(x, e) - \{e\}) = x(dep(x, e) - \{e\})$, pero esto es absurdo, ya que $dep(x, e)$ es minimal.

□

Un par ordenado (e, e') tal que $e' \in dep(x, e) - \{e\}$ es llamado un *par intercambiable* asociado con el vector independente x.

Definición 2.1.7. *Para $x \in P_{(+)}(\rho)$ y $e \in E - sat(x)$ se define*

$$\hat{c}(x, e) := max\{\alpha \in \mathbb{R} : x + \alpha\chi_e \in P_{(+)}(\rho)\}.$$

Este valor se denomina la capacidad de saturación asociada a x y e.

Proposición 2.1.4. *Sea $x \in P_{(+)}(\rho)$ y $e \in E - sat(x)$, la capacidad de saturación asociada a x y e puede ser expresada como:*

$$\hat{c}(x,e) = min\{\rho(X) - x(X) : e \in X \subseteq E\}. \tag{2.4}$$

Además, para cada α con $0 \leq \alpha \leq \hat{c}(x,e)$ tenemos $x + \alpha\chi_e \in P_{(+)}(\rho)$.

Prueba: Sea $\xi = min\{\rho(X) - x(X) : e \in X \subseteq E\}$.

Veamos que $\hat{c}(x,e) \leq \xi$. Sea $X \subseteq E$ con $e \in X$ y sea $\alpha \in \mathbb{R}$ tal que $x + \alpha\chi_e \in P_{(+)}(\rho)$, entonces $x(X) + \alpha\chi_e(X) \leq \rho(X)$, pero como $e \in X$ tenemos $x(X) + \alpha \leq \rho(X)$, así, $\alpha \leq \rho(X) - x(X)$. Esto prueba que $\hat{c}(x,e) \leq \xi$.

Veamos ahora que $\xi \leq \hat{c}(x,e)$. Sea $X \subseteq E$, tenemos dos casos:

i) Si $e \notin X$, entonces $(x + \xi\chi_e)(X) = x(X) \leq \rho(X)$.
ii) Si $e \in X$, entonces $(x+\xi\chi_e)(X) = x(X)+\xi \leq \rho(X)$, pues $\xi \leq \rho(X)-x(X)$ luego $x + \xi\chi_e \in P_{(+)}(\rho)$, con lo cual $\xi \leq \hat{c}(x,e)$.

Concluímos que $\xi = \hat{c}(x,e)$.

Finalmente, sea $\alpha \in \mathbb{R}$ con $0 \leq \alpha \leq \hat{c}(x,e)$, si $x + \alpha\chi_e \notin P_{(+)}(\rho)$, existe $Y \subseteq X$ tal que $x(Y) + \alpha > \rho(Y)$, es decir $\hat{c}(x,e) \geq \alpha > \rho(Y) - x(Y)$, pero esto es absurdo ya que por lo anterior $\hat{c}(x,e) \leq \rho(Y) - x(Y)$.

□

Definición 2.1.8. *Para $x \in P_{(+)}(\rho)$, $e \in sat(x)$ y $e' \in dep(x,e) - \{e\}$ definimos la capacidad de intercambio asociada a x, e y e' como:*

$$\tilde{c}(x,e,e') := max\{\alpha \in \mathbb{R} : x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho)\}$$

Proposición 2.1.5. *Sea $x \in P_{(+)}(\rho)$, $e \in sat(x)$ y $e' \in dep(x,e) - \{e\}$, la capacidad de intercambio asociada a x, e y e', puede ser expresada como:*

$$\tilde{c}(x,e,e') = min\{\rho(X) - x(X) : e \in X \subseteq E, e' \notin X\}. \tag{2.5}$$

Además, $x(e') \geq \tilde{c}(x,e,e')$ y para cada α tal que $0 \leq \alpha \leq \tilde{c}(x,e,e')$ se tiene que $x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho)$.

Prueba: Sea $\xi = min\{\rho(X) - x(X) : e \in X \subseteq E, e' \notin X\}$.

Veamos que $\tilde{c}(x, e, e') \leq \xi$. Sea $\alpha \in \mathbb{R}$ tal que $x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho)$ y sea $X \subseteq E$ con $e \in X$ y $e' \notin X$, tenemos que $x(X) + \alpha(\chi_e - \chi_{e'})(X) \leq \rho(X)$, es decir, $x(X) + \alpha \leq \rho(X)$ o equivalentemente, $\alpha \leq \rho(X) - x(X)$. Esto prueba que $\tilde{c}(x, e, e') \leq \xi$.

Veamos ahora que $\xi \leq \tilde{c}(x, e, e')$. Si $X \subseteq E$ tenemos $\xi \leq \rho(X) - x(X)$, es decir, $x(X) + \xi \leq \rho(X)$. Observamos dos casos:

Caso 1: $e \in X$
i) si $e' \in X$, $x(X) + \xi(\chi_e - \chi_{e'})(X) = x(X) \leq \rho(X)$.
ii) si $e' \notin X$, $x(X) + \xi(\chi_e - \chi_{e'})(X) = x(X) + \xi \leq \rho(X)$
Caso 2: $e \notin X$
i) si $e' \in X$, $x(X) + \xi(\chi_e - \chi_{e'})(X) = x(X) - \xi \leq \rho(X)$ (pues $\xi > 0$)
ii) si $e' \notin X$, $x(X) + \xi(\chi_e - \chi_{e'})(X) = x(X) \leq \rho(X)$,

de aquí, $\xi \leq \tilde{c}(x, e, e')$, luego conluímos que $\xi = \tilde{c}(x, e, e')$.

Ahora, para probar que $x(e') \geq \tilde{c}(x, e, e')$ obsérvese que

$$\begin{aligned}
x(e') < \tilde{c}(x, e, e') &\Rightarrow x(e') < \rho\big(dep(x, e) - \{e'\}\big) - x\big(dep(x, e) - \{e'\}\big) \\
&\Rightarrow x(e') < \rho\big(dep(x, e) - \{e'\}\big) - x\big(dep(x, e)\big) + x(e') \\
&\Rightarrow 0 < \rho\big(dep(x, e) - \{e'\}\big) - x\big(dep(x, e)\big) \\
&\Rightarrow 0 < \rho\big(dep(x, e) - \{e'\}\big) - \rho\big(dep(x, e)\big) \\
&\Rightarrow \rho\big(dep(x, e)\big) < \rho\big(dep(x, e) - \{e'\}\big)
\end{aligned}$$

pero esta última desigualdad no es posible, luego se concluye que $x(e') \geq \tilde{c}(x, e, e')$.

Finalmente, sea α tal que $0 \leq \alpha \leq \tilde{c}(x, e, e')$. Si $x + \alpha(\chi_e - \chi_{e'}) \notin P_{(+)}(\rho)$, entonces existe $Y \subseteq E$ tal que $x(Y) + \alpha(\chi_e - \chi_{e'})(Y) > \rho(Y)$, luego necesaria- mente $e \in Y$ y $e' \notin Y$, de aquí $\alpha > \rho(Y) - x(Y)$, por lo tanto $\tilde{c}(x, e, e') < \alpha$, lo que contradice la hipótesis. Concluímos entonces que $x + \alpha(\chi_e - \chi_{e'}) \in P_{(+)}(\rho)$.

$\square$

Ahora introduciremos algunos conceptos que generalizan la noción de matroide dual.

Definición 2.1.9. *Sea $v \in \mathbb{R}^E$ arbitrario, decimos que v es un vector dominante del polimatroide (E, ρ) si para todo $x \in P_{(+)}(\rho)$ se tiene $x \leq v$ (Aquí $x \leq v$ significa $x(e) \leq v(e)$ para todo $e \in E$).*

Dado un vector dominante v del polimatroide $P = (E, \rho)$ definimos la función: $\rho^*_{(v)} : 2^E \longrightarrow \mathbb{R}$ como:

$$\rho^*_{(v)}(X) := v(X) + \rho(E - X) - \rho(E)$$

para cada $X \subseteq E$.

Se puede ver fácilmente que un vector $v \in 2^E$ es un vector dominante de (E, ρ), si y sólo si, $\rho(\{e\}) \leq v(e)$ para todo $e \in E$.

Proposición 2.1.6. *$P^*(v) = (E, \rho^*_{(v)})$ es un polimatroide.*

Prueba: *i)* $\rho^*_{(v)}(\varnothing) = v(\varnothing) + \rho(E) - \rho(E) = v(\varnothing) = 0$ (pues $v \in \mathbb{R}^E$).

ii) Como v es modular, para todos $X, Y \subseteq E$ se tiene

$$v(Y - X) + v(X) = v(Y),$$

es decir, $v(Y - X) = v(Y) - v(X)$. Luego si $X \subseteq Y \subseteq E$ entonces

$$\begin{aligned}\rho(Y - X) + \rho(E - Y) &\geq \rho\big((Y - X) \cap (E - Y)\big) + \rho\big((Y - X) \cup (E - Y)\big)\\ &= \rho(\varnothing) + \rho(E - X)\\ &= \rho(E - X)\end{aligned}$$

con lo cual,

$$\rho(E - X) - \rho(E - Y) \leq \rho(Y - X) \leq v(Y - X) = v(Y) - v(X)$$

luego $v(X) + \rho(E - X) \leq v(Y) + \rho(E - Y)$, por lo tanto,

$$v(X) + \rho(E - X) - \rho(E) \leq v(Y) + \rho(E - Y) - \rho(E)$$

es decir, $\rho^*_{(v)}(X) \leq \rho^*_{(v)}(Y)$.

iii) Veamos que $\rho^*_{(v)}$ es submodular. Sean $X, Y \subseteq E$,

$$\rho(E - X) + \rho(E - Y) \geq \rho\big((E - X) \cup (E - Y)\big) + \rho\big((E - X) \cap (E - Y)\big)$$

$$= \rho\big(E - (X \cap Y)\big) + \rho\big(E - (X \cup Y)\big)$$

y como

$$v(X) + v(Y) = v(X \cap Y) + v(X \cup Y)$$

entonces

$$\begin{aligned} &v(X) + v(Y) + \rho(E - X) + \rho(E - Y) - 2\rho(E) \\ &\qquad \geq \rho\big(E - (X \cap Y)\big) + \rho\big(E - (X \cup Y)\big) + v(X \cap Y) + v(X \cup Y) - 2\rho(E) \end{aligned}$$

luego

$$\begin{aligned} &v(X) + \rho(E - X) - \rho(E) + v(Y) + \rho(E - Y) - \rho(E) \\ &\qquad \geq v(X \cap Y) + \rho\big(E - (X \cap Y)\big) - \rho(E) + v(X \cup Y) + \rho\big(E - (X \cup Y)\big) - \rho(E). \end{aligned}$$

En otras palabras:

$$\rho^*_{(v)}(X) + \rho^*_{(v)}(Y) \geq \rho^*_{(v)}(X \cap Y) + \rho^*_{(v)}(X \cup Y)$$

De $(i), (ii)$ y (iii) concluímos que $P^*(v)$ es un polimatroide

□

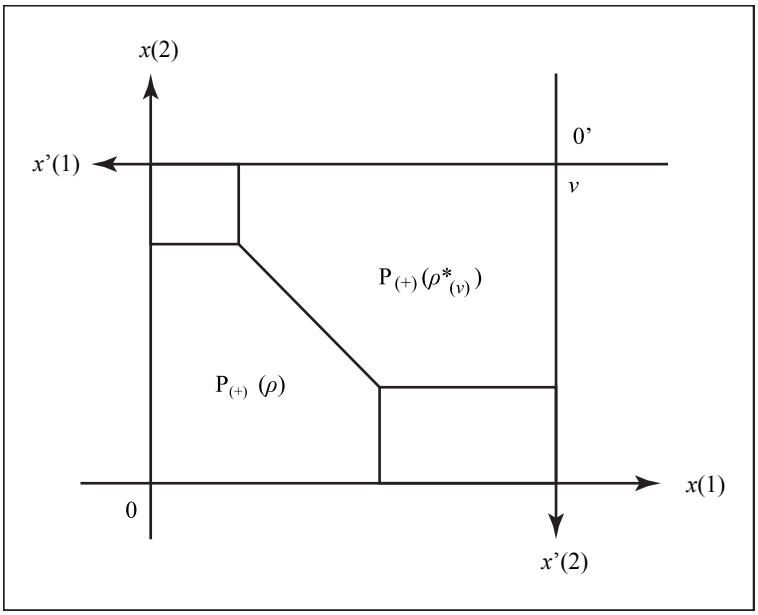

Figura 2.2: Dualidad en polimatroides, caso dos elementos

El polimatroide $P^*(v) = (E, \rho^*_{(v)})$ es llamado el *polimatroide dual* de (E, ρ) asociado al vector v (obsérvese en la figura 2.2 el caso en que E tiene dos elementos). En la siguiente sección veremos algunos ejemplos importantes de polimatroides

2.2. Ejemplos de Polimatroides:

Ejemplo 2.2.1. *(Flujos Multiterminales)*

Sea $\mathcal{N} = (G = (V, A), c, S^+, S^-)$ una red capacitada, es decir, un grafo fundamental $G = (V, A)$, junto con una función de capacidad $c : A \longrightarrow \mathbb{R}_+$, un conjunto de entradas S^+ y un conjunto de salidas S^-, aquí $S^+, S^- \subseteq V$ con $S^+ \cap S^- = \varnothing$. Una función $\varphi : A \longrightarrow \mathbb{R}$, es un *flujo factible* en $\mathcal{N}$, si:

i) Para todo $a \in A$, $0 \leq \varphi(a) \leq c(a)$.

ii) Para todo $v \in V - (S^+ \cup S^-)$ se tiene $\partial\varphi(v) = 0$.

donde $\partial\varphi : V \longrightarrow \mathbb{R}$ es la función definida para cada $v \in V$como:

$$\partial\varphi(v) = \sum_{a \in \delta^+ v} \varphi(a) - \sum_{a \in \delta^- v} \varphi(a)$$

Si $(\partial\varphi)^{S^+}$ denota la restricción de $\partial\varphi$ al conjunto S^+, entonces el conjunto $\{(\partial\varphi)^{S^+} : \varphi \text{ es un flujo factible en } \mathcal{N}\}$ es el poliedro independiente del polimatroide $P = (S^+, \rho)$, donde ρ es la función definida como $\rho(v) = \sum_{a \in \delta^+ v} c(a)$. Similarmente, tenemos un polimatroide en S^-.

Ejemplo 2.2.2. *(Polimatroides Lineales)*

Sea A una matriz y suponga que E es el conjunto de índices de sus columnas. Sea $\{F_1, F_2, ..., F_n\}$ una partición de E, es decir, $F_i \subseteq E$ para todo $i = 1, 2, ..., n$, $F_i \cap F_j = \varnothing$ para $i \neq j$ y $\bigcup F_i = E$.

Sea $\widetilde{E} = \{1, 2, ..., n\}$, para cada $X \subseteq \widetilde{E}$ definimos $A^X := \bigcup\{F_i : i \in X\}$. Definimos además la función $\rho : 2^{\widetilde{E}} \longrightarrow \mathbb{R}$ como $\rho(X) := rank(A^X)$, para todo $X \subseteq \widetilde{E}$ tenemos entonces que $\rho(\varnothing) = 0$, además, si $X \subseteq Y \subseteq \widetilde{E}$ entonces $rank(A^X) \leq rank(A^Y)$.

Por otro lado, para $X, Y \subseteq \widetilde{E}$,

$$rank(A^X) + rank(A^Y) \geq rank(A^{X \cup Y}) + rank(A^{X \cap Y}),$$

luego $(\widetilde{E}, \rho)$ es un polimatroide, llamado *polimatroide lineal.*

Ejemplo 2.2.3. *(Juegos Convexos)*

Sea $E = \{1, 2, ..., n\}$, asumimos que E es un conjunto de n personas, (o jugadores) Una *función característica* v, es una función no negativa definida en el conjunto de coaliciones (subconjuntos) de E con $v(\varnothing) = 0$. El par (E, v) es llamado un *juego de función característica* [13]

Un juego de función característica (E, v) es llamado un *juego convexo* si la función característica v es supermodular, es decir, para $X, Y \subseteq E$

$$v(X) + v(Y) \leq v(X \cup Y) + v(X \cap Y).$$

Definimos el *centro* del juego, como el conjunto

$$\mathcal{C} = \{x \in \mathbb{R}^E : \forall X \subseteq E,\, x(X) \geq v(X),\, X(E) = v(E)\}$$

y la función $\tilde{v} : 2^E \longrightarrow \mathbb{R}$ como

$$\tilde{v}(E - X) := v(E) - v(X) \quad \text{para } X \subseteq E$$

o equivalentemente,

$$\tilde{v}(Y) = v(E) - v(E - Y) \quad \text{para } Y \subseteq E,$$

$\tilde{v}$ es la función rango de un polimatrode en E y el centro coincide con el poliedro base $B(\tilde{v})$ del polimatroide. En efecto:

i) $\tilde{v}(\varnothing) = v(E) - v(E) = 0$

ii) Dados $X_1, X_2 \subseteq E$ tales que $X_1 \subseteq X_2$, entonces por la supermodularidad de v tenemos:

$$v(X_1) + v(X_2 - X_1) \leq v(X_1 \cap (X_2 - X_1)) + v(X_1 \cup (X_2 - X_1))$$

por lo tanto, $v(X_1) + v(X_2 - X_1) \leq v(X_1 \cup X_2) = v(X_2)$, pero $v(X_2 - X_1) \geq 0$, luego $v(X_1) \leq v(X_2)$.

Ahora, si $X \subseteq Y \subseteq E$, tenemos $E - Y \subseteq E - Y$ luego $v(E - Y) \leq v(E - X)$, con lo cual $v(E) - v(E - X) \leq v(E) - v(E - Y)$, esto es, $\tilde{v}(X) \leq \tilde{v}(Y)$.

iii) Sean $X, Y \subseteq E$, utilizando la supermodularidad de v tenemos:

$$v(E - X) + v(E - Y) \leq v\big((E - X) \cup (E - Y)\big) + v\big((E - X) \cap (E - Y)\big).$$

por lo tanto

$$\begin{aligned}
\tilde{v}(X) + \tilde{v}(Y) &= v(E) - v(E - X) + v(E) - v(E - Y) \\
&\geq v(E) - v\big((E - X) \cap (E - Y)\big)) + v(E) - v\big((E - X) \cup (E - Y)\big) \\
&\geq v(E) - v\big(E - (X \cup Y)\big) + v(E) - v\big(E - (X \cap Y)\big) \\
&= \tilde{v}(X \cup Y) + \tilde{v}(X \cap Y)
\end{aligned}$$

Finalmente, para ver que el centro $\mathcal{C}$ coincide con el poliedro base $B(\tilde{v})$, observamos que si $x \in \mathcal{C}$ entonces $x \geq 0$, y si $Y \subseteq E$, $x(E - Y) \geq v(E - Y)$, pero

$$\begin{aligned}
& x(E - Y) \geq v(E - Y) \\
& \quad \Leftrightarrow\ 0 \leq -v(E - Y) + x(E - Y) \\
& \quad \Leftrightarrow\ v(E) \leq v(E) - v(E - Y) + x(E - Y) \\
& \quad \Leftrightarrow\ x(E) \leq v(E) - v(E - Y) + x(E - Y) \quad (\text{pues } x(E) = v(E)) \\
& \quad \Leftrightarrow\ x(Y) + x(E - Y) \leq v(E) - v(E - Y) + x(E - Y) \quad (x \text{ es modular}) \\
& \quad \Leftrightarrow\ x(Y) \leq \tilde{v}(Y)
\end{aligned}$$

de aquí, $x \in B(\tilde{v})$, es decir, $\mathcal{C} \subseteq B(\tilde{v})$. Para probar que $B(\tilde{v}) \subseteq \mathcal{C}$ se utiliza exactamente el mismo razonamiento.

2.3. Relación con matroides:

Los polimatroides se pueden ver como una "versión continua"de los matroides en el sentido en que si cambiamos $\mathbb{R}$ por $\mathbb{Z}$ en las definiciones anteriores, y si ρ es la función rango de un matroide, entonces habrá una correspondencia entre los elementos de $P_{(+)}(\rho)$ y los conjuntos independientes del matroide asociado, así mismo, $B(\rho)$ se identificará con las bases del matroide a través de esta correspondencia.

En efecto, supongamos que $(E, \mathcal{I})$ es el matroide asociado a la función rango ρ. Definamos:

$$P_{M(+)}(\rho) := \{x \in \mathbb{Z}^E \mid x \geq 0, \forall X \subseteq E : x(X) \leq \rho(X)\} \text{ y}$$

$$B_M(\rho) := \{x \in P_{M(+)}(\rho) \mid x(E) = \rho(E)\}$$

Tenemos el siguiente resultado:

Proposición 2.3.1.

i) $I \in \mathcal{I}$, *si y sólo si,* $\chi_I \in P_{M(+)}(\rho)$.
ii) I *es base de* $(E, \mathcal{I})$, *si y sólo si,* $\chi_I \in B_M(\rho)$

Prueba: i) Si $I \in \mathcal{I}$ y $X \subseteq E$, entonces

$$\chi_I(X) = |I \cap X| = \rho(I \cap X) \leq \rho(X),$$

es decir, $\chi_I \in P_{M(+)}(\rho)$.

Recíprocamente, si $I \subseteq X$ es tal que $\chi_I \in P_{M(+)}(\rho)$ entonces

$$\chi_I(X) = |I \cap X| \leq \rho(X).$$

En particular, si $X = I$, entonces $|I \cap I| = |I| \leq \rho(I)$, pero por propiedades de la función rango de un matroide (teorema 1.5.3), $\rho(I) \leq |I|$, luego $\rho(I) = |I|$, con lo cual I será un elemento independiente del matroide.

ii) Si I es base de $(E, \mathcal{I})$, entonces $\rho(E) = |I| = \chi_I(E)$, es decir $\chi_I \in B_M(\rho)$.

Recíprocamente, si $\chi_I \in B_M(\rho)$, entonces $\rho(E) = \chi_I(E) = |I|$, y como por (i), I es independiente, entonces I debe ser una base.

$\square$

Para cualquier $x \in P_{M(+)}(\rho)$ y $a \in E$ se debe tener

$$0 \leq x(a) \leq \rho(\{a\}) \leq |a| = 1,$$

entonces $x = \chi_I$ para algún $I \subseteq E$. Esto, junto con lo anterior, nos permite observar que hay una correspondencia biyectiva, entre los elementos de $P_{M(+)}(\rho)$ y los elementos de $\mathcal{I}$, en donde $B_M(\rho)$ se identifica con las bases del matroide.
Así mismo, para $x \in P_{M(+)}(\rho)$ tenemos la saturación de x como:

$$sat(x) = \bigcup \{X \subseteq E \mid x(X) = \rho(X)\}$$

y la función de saturación $sat : P_{M(+)}(\rho) \longrightarrow 2^E$.

Obsérvese que $P_{M(+)}(\rho) \subseteq P_{(+)}(\rho)$, luego podemos ver de la proposición 2.1.1, que el conjunto $D(x) = \{X \subseteq E \mid x(X) = \rho(X)\}$ es cerrado respecto a la unión y a la intersección , luego $sat(x)$ es el único elemento maximal de $D(x)$. Tenemos además el siguiente resultado:

Proposición 2.3.2. *Para $x \in P_{M(+)}(\rho)$ se tiene que*

$$sat(x) = \{e \in E \mid x + \chi_e \notin P_{M(+)}(\rho)\} \tag{2.6}$$

Prueba: Sea $e \in sat(x)$,

$$(x + \chi_e)(sat(x)) = x(sat(x)) + \chi_e(sat(x)) = \rho(sat(x)) + 1 > \rho(sat(x)),$$

luego $x + \chi_e \notin P_{M(+)}(\rho)$.

Recíprocamente, supongamos que $x + \chi_e \notin P_{M(+)}(\rho)$, existe $Y \subseteq E$ tal que $x(Y) + \chi_e(Y) > \rho(Y)$, luego se debe tener que $e \in Y$. De aquí, $X(Y) + 1 > \rho(Y)$, es

decir $\rho(Y) - X(Y) < 1$.

Pero como ρ y x toman únicamente valores enteros no negativos, entonces se debe tener $\rho(Y) = X(Y)$, y dado que $e \in Y$, entonces $e \in sat(x)$.

□

La siguiente proposición muestra que la función de saturación en un polimatroide es una generalización de la función clausura de un matroide.

Proposición 2.3.3. *Sea $\chi_I \in P_{M(+)}(\rho)$ $(I \in \mathcal{I})$, se tiene:*

$$sat(\chi_I) = Cl(I)$$

en donde $Cl(I)$ denota la clausura de I.

Prueba: Veamos que $sat(\chi_I) \subseteq Cl(I)$. Si $e \in sat(\chi_I)$, entonces por la proposición anterior, $\chi_I + \chi_e \notin P_{M(+)}(\rho)$. Si $e \in I$, entonces $e \in Cl(I)$. Supongamos entonces que $e \notin I$, tenemos $\chi_{I\cup\{e\}} = \chi_I + \chi_e \notin P_{M(+)}(\rho)$, esto quiere decir, por la proposición 2.3.1, que $I \cup \{e\}$ no es independiente, luego $\rho(I \cup \{e\}) = |I| = \rho(I)$. Esto es, $e \in Cl(I)$.

Veamos ahora que $Cl(I) \subseteq sat(\chi_I)$. Sea $e \in Cl(I)$, tenemos

$$\rho(I \cup \{e\}) = |I| = \rho(I).$$

Si $e \in I$, como $\chi_I(I) = |I| = \rho(I)$, entonces $e \in sat(\chi_I)$, si $e \notin I$, entonces $I \cup \{e\}$ no es independiente, es decir, $\chi_{I\cup\{e\}} = \chi_I + \chi_e \notin P_{M(+)}(\rho)$.

Debido a la proposición 2.3.2, concluímos que $e \in sat(\chi_I)$.

□

Para $x \in P_{M(+)}(\rho)$ y $e \in sat(x)$, recordemos que la dependencia $dep(x, e)$ de x respecto a e está definida como

$$dep(x, e) = \bigcap \{X \subseteq E \mid e \in X,\ x(X) = \rho(X)\}$$

Recordamos además, que $dep(x,e)$ es el único elemento minimal del conjunto $D(x,e) = \{X \subseteq E \mid e \in X,\, x(X) = \rho(X)\}$

Proposición 2.3.4. *Para $\chi_I \in P_{M(+)}(\rho)$ $(I \in \mathcal{I})$ y $e \in Cl(I) - I$, $dep\,(\chi_I, e)$ es el único circuito contenido en $I \cup \{e\}$.*

Prueba: Veamos primero que $dep\,(\chi_I, e) \subseteq I \cup \{e\}$. Sea $e' \in dep\,(\chi_I, e) - \{e\}$, por la proposición 2.1.3, $\chi_I(e') > 0$, es decir, $\chi_I(e') = 1$, o equivalentemente $e' \in I$. Tenemos entonces que $dep\,(\chi_I, e) - \{e\} \subseteq I$, pero como $e \in dep\,(\chi_I, e)$, concluimos que $dep\,(\chi_I, e) \subseteq I \cup \{e\}$.

Por otro lado:

$$\begin{aligned}\rho\big(dep\,(\chi_I, e)\big) &= \chi_I\ \big(dep\,(\chi_I, e)\big) = \mid I \cap dep\,(\chi_I, e) \mid \\ &= \mid dep\,(\chi_I, e) - \{e\} \mid = \mid dep\,(\chi_I, e) \mid -1,\end{aligned}$$

esto prueba que $dep\,(\chi_I, e)$ es un conjunto dependiente.

Ahora, sea C un circuito tal que $C \subseteq dep\,(\chi_I, e)$, si $e \notin C$ tenemos $C \subseteq dep\,(\chi_I, e) - \{e\} \subseteq I$, luego $\chi_I(C) = |I \cap C| = |C| > \rho(C)$, lo cual es absurdo ya que $\chi_I \in P_{M(+)}(\rho)$. Se debe tener entonces que $e \in C$. Además, como $C - \{e\} \subseteq dep\,(\chi_I, e) - \{e\} \subseteq I$ y C es dependiente minimal, entonces

$$\chi_I(C) = \mid I \cap C| = \mid C - \{e\}| = \rho\big(C - \{e\}\big) = \rho(C),$$

luego $dep\,(\chi_I, e) \subseteq C$, por lo tanto, $dep\,(\chi_I, e) = C$, concluyéndose asi que $dep\,(\chi_I, e)$ es un circuito.

Finalmente, sea C un circuito contenido en $I \cup \{e\}$, $e \in C$ pues de lo contrario se tendría $C \subseteq I$ y por tanto $C \in \mathcal{I}$, lo cual es absurdo. Si $C \neq dep\,(\chi_I, e)$ entonces por (C_2) del teorema (1.5.2), existe C_0 circuito tal que $C_0 \subseteq \big(C \cup dep\,(\chi_I, e)\big) - \{e\}$, pero como $C \subseteq I \cup \{e\}$ y $dep\,(\chi_I, e) \subseteq I \cup \{e\}$, entonces necesariamente $\big(C \cup dep\,(\chi_I, e)\big) - \{e\} \subseteq I$, luego C_0 es un conjunto independiente, lo cual es contradictorio, es decir, se debe tener $C = dep\,(\chi_I, e)$. Con esto se concluye la demostración. $\square$

La anterior proposición muestra que dado $I \in \mathcal{I}$ y $e \in Cl(I) - I$, existe un único circuito contenido en $I \cup \{e\}$. En teoría de matroides a este circuito se le suele llamar el *circuito fundamental* respecto a I y e.

En el siguiente ejemplo se ilustran algunos de los resultados de esta sección.

Ejemplo 2.3.1. *Consideremos el matroide* $(E, \mathcal{I})$ *donde* $E = \{a, b, c\}$ *e* $\mathcal{I} = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}\}$.

Tenemos:

$\rho(\varnothing) = 0$, $\rho(\{a\}) = \rho(\{b\}) = \rho(\{c\}) = 1$, $\rho(\{a, b\}) = \rho(\{a, c\}) = 2$,
$\rho(\{b, c\}) = 1$ y $\rho(\{a, b, c\}) = 2$.

$P_{(+)}(\rho) = \{0, \chi_{\{a\}}, \chi_{\{b\}}, \chi_{\{c\}}, \chi_{\{a,b\}}, \chi_{\{a,c\}}\}$.

Obsérvese por ejemplo que $\chi_{\{b,c\}} \notin P_{(+)}(\rho)$, pues
$\chi_{\{b,c\}}(\{b, c\}) = 2 > 1 = \rho(\{b, c\})$.

Además:

$sat(\chi_{\{a\}}) = \{a\} = Cl(\{a\})$
$sat(\chi_{\{b\}}) = \{b, c\} = Cl(\{b\})$
$sat(\chi_{\{c\}}) = \{b, c\} = Cl(\{c\})$
$sat(\chi_{\{a,b\}}) = \{a, b, c\} = Cl(\{a, c\})$
$sat(\chi_{\{a,c\}}) = \{a, b, c\} = Cl(\{a, c\})$

Capítulo 3

Sistemas Submodulares

El concepto de polimatroide puede presentarse de manera más general, suprimiendo algunas condiciones sobre la función rango y sobre el conjunto en el cual actúa dicha función, esto se hace a través de la noción de sistema submodular. En este capítulo estudiaremos detalladamente este concepto, no sólo presentaremos definiciones y resultados análogos a los de polimatroides, sino que estudiaremos algunos otros aspectos como reducciones y contracciones, estos últimos tienen que ver con modificaciones o alteraciones aplicadas a un sitema submodular, obteniendo nuevos sistemas submodulares. En este capítulo también se desarrollan algunos ejemplos de sistemas submodulares que se pueden encontrar dentro de la Teoría de Redes, Matrices, y otros que se derivan de las funciones cóncavas.

3.1. Principales conceptos y resultados

Sea E un conjunto finito no vacío y $\mathcal{D} \subseteq 2^E$ una colección de subconjuntos de E cerrado para la unión y la intersección, es decir, para $X, Y \in \mathcal{D}$ se tiene $X \cup Y \in \mathcal{D}$ y $X \cap Y \in \mathcal{D}$. Claramente $\mathcal{D}$ es un retículo distributivo siendo la unión y la intersección sus operaciones de retículo. En la siguiente definición obtenemos de manera más general la noción de submodularidad.

Definición 3.1.1. *Sea $\mathcal{D} \subseteq 2^E$ cerrado para la unión y la intersección, y sea f :*

$\mathcal{D} \longrightarrow \mathbb{R}$. *Decimos que f es una función submodular en $\mathcal{D}$ si para todo par $X, Y \in \mathcal{D}$,*

$$f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y).$$

Tenemos el siguiente lema fundamental sobre funciones submodulares en este tipo de retículos:

Lema 3.1.1. *Sea f una función submodular en un conjunto $\mathcal{D} \subseteq 2^E$ cerrado para la unión e intersección. El conjunto de minimizadores de f*

$$\mathcal{D}_0 := \big\{X \in \mathcal{D} \mid f(X) = \text{mín}\{f(Y) \mid Y \in \mathcal{D}\}\big\}$$

es un subretículo de $\mathcal{D}$

Prueba: Para cada $X, Y \in \mathcal{D}_0$ tenemos $f(X) + f(Y) \geq f(X \cup Y) + f(X \cap Y)$, pero como $f(X) = f(Y)$ entonces $2f(X) \geq f(X \cup Y) + f(X \cap Y)$.

Por otro lado, siendo $f(X) \leq f(X \cup Y)$ y $f(X) \leq f(X \cap Y)$ concluimos que

$$2f(X) = f(X \cup Y) + f(X \cap Y). \tag{3.1}$$

Ahora, si suponemos $f(X \cup Y) > f(X \cap Y)$, entonces $f(X \cup Y) + f(X \cap Y) > 2f(X \cap Y)$, es decir, $2f(X) > 2f(X \cap Y)$, pero esto contradice la minimalidad de $f(X)$. Igualmente, si $f(X \cup Y) < f(X \cap Y)$ entonces $2f(X) > 2f(X \cup Y)$, lo cual también es contradictorio. Se debe tener entonces que $f(X \cup Y) = f(X \cap Y)$.

De la ecuación (3.1) se concluye que $f(X) = f(X \cup Y) = f(X \cap Y)$

□

Definición 3.1.2. *Sea $\mathcal{D} \subseteq 2^E$ cerrado para la unión e intersección tal que $\varnothing, E \in \mathcal{D}$, y sea $f : \mathcal{D} \longrightarrow \mathbb{R}$ una función submodular en $\mathcal{D}$ que satisface $f(\varnothing) = 0$, entonces el par $(\mathcal{D}, f)$ se denomina sistema submodular en E. Se dice además que f es la función rango de $(\mathcal{D}, f)$ y que $f(E)$ es el rango de $(\mathcal{D}, f)$.*

Obsérvese que todo polimatroide (E, ρ) es un sistema submodular, luego se pueden tratar de extender los conceptos relacionados con polimatroides a sitemas submo-

dulares, esto se hace mediante algunas de las definiciones posteriores de ese capítulo.

Definición 3.1.3. *Sea $(\mathcal{D}, f)$ un sistema submodular en E, el conjunto*

$$P(f) := \{x \in \mathbb{R}^E \mid \forall X \in \mathcal{D} : x(X) \leq f(X)\}$$

es llamado el poliedro submodular asociado al sistema submodular $(\mathcal{D}, f)$, y el conjunto

$$B(f) := \{x \in P(f) \mid x(E) = f(E)\}$$

se llama el poliedro base asociado a $(\mathcal{D}, f)$. Un elemento del poliedro base es llamado una base de $(\mathcal{D}, f)$ y un elemento del poliedro submodular es llamado una subbase de $(\mathcal{D}, f)$.

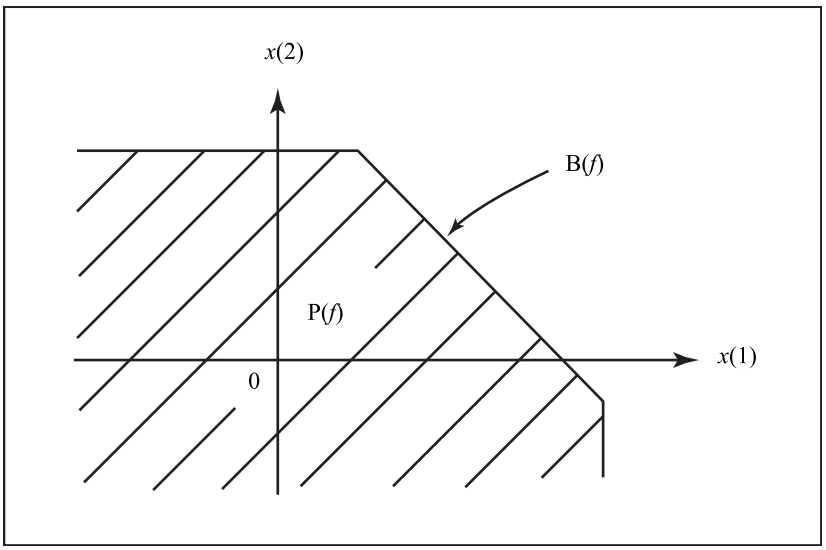

Figura 3.1: Poliedro submodular $P(f)$ y poliedro base $B(f)$

Lema 3.1.2. *Para cada subbase $x \in P(f)$ el conjunto*

$$\mathcal{D}(x) := \{X \in \mathcal{D} \mid x(X) = f(X)\}$$

es un subretículo de $\mathcal{D}$.

Prueba: La función $f - x : \mathcal{D} \longrightarrow \mathbb{R}$ es una función submodular no negativa en $\mathcal{D}$, pues $(f - x)(\varnothing) = 0$ y dados $X, Y \in \mathcal{D}$ se tiene:

$$\begin{aligned}(f - x)(X) + (f - x)(Y) &= f(X) - x(X) + f(Y) - x(Y)\\ &\geq f(X \cup Y) - x(X) + f(X \cap Y) - x(Y)\\ &= f(X \cup Y) - x(X \cup Y) + f(X \cap Y) - x(X \cap Y)\\ &= (f - x)(X \cup Y) + (f - x)(X \cap Y)\end{aligned}$$

Ahora, para $X \in \mathcal{D}$

$$X \in \mathcal{D}(x) \Leftrightarrow x(X) = f(X) \Leftrightarrow f(X) - x(X) = 0$$

$$\Leftrightarrow X \in \big\{X_0 \in \mathcal{D} \mid f(X_0) = \text{mín}\{(f - x)(Y) \mid Y \in \mathcal{D}\}\big\}$$

Observamos entonces que $\mathcal{D}(x)$ es el conjunto de minimizadores de la función $f - x$. Aplicando el lema 3.1.1, concluímos que $\mathcal{D}(x)$ es un subretículo de $\mathcal{D}$.

□

Definición 3.1.4. *El único elemento minimal de $\mathcal{D}(x)$ se llama la saturación de x y es denotado por $sat(x)$, la función $sat : P(f) \longrightarrow 2^E$ se llama la función de saturación del sistema submodular $(\mathcal{D}, f)$.*

Observando la demostración de la proposición 2.1.2, se puede concluir además que

$$sat(x) = \{e \in E \mid \forall\, \alpha > 0, x + \alpha\chi_e \notin P(f)\} \tag{3.2}$$

Lema 3.1.3. *Para cada subbase $x \in P(f)$ y $e \in sat(x)$, el conjunto*

$$\mathcal{D}(x, e) := \{X \in \mathcal{D} \mid e \in X, x(X) = f(X)\}$$

es un subretículo de $\mathcal{D}$.

Prueba: Sea $\mathcal{D}(e) := \{X \in \mathcal{D} \mid e \in X\}$. Nótese que $\mathcal{D}(e)$ es un subretículo de $\mathcal{D}$ y para $X \in \mathcal{D}(e)$,

$$X \in \mathcal{D}(x, e) \Leftrightarrow x(X) = f(X) \Leftrightarrow f(X) - x(X) = 0$$

$$\Leftrightarrow X \in \big\{X_0 \in \mathcal{D} \mid f(X_0) = \text{mín}\{(f - x)(Y) \mid Y \in \mathcal{D}(e)\}\big\}.$$

Luego $\mathcal{D}(x,e)$ es el conjunto de minimizadores de la función submodular $f-x$ en el retículo distributivo $\mathcal{D}(e)$, y asi, por el lema 3.1.1 se concluye lo deseado.

$\square$

El único elemento minimal de $\mathcal{D}(x,e)$ se llama la *dependencia* de x respecto a e y es denotado como $dep(x,e)$. Si para cada $e \in E - sat(x)$ definimos $dep(x,e) = \varnothing$, entonces podemos definir una función $dep : P(f) \times E \longrightarrow 2^E$ la cual es llamada la *función de dependencia* de $(\mathcal{D}, f)$.

Definición 3.1.5. *Para cada $x \in P(f)$ y $e \in E - sat(x)$ definimos la capacidad de saturación asociada a x y e como:*

$$\hat{c}(x,e) := min\{f(X) - x(X) \mid e \in X \in \mathcal{D}\}$$

Definimos además para cada $x \in P(f)$, $e \in sat(x)$ y $e' \in dep(x,e) - \{e\}$ la capacidad de intercambio asociada con x, e y e' como:

$$\tilde{c}(x,e,e') := min\{f(X) - x(X) \mid e \in X \in \mathcal{D}, e' \notin X\}$$

Proposición 3.1.4. *Sea $x \in P(f)$*

i) Para $e \in E - sat(x)$ y α no negativo se tiene $x + \alpha\chi_e \in P(f)$ sii $0 \leq \alpha \leq \hat{c}(x,e)$

ii) Para $e \in sat(x)$, $e' \in dep(x,e) - \{e\}$ y α no negativo se tiene $x + \alpha(\chi_e - \chi_{e'}) \in P(f)$ sii $0 \leq \alpha \leq \tilde{c}(x,e,e')$.

Prueba: i) Si $x + \alpha\chi_e \in P(f)$, entonces $x(X) + \alpha\chi_e(X) \leq f(X)$ para todo $X \in \mathcal{D}$. Si $e \in X$ entonces $x(X) + \alpha\chi_e(X) = x(X) + \alpha \leq f(X)$, así, $\alpha \leq f(X) - x(X)$ para todo $X \in \mathcal{D}$ con $e \in X$, es decir, $\alpha \leq \hat{c}(x,e)$.

Recíprocamente, suponga que $0 \leq \alpha \leq \hat{c}(x,e)$, si $x + \alpha\chi_e \notin P(f)$, existe $Y \in \mathcal{D}$ tal que $x(Y) + \alpha > f(Y)$, es decir, $\hat{c}(x,e) \geq \alpha > f(Y) - x(Y)$ con $e \in Y$, lo cual es absurdo ya que $\hat{c}(x,e) \leq f(Y) - x(Y)$.

ii) Si $x + \alpha(\chi_e - \chi_{e'}) \in P(f)$ entonces para cualquier $X \in \mathcal{D}$ tal que $e \in X$ y $e' \notin X$ tenemos $x(X) + \alpha(\chi_e(X) - \chi_{e'}(X)) \leq f(X)$, es decir, $\alpha \leq f(X) - x(X)$, con lo que se concluye que $\alpha \leq \tilde{c}(x, e, e')$.

Recíprocamente, sea α tal que $0 \leq \alpha \leq \tilde{c}(x, e, e')$ y $X \in \mathcal{D}$. Tenemos dos casos:

Caso 1) $e \in X$.
Si $e' \in X$, entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) \leq f(X)$.
Si $e' \notin X$, entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) + \alpha \leq f(X)$ (pues $\alpha \leq f(X) - x(X)$).

Caso 2) $e \notin X$.
Si $e' \in X$, entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) - \alpha \leq f(X)$.
Si $e' \notin X$, entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) \leq f(X)$.

Concluímos que $x + \alpha(\chi_e - \chi_{e'}) \in P(f)$

□

Con el fin de probar que el poliedro base $B(f)$ es siempre no vacío definiremos un orden parcial en $\mathbb{R}^E$: Dados $x, y \in \mathbb{R}^E$, $x \leq y$ sii $x(e) \leq y(e)$ para todo $e \in E$. El siguiente teorema establece entonces esta propiedad del poliedro base.

Teorema 3.1.5. *Sea $(\mathcal{D}, f)$ un sistema submodular en E, El poliedro base $B(f)$ es el conjunto de subbases maximales de $(\mathcal{D}, f)$.*

Prueba: Denotemos por $B'(f)$ al conjunto de subbases maximales de $(\mathcal{D}, f)$. Sean $x \in B(f)$ y $y \in P(f)$ tales que $x \leq y$. Si $y(e) > x(e)$ para algún $e \in E$ entonces $y(E) > x(E) = f(E)$ lo cual es absurdo. Tenemos entonces que x es maximal en $(\mathcal{D}, f)$ y consecuentemente, $B(f) \subseteq B'(f)$.

Recíprocamente suponga que $x \in B'(f)$, si $e \in E$ y $\alpha > 0$ entonces $x < x + \alpha\chi_e$, luego por la maximalidad de x se debe tener $x + \alpha\chi_e \notin P(f)$, así por la ecuación (3.2), $e \in sat(x)$. De aquí, $E = sat(x)$ o en otras palabras $x(E) = f(E)$, concluyéndose así que $x \in B(f)$, es decir, $B'(f) \subseteq B(f)$

□

A partir del teorema anterior se observa que para cada subbase x de $(\mathcal{D}, f)$ existe una base y de $(\mathcal{D}, f)$ tal que $x \leq y$, de aquí $B(f) \neq \varnothing$.

Definición 3.1.6. *Sea $\mathcal{D} \subseteq 2^E$ cerrado para la unión y la intersección, una función $g : \mathcal{D} \longrightarrow \mathbb{R}$ se llama una función supermodular en $\mathcal{D}$ si $-g$ es una función submodular en $\mathcal{D}$, es decir, para todos $X, Y \in \mathcal{D}$ se tiene:*

$$g(X) + g(Y) \leq g(X \cup Y) + g(X \cap Y).$$

si $\varnothing, E \in \mathcal{D}$ y $g(\varnothing) = 0$, el par $(\mathcal{D}, g)$ se llama un sistema supermodular en E. Una función la cual es submodular y supermodular en $\mathcal{D}$ se llama un función modular en $\mathcal{D}$.

Definición 3.1.7. *Dado un sistema supermodular $(\mathcal{D}, g)$ en E, los conjuntos*

$$P(g) := \{x \in \mathbb{R}^E \mid \forall X \in \mathcal{D} : x(X) \geq g(X)\}$$

$$B(g) := \{x \in P(f) \mid x(E) = g(E)\}$$

se llaman respectivamente el poliedro supermodular y el poliedro base asociados al sistema supermodular $(\mathcal{D}, g)$. Un vector en $P(g)$ es llamado una superbase y un vector en $B(g)$ es llamado una base de $(\mathcal{D}, g)$.

Proposición 3.1.6. *Sea $(\mathcal{D}, f)$ un sistema submodular en E, si definimos el conjunto $\overline{\mathcal{D}} := \{E - X \mid X \in \mathcal{D}\}$ y la función $f^{\#} : \overline{\mathcal{D}} \longrightarrow \mathbb{R}$ como $f^{\#}(E - X) = f(E) - f(X)$, $(X \in \mathcal{D})$ entonces el par $(\mathcal{D}, f^{\#})$ es un sistema supermodular en E.*

Prueba: Dados $X, Y \in \mathcal{D}$, como $(E - X) \cap (E - Y) = E - (X \cap Y)$ y $(E - X) \cup (E - Y) = E - (X \cup Y)$ entonces $\mathcal{D}$ es cerrado para uniones e intersecciones.

Por otro lado, $f^{\#}(\varnothing) = f^{\#}(E - E) = f(E) - f(E) = 0$ y para $Y \in \overline{\mathcal{D}}$, $f^{\#}(Y) = f(E) - f(E - Y)$, luego para $Y_1, Y_2 \in \mathcal{D}$, si Y^c denota al conjunto $E - Y$ para cada $Y \subseteq E$ entonces:

$$\begin{aligned} f^{\#}(Y_1) + f^{\#}(Y_2) &= f(E) - f(E - Y_1) + f(E) - f(E - Y_2) \\ &\leq f(E) - f(Y_1^c \cap Y_2^c) + f(E) - f(Y_1^c \cup Y_2^c) \\ &= f(E) - f((Y_1 \cup Y_2)^c) + f(E) - f((Y_1 \cap Y_2)^c) \\ &= f^{\#}(Y_1 \cup Y_2) + f^{\#}(Y_1 \cap Y_2) \end{aligned}$$

□

$(\overline{\mathcal{D}}, f^{\#})$ se llama el *poliedro supermodular dual* de $(\mathcal{D}, f)$. Igualmente se puede definir el sistema submodular dual $(\overline{\mathcal{D}}, g^{\#})$ de un sistema supermodular $(\mathcal{D}, g)$ donde $\overline{\mathcal{D}}$ y $g^{\#}$ se definen similarmente. Tenemos además el siguiente lema respecto a la dualidad de sistemas submodulares.

Lema 3.1.7. *Si $(\mathcal{D}, f)$ es un sistema submodular entonces $B(f) = B(f^{\#})$ y $(f^{\#})^{\#} = f$.*

Prueba:

$$\begin{aligned} x \in B(f) &\Leftrightarrow x(X) \leq f(X) \text{ para todo } X \in \mathcal{D} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x(E - Y) \leq f(E - Y) \text{ para todo } Y \in \overline{\mathcal{D}} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x(E) - x(Y) \leq f(E - Y) \text{ para todo } Y \in \overline{\mathcal{D}} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x(Y) \geq x(E) - f(E - Y) \text{ para todo } Y \in \overline{\mathcal{D}} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x(Y) \geq f(E) - f(E - Y) \text{ para todo } Y \in \overline{\mathcal{D}} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x(Y) \geq f^{\#}(Y) \text{ para todo } Y \in \overline{\mathcal{D}} \text{ y } x(E) = f(E) \\ &\Leftrightarrow x \in B(f^{\#}). \end{aligned}$$

Por otro lado,

$(f^{\#})^{\#}(X) = f^{\#}(E) - f^{\#}(E - X) = f(E) - f(\varnothing) - f(E) + f(E - (E - X))$
$= f(X)$, y como $\overline{\overline{\mathcal{D}}} = \mathcal{D}$, se concluye lo deseado.

□

3.2. Reducciones y contracciones

A partir de un sistema submodular arbitrario se pueden generar nuevos sistemas submodulares de diferentes maneras, aquí utilizamos un elemento del retículo

asociado para generar nuevos sistemas submodulares llamados reducciones y contracciones. Además, los resultados de esta sección, tienen vital importancia para la teoría desarrollada en el capítulo 4.

Definición 3.2.1. *Sea $(\mathcal{D}, f)$ un sistema submodular en un conjunto E y $A \in \mathcal{D}$, definimos $\mathcal{D}^A := \{X \in \mathcal{D} \mid X \subseteq A\}$ y $f^A(X) := f(X)$ para cada $X \in \mathcal{D}^A$. El par $(\mathcal{D}^A, f^A)$ es claramente un sistema submodular y es llamado la reducción o la restricción de $(\mathcal{D}, f)$ al conjunto A. Este sistema submodular es denotado también como $(\mathcal{D}, f) \cdot A$.*

Definición 3.2.2. *Sea $A \in \mathcal{D}$, definimos $\mathcal{D}_A := \{X - A \mid A \subseteq X \in \mathcal{D}\}$ y $f_A(X) := f(X \cup A) - f(A)$, para $(X \in \mathcal{D}_{\mathcal{A}})$, o de la misma manera, $f_A(X - A) = f(X) - f(A)$ para $A \subseteq X \in \mathcal{D}$.*

Proposición 3.2.1. *Para cada $A \in \mathcal{D}$, $(\mathcal{D}_A, f_A)$ es un sistema submodular en $E - A$.*

Prueba: Para $X, Y \in \mathcal{D}$ con $A \subseteq X$, $A \subseteq Y$ tenemos $A \subseteq X \cup Y$, $A \subseteq X \cap Y$, $(X - A) \cup (Y - A) = (X \cup Y) - A$ y $(X - A) \cap (Y - A) = (X \cap Y) - A$, luego $\mathcal{D}_A$ es cerrado para uniones e intersecciones, además

$$\begin{aligned} f_A(X - A) + f_A(Y - A) &= f(X) - f(A) + f(Y) - f(A) \\ &\geq f(X \cup Y) - f(A) + f(X \cap Y) - f(A) \\ &= f_A((X \cup Y) - A) + f_A((X \cap Y) - A) \end{aligned}$$

Como vemos f_A es submodular en $\mathcal{D}_A$

□

El sistema submodular $(\mathcal{D}_A, f_A)$ es llamado la *contracción* de $(\mathcal{D}, f)$ por A y es denotado también como $(\mathcal{D}, f)/A$. Un sistema submodular obtenido de repetidas reducciones y contracciones es llamado un *menor* de $(\mathcal{D}, f)$. Tenemos el siguiente resultado en relación a reducciones y contracciones.

Lema 3.2.2. *Sea $(\mathcal{D}, f)$ un sistema submodular en un conjunto E, sea x^A una base de la reducción $(\mathcal{D}, f) \cdot A$ y x_A una base de la contracción $(\mathcal{D}, f)/A$, entonces la suma directa $\hat{x} = x^A \oplus x_A$ definida por*

$$x^A \oplus x_A\,(e) = \begin{cases} x^A(e) & \text{si } e \in A \\ x_A(e) & \text{si } e \in E - A \end{cases} \tag{3.3}$$

es una base de $(\mathcal{D}, f)$. Recíprocamente, para cada base $\hat{x}$ de $(\mathcal{D}, f)$ satisfaciendo $\hat{x}(A) = f(A)$, restringiendo $\hat{x}$ en A (o en $E - A$) se obtiene una base de $(\mathcal{D}, f) \cdot A$ (o de $(\mathcal{D}, f)/A$).

Prueba: Si x^A es una base de $(\mathcal{D}, f) \cdot A$ y x_A es una base de $(\mathcal{D}, f) \cdot A$, entonces para cada $X \in \mathcal{D}$ tenemos:

$$\begin{aligned} \hat{x}(X) &= \hat{x}(X \cap A) + \hat{x}(X - A) = x^A(X \cap A) + x_A(X - A) \\ &\leq f^A(X \cap A) + f_A(X - A) = f(X \cap A) + f(X \cup A) - f(A) \\ &\leq f(X) \end{aligned}$$

además,

$$\begin{aligned} \hat{x}(E) &= \hat{x}(A) + \hat{x}(E - A) = x^(A) + x_A(E - A) \\ &= f(A) + f_A(E - A) = f(A) + f(E) - f(A) \\ &= f(E) \end{aligned}$$

Luego como vemos $\hat{x}$ es una base de $(\mathcal{D}, f)$.

Recíprocamente, si $\hat{x}$ es una base de $(\mathcal{D}, f)$ y satisface $\hat{x}(A) = f(A)$, entonces para cada $X \in \mathcal{D}^A$ y $Y \in \mathcal{D}_A$ tenemos $x^A(X) = \hat{x}(X) \leq f(X)$ y

$$\begin{aligned} x_A(Y) &= \hat{x}(Y \cup A) - \hat{x}(A) \\ &\leq f(Y \cup A) - f(A) \\ &\leq f(Y) - f(Y \cap A) \\ &= f(Y) - f(\varnothing) = f(Y) \end{aligned}$$

□

3.3. Ejemplos de sistemas submodulares:

Además de los polimatroides, los cuales son también sistemas submodulares, tenemos los siguientes ejemplos de sistemas submodulares no polimatroidales.

Ejemplo 3.3.1. *(Funciones Corte)*

Sea $\mathcal{N} = (\, G = (V, A)\,,\, \underline{c}\,,\, \overline{c}\,)$ una red capacitada con grafo fundamental $G = (V, A)$ y funciones de capacidad inferior y superior $\underline{c} : A \longrightarrow \mathbb{R} \cup \{-\infty\}$ y $\overline{c} : A \longrightarrow \mathbb{R} \cup \{\infty\}$ con $\underline{c}(a) \leq \overline{c}(a)$ para todo $a \in A$. Definimos la función $\mathcal{K}_{\underline{c},\,\overline{c}} : 2^V \longrightarrow \mathbb{R} \cup \{\infty\}$ como:

$$\mathcal{K}_{\underline{c},\,\overline{c}}\,(U) = \sum_{a \in\, \Delta^+ U} \overline{c}(a) - \sum_{a \in\, \Delta^- U} \underline{c}(a)$$

para cada $U \subseteq V$. Aquí, $\Delta^+ U$ es el conjunto de arcos que salen de U en G y $\Delta^- U$ es el conjunto de arcos entrando a U en G. Para cada $U, W \in 2^V$ tal que $\mathcal{K}_{\underline{c},\,\overline{c}}\,(U) < \infty$ y $\mathcal{K}_{\underline{c},\,\overline{c}}\,(W) < \infty$ tenemos $\mathcal{K}_{\underline{c},\,\overline{c}}\,(U \cup W) < \infty$ y $\mathcal{K}_{\underline{c},\,\overline{c}}\,(U \cap W) < \infty$, ya que siendo $\mathcal{K}_{\underline{c},\,\overline{c}}\,(U) < \infty$ entonces $\overline{c}(a) < \infty$ para cada $a \in \Delta^+ U$ y $\underline{c}(a) > -\infty$ para cada $a \in \Delta^- U$, además $\Delta^+(U \cup W) \subseteq (\Delta^+ U) \cup (\Delta^+ W)$ y $\Delta^+(U \cap W) \subseteq (\Delta^+ U) \cap (\Delta^+ W)$ (La misma propiedad se tiene para Δ^-). Luego como vemos, el conjunto $\mathcal{D}(\underline{c}, \overline{c})$ definido como

$$\mathcal{D}(\underline{c}, \overline{c}) = \{U \subseteq V \mid \mathcal{K}_{\underline{c},\,\overline{c}}\,(U) < \infty\}$$

es cerrado para uniones e intersecciones, obsérvese que además $\varnothing, V \in \mathcal{D}(\underline{c}, \overline{c})$. Se puede ver también, realizando los cálculos cuidadosamente que para $U, W \in \mathcal{D}(\underline{c}, \overline{c})$ tenemos

$$\begin{aligned}
&\mathcal{K}_{\underline{c},\,\overline{c}}\,(U) + \mathcal{K}_{\underline{c},\,\overline{c}}\,(W) - \mathcal{K}_{\underline{c},\,\overline{c}}\,(U \cup W) - \mathcal{K}_{\underline{c},\,\overline{c}}\,(U \cap W) \\
&\quad = \sum_{a \in \Delta^+(U-W) \cap \Delta^-(W-U)} \overline{c}(a) \quad - \sum_{a \in \Delta^-(W-U) \cap \Delta^+(U-W)} \underline{c}(a) \\
&\qquad + \sum_{a \in \Delta^+(W-U) \cap \Delta^-(U-W)} \overline{c}(a) \quad - \sum_{a \in \Delta^-(U-W) \cap \Delta^+(W-U)} \underline{c}(a)
\end{aligned}$$

$$
\begin{aligned}
&= \sum \{\overline{c}(a) - \underline{c}(a) \mid a \in A,\ \partial^+ a \in U - W,\ \partial^- a \in W - U\} \\
&\qquad + \sum \{\overline{c}(a) - \underline{c}(a) \mid a \in A,\ \partial^- a \in U - W,\ \partial^+ a \in W - U\} \\
&\geq 0
\end{aligned}
$$

es decir, $\mathcal{K}_{\underline{c},\overline{c}}$ es una función submodular en $\mathcal{D}(\underline{c},\overline{c})$, con $\mathcal{K}_{\underline{c},\overline{c}}(\varnothing) = \mathcal{K}_{\underline{c},\overline{c}}(V) = 0$, o en otras palabras, $(\mathcal{D}(\underline{c},\overline{c}), \mathcal{K}_{\underline{c},\overline{c}})$ es un sistema submodular que claramente es no polimatroidal pues por ejemplo $\mathcal{K}_{\underline{c},\overline{c}}$ puede tomar valores negativos y en general no es monótona no decreciente. La función $\mathcal{K}_{\underline{c},\overline{c}}$ es llamada la *función corte* asociada a la red $\mathcal{N}$.

Un *flujo factible* φ en $\mathcal{N}$ es una función $\varphi : A \longrightarrow \mathbb{R}$ tal que $\underline{c}(a) \leq \varphi(a) \leq \overline{c}(a)$ para cada $a \in A$, recordemos que la *frontera* $\partial\varphi : V \longrightarrow \mathbb{R}$ de φ está definida como:

$$\partial\varphi(v) = \sum_{a \in \delta^+ v} \varphi(a) - \sum_{a \in \delta^- v} \varphi(a) \qquad (v \in V)$$

Definimos además el conjunto de funciones frontera de flujos factibles de la red $\mathcal{N}$ como:

$$\partial\Phi := \{\partial\varphi \mid \varphi \textit{ es flujo factible en } \mathcal{N}\}$$

Si $B(\mathcal{K}_{\underline{c},\overline{c}})$ denota al poliedro base del sistema submodular $(\mathcal{D}(\underline{c},\overline{c}), \mathcal{K}_{\underline{c},\overline{c}})$ entonces tenemos $\partial\Phi = B(\mathcal{K}_{\underline{c},\overline{c}})$.

En efecto, sea $\varphi : A \longrightarrow \mathbb{R}$ tal que para todo $a \in A$, $\underline{c}(a) \leq \varphi(a) \leq \overline{c}(a)$, para cada $U \subseteq V$ se tiene

$$\partial\varphi(U) = \sum_{v \in U} \partial\varphi(v) = \sum_{a \in \Delta^+ U} \varphi(a) - \sum_{a \in \Delta^- U} \varphi(a)$$

pues si $a \in A$ es tal que $\partial^+(a)$ y $\partial^-(a)$ están en U entonces los términos $\varphi(a)$ y $-\varphi(a)$ aparecen cada uno una sóla vez en el cálculo de $\partial\varphi(U)$, luego estos términos se cancelarán.

Ahora, como $\underline{c}(a) \leq \varphi(a) \leq \overline{c}(a)$ entonces $-\varphi(a) \leq -\underline{c}(a)$ para cada $a \in A$, luego

$$\begin{aligned}\partial\varphi(U) &= \sum_{a\in\Delta^+U} \varphi(a) - \sum_{a\in\Delta^-U} \varphi(a) \\ &\leq \sum_{a\in\Delta^+U} \overline{c}(a) - \sum_{a\in\Delta^-U} \underline{c}(a) = \mathcal{K}_{\underline{c},\overline{c}}(U)\end{aligned}$$

y como $\partial\varphi(V) = \mathcal{K}_{\underline{c},\overline{c}}(V) = 0$ entonces hemos mostrado que $\partial\Phi \subseteq B(\mathcal{K}_{\underline{c},\overline{c}})$

Recíprocamente, sea $x \in B(\mathcal{K}_{\underline{c},\overline{c}})$, considere un nuevo vértice $s \notin V$ con nuevos arcos (s, v) para cada $v \in V$, definimos $\underline{c}(s, v) = \overline{c}(s, v) = x(v)$ para cada $v \in V$. Denotamos esta nueva red como

$$\mathcal{N}' = \big(G' = (V \cup \{s\}, A \cup \{(s, v) : v \in V\}), \underline{c}, \overline{c}\big)$$

Por el teorema 1.3.1 (Hoffman) existe una circulación factible φ en $\mathcal{N}$ si y sólo si para todo $U \subseteq V \cup \{s\}$ se tiene

$$\sum_{a\in\Delta^+U} \overline{c}(a) \geq \sum_{a\in\Delta^-U} \underline{c}(a)$$

donde Δ^+U y Δ^-U se toman respecto a G'. Pero esta última desigualdad se tiene, ya que para $U \subseteq V \cup \{s\}$, si $s \notin U$

$$\begin{aligned}\sum_{\substack{a\in\Delta^+U \\ \text{en } G'}} \overline{c}(a) \geq \sum_{\substack{a\in\Delta^-U \\ \text{en } G'}} \underline{c}(a) &\Leftrightarrow \sum_{\substack{a\in\Delta^+U \\ \text{en G}}} \overline{c}(a) \geq \sum_{\substack{a\in\Delta^-U \\ \text{en G}}} \underline{c}(a) + \sum_{v\in U} x(v) \\ &\Leftrightarrow \sum_{v\in U} x(v) \leq \sum_{\substack{a\in\Delta^+U \\ \text{en G}}} \overline{c}(a) - \sum_{\substack{a\in\Delta^-U \\ \text{en G}}} \underline{c}(a) \\ &\Leftrightarrow x(U) \leq \mathcal{K}_{\underline{c},\overline{c}}(U)\end{aligned}$$

Pero la última desigualdad se tiene ya que $x \in B(\mathcal{K}_{\underline{c},\overline{c}})$. En el caso en que $s \in U$,

sea $U_0 = U - \{s\}$, tenemos:

$$\sum_{\substack{a\in\Delta^+U\\ \text{en } G'}} \overline{c}(a) \geq \sum_{\substack{a\in\Delta^-U\\ \text{en } G'}} \underline{c}(a) \Leftrightarrow \sum_{\substack{a\in\Delta^+U_0\\ \text{en G}}} \overline{c}(a) + \sum_{v\in V-U_0} x(v) \geq \sum_{\substack{a\in\Delta^-U\\ \text{en G}}} \underline{c}(a)$$

$$\Leftrightarrow \sum_{\substack{a\in\Delta^+U_0\\ \text{en G}}} \overline{c}(a) + x(V) - x(U_0) \geq \sum_{\substack{a\in\Delta^-U\\ \text{en G}}} \underline{c}(a)$$

$$\Leftrightarrow \sum_{\substack{a\in\Delta^+U_0\\ \text{en G}}} \overline{c}(a) - \sum_{\substack{a\in\Delta^-U_0\\ \text{en G}}} \underline{c}(a) \geq x(U_0) - \mathcal{K}_{\underline{c},\overline{c}}(V)$$

$$\Leftrightarrow \mathcal{K}_{\underline{c},\overline{c}}(U_0) \geq x(U_0).$$

Restringiendo φ a A obtenemos un flujo factible en $\mathcal{N}$ cuya frontera es igual a x.

Ejemplo 3.3.2. *(Funciones submodulares derivadas de funciones cóncavas)*

Sea E un conjunto finito y $g : E \longrightarrow \mathbb{R}$, definamos la función $f : 2^E \longrightarrow \mathbb{R}$ por $f(X) = g(|X|)$ para $X \subseteq E$. Por el teorema 1.1.2 f es una función submodular en 2^E. Luego como $f(\varnothing) = 0$ el par $(2^E, f)$ es un sistema submodular en E que generalmente no es polimatroidal ya que g no necesariamente es no decreciente.

Un sistema submodular de este tipo se puede derivar del concepto de *mayorización*:

Dado un vector $x \in \mathbb{R}^E$ ordenemos las componentes $x(e)$ $(e \in E)$ en orden descendente, y para $i = 1, 2, ..., |E|$, denotemos por $x_{[i]}$ al iésimo elemento en el orden descrito. Dados dos vectores $x, y \in \mathbb{R}^E$, decimos que x es *mayorizado* por y si

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]} \qquad (k = 1, 2, ..., |E| - 1) \text{ y}$$

$$\sum_{i=1}^{|E|} x_{[i]} = \sum_{i=1}^{|E|} y_{[i]}$$

Definamos ahora para $y \in \mathbb{R}^E$ la función $f_y : 2^E \longrightarrow \mathbb{R}$ por

$$f_y(X) := \sum_{i=1}^{|X|} y_{[i]}$$

es decir, $f_y(X) = g(|X|)$ donde $g(n) = \sum_{i=1}^{n} y_{[i]}$. Para ver que g es una función cóncava basta ver que $g(n+1) - g(n) \geq g(n+2) - g(n+1)$ para $n = 1, 2, ...|E| - 2$. Esto se puede verificar fácilmente:

$$g(n+1) - g(n) = \sum_{i=1}^{n+1} y_{[i]} - \sum_{i=1}^{n} y_{[i]} = y_{[n+1]}$$

de igual manera, $g(n+2) - g(n+1) = y_{[n+2]}$, pero por hipótesis $y_{[n+1]} \geq y_{[n+2]}$.

De aquí se concluye que $(2^E, f_y)$ es un sistema submodular.

Otra observación que tenemos es que un vector $x \in \mathbb{R}^E$ es mayorizado por el vector y si y sólo si $x \in B(f_y)$, en efecto, si x es mayorizado por y entonces para $X \subseteq E$,

$$x(X) = \sum_{e \in X} x(e) = \sum_{i=1}^{|X|} x_{[i]} \leq \sum_{i=1}^{|X|} y_{[i]} = f_y(X)$$

en donde para el caso $X = E$ se tiene claramente la igualdad. Se concluye que $x \in B(f_y)$.

Recíprocamente, si $x \in B(f_y)$ entonces para $X \subseteq E$ se tiene $x(X) \leq f_y(X)$, es decir,

$$\sum_{i=1}^{|X|} x_{[i]} \leq \sum_{i=1}^{|X|} y_{[i]}$$

y como $x(E) = f_y(E)$ entonces

$$\sum_{i=1}^{|E|} x_{[i]} = \sum_{i=1}^{|E|} y_{[i]}$$

luego x es mayorizado por y.

Ejemplo 3.3.3. *(Matrices Monge)*

Sea $A = [a_{ij}]_{m\times n}$ una matriz de tamaño $m \times n$. Decimos que A es una matriz *Monge* si para p, q, r, s tales que $1 \leq p < q \leq m$ y $1 \leq r < s \leq n$ tenemos $a_{ps} + a_{qr} \geq a_{qs} + a_{pr}$. Sea M la suma directa entre el conjunto fila $R = \{1, 2, .., m\}$ y el conjunto columna $C = \{1, 2, ..., n\}$, es decir, $M = R \oplus C = \{(1, 1), (1, 2), ..., (1, m), (2, 1), (2, 2), ..., (2, n)\}$. Para $i \in R$ y $j \in C$ definimos $R(i) = \{i' \mid 1 \leq i' \leq i\}$ y $C(j) = \{j' \mid 1 \leq j' \leq j\}$ y definimos además $I(i, j) = R(i) \oplus C(j)$.

El conjunto $\mathcal{D} := \{I(i, j) \mid 1 \leq i \leq m, 1 \leq j \leq n\} \cup \{\varnothing\}$ es cerrado para uniones e intersecciones ya que si $I(i, j), I(k, l) \in \mathcal{D}$ entonces

$$I(i, j) \cap I(k, l) = I\big(\text{mín}\{i, k\}, \text{mín}\{j, l\}\big) \quad \text{e}$$

$$I(i, j) \cup I(k, l) = I\big(\text{máx}\{i, k\}, \text{máx}\{j, l\}\big)$$

Si definimos $f : \mathcal{D} \longrightarrow \mathbb{R}$ por $f(I(i, j)) = a_{ij}$ y $f(\varnothing) = 0$ entonces f es una función submodular en $\mathcal{D}$, en efecto, si $X = I(i, j)$ y $Y = I(k, l)$ son elementos en $\mathcal{D}$ entonces tenemos tres casos:

Caso 1: $i = k$.

Si $j = l$ entonces $f(X) + f(Y) = a_{ij} + a_{ij} = f(X \cup Y) + f(X \cap Y)$
Si $j < l$ entonces $f(X) + f(Y) = a_{ij} + a_{il} = f(X \cap Y) + f(X \cup Y)$
Si $j > l$ entonces $f(X) + f(Y) = a_{ij} + a_{il} = f(X \cup Y) + f(X \cap Y)$

Caso 2: $i < k$.

Si $j = l$ entonces $f(X) + f(Y) = a_{ij} + a_{kj} = f(X \cap Y) + f(X \cup Y)$
Si $j < l$ entonces $f(X) + f(Y) = a_{ij} + a_{kl} = f(X \cap Y) + f(X \cup Y)$
Si $j > l$ entonces $f(X) + f(Y) = a_{ij} + a_{kl} \geq a_{il} + a_{kj}$
$= f(X \cap Y) + f(X \cup Y)$

Caso 3: $i > k$.

En este caso se tiene también $f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$ (por los casos anteriores).

Concluímos entonces que el par $(\mathcal{D}, f)$ es un sistema submodular.

Capítulo 4

Algoritmo Greedy

Dado un sistema submodular $(\mathcal{D}, f)$ y una función cualquiera $w : E \longrightarrow \mathbb{R}$, consideremos el problema de optimización lineal:

$$\mathbf{P}_w : \text{ Minimizar} \sum_{e \in E} w(e)x(e) \qquad \text{sujeto a } \; x \in B(f) \tag{4.1}$$

Obsérvese que la condición $x \in B(f)$ está determinada por el conjunto de desigualdades y una igualdad:

$$x(X) \leq f(X) \;\; (X \in \mathcal{D})$$

$$x(E) = f(E)$$

por lo tanto $\mathbf{P}_w$ es un problema típico de optimización lineal.

Para el problema $\mathbf{P}_w$ de la ecuación (4.1), la función w es llamada la *función de peso* de $\mathbf{P}_w$ y una solución óptima de $\mathbf{P}_w$ es llamada una *base de mínimo peso* de $(\mathcal{D}, f)$ respecto a w. Así mismo, una *base de máximo peso* de $(\mathcal{D}, f)$ respecto a w es una solución óptima del problema $\mathbf{P}_{-w}$.

El objetivo de este capítulo es resolver este problema, veremos que la existencia de la solución óptima depende de algunas propiedades que debe tener por un lado el sistema submodular $(\mathcal{D}, f)$, y por otro la función de peso w, es por ello que tendremos

que presentar en principio algunos resultados relacionados con retículos distributivos y posets, además de algunas propiedades del poliedro base que son inducidas por características propias del retículo distributivo $\mathcal{D}$. Una vez verificadas las condiciones sobre el sistema submodular $(\mathcal{D}, f)$ y la función de peso w, se probará la existencia de la solución óptima y se presentará un algoritmo, denominado el algoritmo greedy o "voraz"[1] que determina dicha solución. Este algoritmo depende en principio de hallar extensiones lineales de un poset $\mathcal{P}$ (órdenes totales que preservan el orden en $\mathcal{P}$), dichas extensiones deben tener además ciertas características relacionadas con la función de peso w. El algoritmo presentará en seguida y directamente, las coordenadas de una solución óptima. Cabe destacar que aunque los algoritmos greedy generalmente no conducen siempre a una solución óptima, en este caso en particular el algoritmo si exhibirá una solución óptima.

Finalmente, presentaremos un ejemplo ilustrando los resultados del capítulo.

De aquí en adelante, cuando hablemos de un retículo distributivo $\mathcal{D} \subseteq 2^E$ asumiremos que las operaciones de retículo son la unión y la intersección de conjuntos.

4.1. Retículos distributivos y posets.

Definición 4.1.1. *Sea $\mathcal{D} \subseteq 2^E$ un retículo distributivo con $\varnothing, E \in \mathcal{D}$. Un conjunto $\mathcal{C} = \{S_0, S_1, ...S_k\}$ de elementos de $\mathcal{D}$ tales que $S_0 \subset S_1 \subset ... \subset S_k$ es llamado una cadena en $\mathcal{D}$, decimos además que k es la longitud de $\mathcal{C}$. Si no existe una cadena en $\mathcal{D}$ que contiene a $\mathcal{C}$ como subconjunto propio, entonces decimos que $\mathcal{C}$ es una cadena maximal de $\mathcal{D}$.*

Definición 4.1.2. *Para cada $e \in E$ definimos el conjunto*

$$D(e) := \bigcap \{X \in \mathcal{D} \mid e \in X\}$$

Obsérvese que $D(e)$ es el único elemento minimal de $\mathcal{D}$ que contiene a e y que para cada $e' \in D(e)$ se tiene $D(e') \subseteq D(e)$.

Sea $G(\mathcal{D}) = (E, A(\mathcal{D}))$ un grafo dirigido cuyo conjunto de vértices es E y cuyo conjunto de aristas es el conjunto $A(\mathcal{D})$ definido como:

$$A(D) := \{(e, e') \mid e \in E, e' \in D(e)\}$$

Descompongamos el grafo $G(\mathcal{D})$ en componentes fuertemente conexas $G_i = (F_i, A_i)$, $(i \in I)$. Sea $\preceq_{\mathcal{D}}$ el orden parcial en el conjunto de componentes fuertemente conexas inducido por la descomposición. Si (e, e') y (e', e'') están en $A(\mathcal{D})$ entonces (e, e'') está en $A(D)$, esto implica que para cada dos vértices e_1, e_2 en una componente G_i, el arco (e_1, e_2) está en G_i. Luego además, si $G_i \preceq_{\mathcal{D}} G_j$ para algunos $i, j \in I$ entonces existe un arco de cada vértice de G_j a cada vértice de G_i.

Si definimos el conjunto $\Pi(\mathcal{D}) = \{F_i \mid i \in I\}$, entonces $\Pi(\mathcal{D})$ es una partición de E, luego el orden parcial $\preceq_{\mathcal{D}}$ se puede inducir al conjunto $\Pi(\mathcal{D})$ identificando G_i con F_i para cada $i \in I$.

Ejemplo 4.1.1. *Sea* $E = \{1, 2, 3, 4, 5\}$ *y* $\mathcal{D} = \big\{\varnothing, E, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 4, 5\}\big\}$, *el grafo* $G(\mathcal{D})$ *y su descomposición se pueden ver en la siguiente figura:*

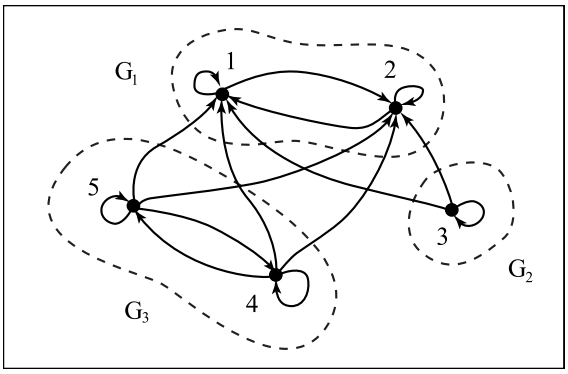

Figura 4.1: Descomposición del grafo asociado a $\mathcal{D}$.

Aquí $F_1 = \{1, 2\}, F_2 = \{3\}$ y $F_3 = \{4, 5\}$, obsérvese que $F_1 \preceq_{\mathcal{D}} F_2$ y $F_1 \preceq_{\mathcal{D}} F_3$.

Definición 4.1.3. *El poset* $\mathcal{P}(\mathcal{D}) = \big(\Pi(\mathcal{D}), \preceq_{\mathcal{D}}\big)$ *es llamado el poset derivado del retículo distributivo* $\mathcal{D}$.

Recordemos que un ideal de un poset arbitrario $\mathcal{P} = (P, \preceq)$ es un conjunto $J \subseteq P$ tal que para cada $e, e' \in \mathcal{D}$ tenemos $e \preceq e' \in J \Rightarrow e \in J$.

En relación al poset $\mathcal{P}(\mathcal{D})$ de la definición anterior tenemos el siguiente teorema:

Teorema 4.1.1. *(Birkhoff) Sea* $\mathcal{D} \subseteq 2^E$ *un retículo distributivo con* $\varnothing, E \in \mathcal{D}$*, para el poset* $\mathcal{P}(\mathcal{D})$ *se tiene:*

i) Para cada ideal J *de* $\mathcal{P}(\mathcal{D})$

$$\bigcup\{F \mid F \in J\} \in \mathcal{D}$$

ii) Dado $X \in \mathcal{D}$ *existe un ideal* J *de* $\mathcal{P}(\mathcal{D})$ *tal que*

$$X = \bigcup\{F \mid F \in J\}$$

Prueba: (i) Sea $X = \bigcup\{F \mid F \in J\}$. Afirmamos que $D(e) \subseteq X$ para todo $e \in X$.

En efecto, si $e' \in D(e)$ entonces $(e, e') \in A(\mathcal{D})$ y como $e \in X$, entonces $e \in F$ para algún $F \in J$. Si $e' \in F$ entonces $e' \in X$, suponga entonces que $e \notin F$ y sea $F_i \in \Pi(\mathcal{D})$ con $F_i \neq F$, tal que $e' \in F_i$. (e, e') es un camino dirigido de un vértice de F a un vértice de F_i, con lo cual $F_i \preceq_{\mathcal{D}} F$, de aquí, $F_i \in J$, luego $e' \in X$. Se concluye que $D(e) \subseteq X$.

La afirmación anterior nos permite escribir $X = \bigcup\{D(e) \mid e \in X\}$ y de aquí, $X \in \mathcal{D}$.

(ii) Dados $X \in \mathcal{D}$ y $F \in \Pi(\mathcal{D})$, ellos no se cruzan, es decir, $F \subseteq X$ o $F \subseteq E - X$, en efecto:

Caso 1) Supongamos que $F \cap X = \varnothing$. Sea $e \in F \cap X$, si $e' \in F$ entonces $e' \in D(e)$ (pues $F \in \Pi(\mathcal{D})$), pero $D(e) \subseteq X$, es decir, $e' \in X$. En otras palabras $F \subseteq X$.

Caso 2) Si $F \cap X \neq \varnothing$ entonces $F \subseteq E - X$.

Definiendo $J = \{F \in \Pi(\mathcal{D}) \mid F \subseteq X\}$ entonces J es una partición de X y claramente $X = \bigcup\{F \mid F \in J\}$. Por otro lado, sean $F_1, F_2 \in \Pi(\mathcal{D})$ tales que $F_2 \subseteq X$ y $F_1 \preceq_{\mathcal{D}} F_2$. Si $e' \in F_1$, para $e \in F_2$ tenemos $(e, e') \in A(\mathcal{D})$, luego $e' \in D(e)$, pero como $e \in X$ entonces $D(e) \subseteq X$, así, $e' \in X$ o en otras palabras $F_1 \subseteq X$. De aquí, que J es el ideal deseado.

$\square$

El teorema anterior nos muestra que hay una correspondencia uno a uno entre $\mathcal{D}$ y el conjunto de todos los ideales de $\mathcal{P}(\mathcal{D})$. El siguiente teorema muestra también una correspondencia importante entre conjuntos.

Teorema 4.1.2. *Hay una correspondencia uno a uno entre el conjunto de retículos distributivos $\mathcal{D} \subseteq 2^E$ con $\varnothing, E \in \mathcal{D}$, y el conjunto de posets $\mathcal{P} = (P, \preceq)$ en particiones de E*

Prueba: Hemos visto que dado un retículo distributivo $\mathcal{D} \subseteq 2^E$ con $\varnothing, E \in \mathcal{D}$, $\mathcal{P}(\mathcal{D})$ es un poset en una partición de E. Ahora, sea $\mathcal{P} = (P, \preceq)$ un poset en una partición de E, si definimos

$$\mathcal{D}(\mathcal{P}) := \{\widetilde{J} \mid J \text{ es un ideal de } \mathcal{P}\}$$

en donde $\widetilde{J} := \bigcup\{F \mid F \in J\}$, entonces $\mathcal{D}(\mathcal{P})$ es un retículo distributivo con $\varnothing, E \in \mathcal{D}(\mathcal{P})$, en efecto, sean $\widetilde{X}, \widetilde{Y} \in \mathcal{D}(\mathcal{P})$, existen X, Y ideales de $\mathcal{P}$ tales que $\widetilde{X} = \bigcup\{F \mid F \in X\}$ y $\widetilde{Y} = \bigcup\{F \mid F \in Y\}$. $X \cup Y$ es un ideal de $\mathcal{P}$ y $\widetilde{X} \cup \widetilde{Y} = \bigcup\{F \mid F \in X \cup Y\} \in \mathcal{D}(\mathcal{P})$, para probar que $\widetilde{X} \cap \widetilde{Y} \in \mathcal{D}(\mathcal{P})$ observamos que

$$\begin{aligned} x \in \widetilde{X} \cap \widetilde{Y} &\Leftrightarrow x \in F \text{ para algún } F \in X \text{ y } x \in G \text{ para algún } G \in Y \\ &\Leftrightarrow x \in F \cap G \text{ para algunos } F \in X, G \in Y \\ &\Leftrightarrow x \in \bigcup\{H \mid H \in I\} \end{aligned}$$

en donde $I = \{F \cap G \mid F \in X, G \in Y\}$. Pero claramente $I = X \cap Y$ ya que P es una partición en E, concluyéndose así, que $\widetilde{X} \cap \widetilde{Y} \in \mathcal{D}(\mathcal{P})$.

El teorema 4.1.1 nos indica además que $\mathcal{P}(\mathcal{D}(\mathcal{P})) = \mathcal{P}$. También podemos ver que la aplicación que asigna a cada $\mathcal{D}$ su poset derivado $\mathcal{P}(\mathcal{D})$ es una correspondenca uno a uno, pues si $\mathcal{D}_1$ y $\mathcal{D}_2$ son retículos distributivos con $\varnothing, E \in \mathcal{D}_1 \cap \mathcal{D}_2$ y suponemos $\mathcal{P}(\mathcal{D}_1) = \mathcal{P}(\mathcal{D}_2)$, entonces por el teorema 4.1.1:

$$\begin{aligned}\mathcal{D}_1 &= \{\widetilde{J} \mid J \text{ es un ideal de } \mathcal{P}(\mathcal{D}_1)\}\\ &= \{\widetilde{J} \mid J \text{ es un ideal de } \mathcal{P}(\mathcal{D}_2)\} = \mathcal{D}_2\end{aligned}$$

□

Otro importante resultado se presenta a continuación:

Teorema 4.1.3. *Dado un retículo distributivo $\mathcal{D} \subseteq 2^E$ con $\varnothing, E \in \mathcal{D}$. Si*

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset ... \subset S_k = E$$

es una cadena maximal arbitraria de $\mathcal{D}$, entonces

$$\Pi(\mathcal{D}) = \{S_i - S_{i-1} \mid i = 1, 2, ..., k\} \tag{4.2}$$

Prueba: Sea $j \in \{1, 2, ..., k\}$ y $e_1, e_2 \in S_j - S_{j-1}$ tales que $e_2 \notin D(e_1)$, tenemos entonces que

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset ... \subset S_{j-1} \subset S_{j-1} \cup D(e_1) \subset S_j \subset ... \subset S_k = E$$

es una cadena que contiene a $\mathcal{C}$ y que es distinta de $\mathcal{C}$, lo cual es absurdo ya que $\mathcal{C}$ es maximal, esto implica que $S_j - S_{j-1} \subseteq F_0$ para algún $F_0 \in \Pi(\mathcal{D})$.

Por otro lado sea $F \in \Pi(\mathcal{D})$ y $j \in \{1, 2, ..., k\}$ tal que S_j es el mínimo conjunto de $\mathcal{C}$ que contiene a F, como $S_j - S_{j-1} \cap F \neq \varnothing$ y $\Pi(\mathcal{D})$ es una partición de E entonces por la primera observación de esta prueba se debe tener $S_j - S_{j-1} \subseteq F$, pero si existe un elemento $e \in F$ tal que $e \notin S_j - S_{j-1}$ entonces para cada $e' \in S_j - S_{j-1}$ tenemos $e' \in D(e) \subseteq S_j$, lo cual es absurdo. Concluímos entonces que $F = S_j - S_{j-1}$ y así $\Pi(\mathcal{D}) \subseteq \{S_i - S_{i-1} \mid i = 1, 2, ..., k\}$, pero siendo ambos conjuntos particiones de E, necesariamente se tiene la igualdad (4.2).

□

Definición 4.1.4. *Un retículo distributivo $\mathcal{D}$ es llamado simple si la partición $\Pi(\mathcal{D})$ está compuesta de singletons, es decir $\Pi(\mathcal{D}) = \big\{\{e\} : e \in E\big\}$. Un sistema submodular $(\mathcal{D}, f)$ en donde $\mathcal{D}$ es un retículo distributivo simple es llamado un sistema submodular simple.*

Para un retículo distributivo simple $\mathcal{D}$, podemos considerar entonces a $\mathcal{P}(\mathcal{D})$ como un poset en E y escribimos $\mathcal{P}(\mathcal{D}) = (E, \preceq_{\mathcal{D}})$. Recíprocamente, dado un poset $\mathcal{P} = (E, \preceq)$ en E, el conjunto de todos ideales de $\mathcal{P}$ forman un retículo distributivo $\mathcal{D} \subseteq 2^E$ el cual es simple ya que para $e \in E$, se tiene $D(e) = \{e' \in E \mid e' \preceq e\}$, luego si $e' \in D(e)$ y $e \in D(e')$ entonces $e \preceq e'$ y $e' \preceq e$, así $e = e'$ concluyéndose que $\Pi(\mathcal{D}) = \big\{\{e\} : e \in E\big\}$. En este caso escribimos $\mathcal{D} = 2^{\mathcal{P}}$.

Teorema 4.1.4. *Dado un sistema submodular arbitrario $(\mathcal{D}, f)$ en E, si definimos*

$$\widehat{X} := \{F \in \Pi(\mathcal{D}) \mid F \subseteq X\} \text{ para cada } X \in \mathcal{D},$$

$\widehat{\mathcal{D}} := \{\widehat{X} \mid X \in \mathcal{D}\}$ y $\widehat{f}(\widehat{X}) := f(X)$ $(X \in \mathcal{D})$, entonces $(\widehat{\mathcal{D}}, \widehat{f})$ es un sistema submodular simple en $\Pi(\mathcal{D})$ y es llamado la simplificación de $(\mathcal{D}, f)$.

Prueba: Si escribimos $\mathcal{P} = \big(\Pi(\mathcal{D}), \preceq_{\mathcal{D}}\big)$ entonces hemos visto que $\mathcal{D}(\mathcal{P}) = \{J \mid J$ es un ideal de $\mathcal{P}\}$ es un retículo simple en $\Pi(\mathcal{D})$, pero $\mathcal{D}(\mathcal{P}) = \widehat{\mathcal{D}}$ ya que si $J \in \mathcal{D}(\mathcal{P})$, por el teorema 4.1.1, $X = \bigcup\{F \mid F \in J\} \in \mathcal{D}$, y claramente $J = \widehat{X}$, es decir, $J \in \widehat{\mathcal{D}}$.

Recíprocamente, si $\widehat{X} \in \widehat{\mathcal{D}}$, entonces $\widehat{X} = \{F \in \Pi(\mathcal{D}) \mid F \subseteq X\}$ para algún $X \in \mathcal{D}$, luego por el teorema 4.1.1 existe un ideal J de $\mathcal{P}(\mathcal{D})$ tal que $X = \bigcup\{F \mid F \in J\}$, pero esta última es una unión disyunta de elementos de $\Pi(\mathcal{D})$, luego necesariamente $\widehat{X} = J$, es decir, $\widehat{X} \in \mathcal{D}(\mathcal{P})$.

La submodularidad de $\widehat{f}$ se deriva de la submodularidad de f.

□

4.2. Algoritmo greedy

En esta sección presentaremos el algoritmo greedy que soluciona el problema $\mathbf{P}_w$ de la ecuación (4.1), aunque antes de ello presentaremos, utilizando la teoría desarrollada en la sección anterior, algunas propiedades del poliedro base $B(f)$ en términos de las propiedades del retículo distributivo $\mathcal{D}$ así como algunos resultados que no sólo permiten simplificar el problema, sino que garantizan, bajo algunas condiciones, la existencia de la solución óptima.

Teorema 4.2.1. *El poliedro base $B(f)$ de un sistema submodular $(\mathcal{D}, f)$ es puntiagudo (es decir, tiene al menos un punto extremo) sii $\mathcal{D}$ es simple, en otras palabras $\mathcal{D} = 2^{\mathcal{P}}$ para algún poset $\mathcal{P} = (E, \preceq)$.*

Prueba: Del lema 1.4.1, un poliedro es puntiagudo si y sólo si su cono característico no contiene ninguna recta, para el caso del poliedro base $B(f)$, su cono característico será el conjunto solución del sistema de desigualdades y una ecuación:

$$x(X) \leq 0 \quad (X \in \mathcal{D})$$

$$x(E) = 0$$

luego también a partir del lema 1.4.1, $B(f)$ es puntiagudo sii el sistema de ecuaciones

$$x(X) = 0 \quad (X \in \mathcal{D}) \tag{4.3}$$

tiene única solución $x = 0$. Podemos ver que (4.3) es equivalente al sistema:

$$x(F) = 0 \quad (F \in \Pi(\mathcal{D})) \tag{4.4}$$

En efecto, para $X \in \mathcal{D}$, por el teorema 4.1.1, $X = \bigcup\{F \mid F \in J\}$ para algún ideal J de $\mathcal{P}(\mathcal{D})$, luego si $x(F) = 0$ para todo $F \in \Pi(\mathcal{D})$ entonces $x(X) = 0$ para todo $X \in \mathcal{D}$ $\big($pues $x(X) = \sum_{F \in J} x(F)\big)$. Recíprocamente, suponga que $x(X) = 0$ para todo $X \in \mathcal{D}$, sea

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset ... \subset S_k = E$$

una cadena maximal. Dado $F \in \Pi(\mathcal{D})$ por el teorema 4.1.3, $F = S_i - S_{i-1}$ para algún $i \in \{1, 2, ..., k\}$, luego como x es modular,

$$\begin{aligned} 0 = x(S_i) &= x((S_i - S_{i-1}) \cup S_{i-1}) = x(S_i - S_{i-1}) + x(S_{i-1}) \\ &= x(S_i - S_{i-1}) = x(F), \end{aligned}$$

y así se concluye que (4.3) es equivalente a (4.4).

Por otro lado, el sistema (4.4) tiene única solución $x = 0$ sii $|F| = 1$ para todo $F \in \Pi(\mathcal{D})$, esto se comprueba observando que $F_i \cap F_j = \varnothing$ para $F_i, F_j \in \Pi(\mathcal{D})$ $(i \neq j)$ y si $F = \{e_1, e_2, ...e_p\}$ con $p \geq 2$ entonces $x(e_1) + x(e_2) + \ldots + x(e_p) = 0$ es una ecuación independiente en (4.4) que tiene infinitas soluciones.

Concluimos que $\mathcal{D}$ debe ser simple.

□

Teorema 4.2.2. *El poliedro base $B(f)$ es acotado sii $\mathcal{D} = 2^E$.*

Prueba: Supongamos que $B(f)$ es acotado y que $\mathcal{D} \neq 2^E$, existen elementos distintos $e, e' \in E$ tales que $e' \in D(e)$, luego dada cualquier base $x \in B(f)$ el rayo $\{x + \alpha(\chi_e - \chi_{e'}) \mid \alpha \geq 0\}$ está contenido en $B(f)$ ya que para cada $X \in \mathcal{D}$, si $e \in X$, entonces $e' \in X$, luego

$$x(X) + \alpha(\chi_e(X) - \chi_{e'}(X)) = x(X) \leq f(X),$$

si $e \notin X$, entonces

$$x(X) + \alpha(\chi_e(X) - \chi_{e'}(X)) = x(X) - \alpha\chi_{e'}(X) \leq x(X) \leq f(X),$$

además, $x(E) + \alpha(\chi_e(E) - \chi_{e'}(E)) = x(E) = f(E)$. Se comprueba entonces que $B(f)$ no es acotado, lo cual es contradictorio.

Recíprocamente, supongamos que $\mathcal{D} = 2^E$. Si $e \in E$, escribamos a E en la forma $E = \{e, e_1, e_2, ...e_n\}$, el sistema

$$\begin{cases} x(e) \leq f(\{e\}) & (e \in E) \\ x(E) = f(E) \end{cases} \tag{4.5}$$

es equivalente al sistema

$$\begin{cases} x(e) \leq f(\{e\}) \\ x(e_i) \leq f(\{e_i\}) \\ x(E) = f(E) \end{cases} \tag{4.6}$$

luego si x es solución de este sistema entonces

$$\sum_{i=1}^{n} x(e_i) \leq \sum_{i=1}^{n} f(\{e_i\})$$

con lo cual

$$0 \geq \sum_{i=1}^{n} x(e_i) - \sum_{i=1}^{n} f(\{e_i\})$$

así,

$$x(e) \geq x(e) + \sum_{i=1}^{n} x(e_i) - \sum_{i=1}^{n} f(\{e_i\})$$

es decir,

$$x(e) \geq x(E) - \sum_{i=1}^{n} f(\{e_i\})$$

o en otras palabras,

$$x(E) - \sum_{i=1}^{n} f(\{e_i\}) \leq x(e) \leq f(\{e\}) \leq \operatorname*{máx}_{e' \in E} f(\{e'\})$$

y como e es un elemento arbitrario en E, esta última desigualdad prueba que el conjunto solución de (4.5) es acotado.

Finalmente, como $B(f)$ está contenido en el conjunto solución de (4.5) concluímos que $B(f)$ es acotado.

□

Tenemos el siguiente teorema en relación a la solución del problema $\mathbf{P}_w$ de la ecuación (4.1).

Teorema 4.2.3. *Suponga que un sistema submodular $(\mathcal{D}, f)$ no es simple y que existe $F \in \Pi(\mathcal{D})$ tal que para algunos $e', e'' \in F$ se tiene $w(e') \neq w(e'')$, entonces el problema $\mathbf{P}_w$ no tiene solución óptima.*

Prueba: Sean e', e'' diferentes tales que $e', e'' \in F \in \Pi(\mathcal{D})$ y $w(e') \neq w(e'')$. Sea $x \in B(f)$ tal que x es solución del problema $\mathbf{P}_w$, es decir, la expresión

$$\sum_{e \in E} w(e)x(e) = \left\{ \sum_{\substack{e \in E \\ e \neq e', e''}} w(e)x(e) \right\} + w(e')x(e') + w(e'')x(e'')$$

es mínima en $B(f)$. Supongamos sin pérdida de generalidad que $w(e') < w(e'')$, entonces para $\alpha > 0$ se tiene:

$$\begin{aligned}
&\sum_{e \in E} w(e)\Big(x(e) + \alpha\big(\chi_{e'}(e) - \chi_{e''}(e)\big)\Big) \\
&\quad = \left\{ \sum_{\substack{e \in E \\ e \neq e', e''}} w(e)x(e) \right\} + w(e')\big(x(e') + \alpha\big) + w(e'')\big(x(e'') - \alpha\big) \\
&\quad = \left\{ \sum_{\substack{e \in E \\ e \neq e', e''}} w(e)x(e) \right\} + w(e')x(e') \ + \ w(e'')x(e'') \ + \ \big(w(e') - w(e'')\big)\alpha \\
&\quad < \sum_{e \in E} w(e)x(e)
\end{aligned}$$

pero esto es absurdo ya que $x + \alpha(\chi_{e'} - \chi_{e''}) \in B(f)$. Esto último se comprueba observando que para cada $X \in \mathcal{D}$, si $e' \in X$, entonces $e'' \in X$ (pues $e'' \in D(e')$), luego $x(X) + \alpha(\chi_{e'}(X) - \chi_{e''}(X)) = x(X) \leq f(X)$. De aquí, si $e' \notin X$ entonces

$$x(X) + \alpha(\chi_{e'}(X) - \chi_{e''}(X)) = x(X) - \alpha\chi_{e''}(X) \leq x(X) \leq f(X),$$

además, $x(E) + \alpha(\chi_{e'}(E) - \chi_{e''}(E)) = x(E)$.

□

Del teorema anterior, si $\mathbf{P}_w$ tiene alguna solución óptima entonces la función $w : E \longrightarrow \mathbb{R}$ debe ser constante en cada $F \in \Pi(\mathcal{D})$. El siguiente teorema nos muestra que para solucionar el problema $\mathbf{P}_w$ es suficiente considerar la simplificación del

sistema submodular $(\mathcal{D}, f)$.

Teorema 4.2.4. *Sea $(\mathcal{D}, f)$ un sistema submodular y suponga que en el problema $\mathbf{P}_w$ de la ecuación (4.1), w es constante en cada $F \in \Pi(\mathcal{D})$. Sea $(\widehat{\mathcal{D}}, \widehat{f})$ la simplificación de $(\mathcal{D}, f)$ y considere el problema:*

$$\widehat{\mathbf{P}}_w : \ \textit{Minimizar} \sum_{F \in \Pi(\mathcal{D})} w(F)\widehat{x}(F) \qquad \textit{sujeto a} \ \ \widehat{x} \in B(\widehat{f}) \tag{4.7}$$

donde $w(F)$ es el valor que toma w en F.

$\widehat{\mathbf{P}}_w$ tiene solución óptima sii $\mathbf{P}_w$ la tiene.

Prueba: Primero observamos lo siguiente:

$$\begin{aligned}
\widehat{x} \in B(\widehat{f}\,) &\Leftrightarrow \widehat{x} \in P(\widehat{f}\,) \quad \text{y} \quad \widehat{x}(\Pi(\mathcal{D})) = \widehat{f}(\Pi(\mathcal{D})) \\
&\Leftrightarrow \widehat{x}(\widehat{X}) \leq \widehat{f}(\widehat{X}) \text{ para todo } \widehat{X} \in \widehat{\mathcal{D}} \text{ y } \widehat{x}(\Pi(\mathcal{D})) = f(E) \\
&\Leftrightarrow \sum_{\substack{F \in \Pi(\mathcal{D}) \\ F \subseteq X}} \widehat{x}(F) \leq f(X) \text{ para todo } X \in \mathcal{D} \text{ y } \sum_{F \in \Pi(\mathcal{D})} \widehat{x}(F) = f(E)
\end{aligned} \tag{4.8}$$

Supongamos que hemos encontrado una solución óptima $\widehat{x}_*$ de $\widehat{\mathbf{P}}_w$, sea $\Pi(\mathcal{D}) = \{F_1, F_2, ..., F_m\}$ y $w_i = w(F_i)$ para cada $i = 1, 2, ..., m$, dada $x \in B(f)$

$$\begin{aligned}
\sum_{e \in E} w(e)x(e) &= \sum_{e \in F_1} w_1 x(e) + \cdots + \sum_{e \in F_m} w_m x(e) \\
&= w_1 x(F_1) + \cdots + w_m x(F_m)
\end{aligned} \tag{4.9}$$

y la función $\widehat{x} : \Pi(\mathcal{D}) \longrightarrow \mathbb{R}$ definida por $\widehat{x}(F) = x(F)$ para cada $F \in \Pi(\mathcal{D})$ está en $B(\widehat{f})$ ya que dado $X \in \mathcal{D}$, con $\widehat{X} = \{F_{i_1}, ..., F_{i_k}\}$ tenemos:

$$\widehat{x}(\widehat{X}) = \sum_{j=1}^{k} \widehat{x}(F_{i_j}) = \sum_{j=1}^{k} x(F_{i_j}) = x\left(\bigcup_{j=1}^{k} F_{i_j}\right) = x(X) \leq f(X) = \widehat{f}(\widehat{X})$$

es decir $\widehat{x} \in P(\widehat{f})$. Además

$$\widehat{x}(\Pi(\mathcal{D})) = \sum_{i=1}^{m} \widehat{x}(F_i) = \sum_{i=1}^{m} x(F_i) = x(E) = f(E) = \widehat{f}\,(\Pi(\mathcal{D}))$$

de aquí se concluye que $\widehat{x} \in B(\widehat{f})$, y con (4.9) tenemos que para $x \in B(f)$:

$$\sum_{e \in E} w(e)x(e) = \sum_{F \in \Pi(\mathcal{D})} w(F)\widehat{x}(F) \geq \sum_{F \in \Pi(\mathcal{D})} w(F)\widehat{x}_*(F). \tag{4.10}$$

Por otro lado, si definimos la función $x_* \in \mathbb{R}^E$ para cada $e \in E$ como $x_*(e) = \widehat{x}_*(F)/|F|$ donde $e \in F \in \Pi(\mathcal{D})$, entonces para cada $F \in \Pi(\mathcal{D})$, $x_*(F) = \widehat{x}_*(F)$. Además para cada $X \in \mathcal{D}$, si suponemos $\widehat{X} = \{F_{i_1}, ..., F_{i_k}\}$ tenemos

$$x_*(X) = \sum_{j=1}^{k} x_*(F_{i_j}) = \sum_{j=1}^{k} \widehat{x}_*(F_{i_j}) = \widehat{x}_*(\widehat{X}) \leq \widehat{f}(\widehat{X}) = f(X)$$

y

$$x_*(E) = \sum_{i=1}^{m} x_*(F_i) = \sum_{i=1}^{m} \widehat{x}_*(F_i) = \widehat{x}_*(\Pi(\mathcal{D})) = \widehat{f}(\Pi(\mathcal{D})) = f(E)$$

por lo tanto $x_* \in B(f)$. Por otro lado,

$$\begin{aligned}\sum_{e \in E} w(e)x_*(e) &= w_1 x_*(F_1) + \cdots + w_m x_*(F_m) \\ &= w_1 \widehat{x}_*(F_1) + \cdots + w_m \widehat{x}_*(F_m) = \sum_{F \in \Pi(\mathcal{D})} w(F)\widehat{x}_*(F),\end{aligned}$$

luego por la ecuación (4.10), x_* es una solución óptima del problema $\mathbf{P}_w$ de la ecuación (4.1).

Para probar el recíproco suponga que x_* es una solución óptima de $\mathbf{P}_w$, definimos $\widehat{x}_* : \Pi(\mathcal{D}) \longrightarrow \mathbb{R}$ por $\widehat{x}_*(F) = x_*(F)$ para cada $F \in \Pi(\mathcal{D})$. Para $X \in \mathcal{D}$ tenemos:

$$\sum_{\substack{F \in \Pi(\mathcal{D}) \\ F \subseteq X}} \widehat{x}_*(F) = \sum_{\substack{F \in \Pi(\mathcal{D}) \\ F \subseteq X}} x_*(X) = x_*(F) \leq f(X) \quad \text{y}$$

$$\sum_{F\in\,\Pi(\mathcal{D})} \widehat{x}_*(F) = \sum_{e\in E} x_*(e) = x_*(E) = f(E)$$

luego de (4.8), $\widehat{x}_* \in B(\widehat{f})$, además

$$\sum_{e\in E} w(e)x_*(e) = \sum_{F\in\,\Pi(\mathcal{D})} w(F)\widehat{x}_*(F).$$

Ahora, para $\widehat{x} \in B(\widehat{f})$, sea $x \in \mathbb{R}^E$ definida por $x(e) = \widehat{x}(F)/|F|$ donde $e \in F \in \Pi(\mathcal{D})$, obsérvese que para cada $F \in \Pi(\mathcal{D})$:

$$x(F) = \sum_{e\in F} x(e) = \sum_{e\in F} \frac{\widehat{x}(F)}{|F|} = |F|\frac{\widehat{x}(F)}{|F|} = \widehat{x}(F)$$

luego

$$\sum_{F\in\,\Pi(\mathcal{D})} w(F)\widehat{x}(F) = \sum_{F\in\,\Pi(\mathcal{D})} w(F)x(F) = \sum_{e\in E} w(e)x(e)$$

$$\leq \sum_{e\in E} w(e)x_*(e) = \sum_{F\in\,\Pi(\mathcal{D})} w(F)\widehat{x}_*(F)$$

con esto se concluye que $\widehat{x}_*$ es una solución óptima de $\widehat{\mathbf{P}}_w$.

□

En virtud del teorema anterior, podemos suponer de aquí en adelante (sin pérdida de generalidad) que en el problema $\mathbf{P}_w$ de la ecuación (4.1), el retículo $\mathcal{D}$ es simple, es decir, supondremos de aquí en adelante que $\mathcal{D} = 2^{\mathcal{P}}$ donde $\mathcal{P} = (E, \preceq)$ es un poset en E. El siguiente teorema muestra condiciones sobre $\mathcal{P}$ y w para que el problema $\mathbf{P}_w$ tenga solución óptima.

Teorema 4.2.5. *El problema $\mathbf{P}_w$ de la ecuación (4.1) tiene solución óptima sii $w : E \longrightarrow \mathbb{R}$ es una función monótona no decreciente, es decir, para todos $e, e' \in E$ se tiene $e \preceq e' \Rightarrow w(e) \leq w(e')$.*

Prueba: ($\Leftarrow$) Suponga que w es monótona no decreciente de $\mathcal{P} = (E, \preceq)$ en $(\mathbb{R}, \leq)$.

Sean $w_1 < w_2 < \cdots < w_p$ los distintos valores que toma w en E. Definimos

$$A_i := \{e \in E \mid w(e) \leq w_i\} \quad \text{para } i = 1, 2, ..., p. \tag{4.11}$$

Obsérvese que como w es monótona no decreciente, entonces cada A_i es un ideal de $\mathcal{P}$, luego los conjuntos A_i $(i = 1, 2, ..., p)$ forman una cadena $A_1 \subset A_2 \subset \cdots \subset A_p$ de $\mathcal{D}$.

Para cada $i = 1, 2, ..., p$ definamos

$$(\mathcal{D}_i, f_i) := \big((\mathcal{D}, f) \cdot A_i\big)/A_{i-1} \tag{4.12}$$

donde $A_0 := \varnothing$, es decir, $\mathcal{D}_i = \{X - A_{i-1} \mid X \in \mathcal{D}, X \subseteq A_i\}$ y $f_i(X - A_{i-1}) = f(X) - f(A_{i-1})$ $(X - A_{i-1} \in \mathcal{D}_i)$.

Tomemos una base x_i de $(\mathcal{D}_i, f_i)$ para cada $i = 1, 2, ..., p$, entonces por el lema 3.2.2, $x = x_1 \oplus x_2 \oplus \cdots \oplus x_p$, donde la suma $\oplus$ se definió en la ecuación (3.3), es una base de $(\mathcal{D}, f)$ y en particular cumple:

$$x(A_i) = f(A_i) \quad \text{para cada } i = 1, 2, ..., p. \tag{4.13}$$

Para cada $y \in B(f)$ tenemos:

$$\begin{aligned}
\sum_{e \in E} w(e)y(e) - \sum_{e \in E} w(e)x(e) &= \sum_{i=1}^{p} \sum_{e \in A_i - A_{i-1}} w_i(y(e) - x(e)) \\
&= \sum_{i=1}^{p} \Big\{ w_i\big(y(A_i) - x(A_i)\big) - w_i\big(y(A_{i-1}) - x(A_{i-1})\big) \Big\} \\
&= \sum_{i=1}^{p-1} \Big\{ w_i\big(y(A_i) - x(A_i)\big) - w_i\big(y(A_{i-1}) - x(A_{i-1})\big) \Big\} \\
&\qquad + w_p\big(y(A_p) - x(A_p)\big) - w_p\big(y(A_{p-1}) - x(A_{p-1})\big) \\
&= \sum_{i=1}^{p-1} \Big\{ (w_i - w_{i+1})\big(y(A_i) - x(A_i)\big) \Big\} + w_p\big(y(A_p) - x(A_p)\big)
\end{aligned}$$

$$= \sum_{i=1}^{p-1} (w_{i+1} - w_i)\big(f(A_i) - y(A_i)\big) \geq 0$$

Esto muestra la optimalidad de la función x.

($\Rightarrow$) Suponga que $\mathbf{P}_w$ tiene solución óptima. Si w no es monótona no decreciente, existen elementos $e, e' \in E$ tales que $e' \preceq e$ y $w(e') > w(e)$, luego para cada $X \in \mathcal{D}$, con $e \in X$, tenemos que $e' \in X$ (pues X es ideal de $\mathcal{P}$). Así, dados $x \in B(f)$, $X \in \mathcal{D}$ y $\alpha > 0$:

i) si $e \in X$, $x(X) + \alpha(\chi_e(X) - \chi_{e'}(X)) = x(X) \leq f(X)$

ii) si $e \notin X$, $x(X) + \alpha(\chi_e(X) - \chi_{e'}(X)) \leq x(X) \leq f(X)$

además $x(E) + \alpha(\chi_e(E) - \chi_{e'}(E)) = f(E)$

En conclusión, la función $x + \alpha(\chi_e - \chi_{e'}) \in B(f)$ para cada $\alpha > 0$. Así mismo, como $w(e) < w(e')$ entonces:

$$\begin{aligned}
&\sum_{a \in E} w(a)\Big(x(a) + \alpha\big(\chi_e(a) - \chi_{e'}(a)\big)\Big) \\
&\qquad = \left\{ \sum_{\substack{a \in E \\ a \neq e, e'}} w(a)x(a) \right\} + w(e)\big(x(e) + \alpha\big) + w(e')\big(x(e') - \alpha\big) \\
&\qquad = \left\{ \sum_{\substack{a \in E \\ a \neq e, e'}} w(a)x(a) \right\} + w(e)x(e) \ + \ w(e')x(e') \ + \ \big(w(e) - w(e')\big)\alpha \\
&\qquad = \left\{ \sum_{a \in E} w(a)x(a) \right\} + \big(w(e) - w(e')\big)\alpha \\
&\qquad < \sum_{a \in E} w(a)x(a) \qquad (4.14)
\end{aligned}$$

luego el problema $\mathbf{P}_w$ no tendría solución óptima, lo cual es absurdo.

□

Teorema 4.2.6. *Suponga que w es una función monótona no decreciente de $\mathcal{P} = (E, \preceq)$ a $(\mathbb{R}, \leq)$. Para cada $i = 1, 2, ..., p$, sean A_i y $(\mathcal{D}_i, f_i)$ definidos por las fórmulas (4.11) y (4.12). El conjunto de soluciones óptimas del problema $\mathbf{P}_w$ están dadas por:*

$$B(f_1) \oplus B(f_2) \oplus \cdots \oplus B(f_p)$$
$$= \{x_1 \oplus x_2 \oplus \cdots \oplus x_p \mid x_i \in B(f_i)\ (i = 1, 2, ..., p)\}$$

donde la suma directa $\oplus$ está definida por la fórmula (3.3). Esto es, $x \in B(f)$ es solución óptima de $\mathbf{P}_w$ sii x restringido a $A_i - A_{i-1}$ es una base de $\big((\mathcal{D}, f) \cdot A_i\big)/A_{i-1}$, para cada $i = 1, 2, ..., p$.

Prueba: Se sigue de la prueba de "$\Leftarrow$"del teorema 4.2.5 que $x \in B(f_1) \oplus B(f_2) \oplus \cdots \oplus B(f_p)$ es una solución óptima de $\mathbf{P}_w$.

Recíprocamente, suponga que x es una solución óptima de $\mathbf{P}_w$, tenemos

$$dep(x, e) \subseteq A_i \quad \text{para cada } e \in A_i,\ i = 1, 2, ..., p. \tag{4.15}$$

en efecto, sea $e' \in dep(x, e)$, si $e' \notin A_i$, entonces $w(e') > w_i \geq w(e)$. Sea

$$\alpha = \text{mín}\{f(X) - x(X) \mid e \in X, e' \notin X\} = \tilde{c}(x, e, e')$$

$\alpha > 0$ ya que si existiese X tal que $f(X) = x(X)$ con $e \in X, e' \notin X$, tendríamos $dep(x, e) \subseteq X$ con lo cual $e' \notin dep(x, e)$, lo que es absurdo. Ahora, $x + \alpha(\chi_e - \chi_{e'}) \in B(f)$, ya que dado $X \in \mathcal{D}$ tenemos dos casos:

Caso 1. $e \in X$.
Si $e' \in X$ entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) \leq f(X)$.
Si $e' \notin X$ entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) + \alpha \leq f(X)$
(pues $\alpha \leq f(X) - x(X)$)

Caso 2. $e \notin X$.
Si $e' \in X$ entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) - \alpha \leq f(X)$.

Si $e' \notin X$ entonces $x(X) + \alpha(\chi_e - \chi_{e'})(X) = x(X) \leq f(X)$.

y claramente, $x(E) + \alpha(\chi_e - \chi_{e'})(E) = x(E) = f(E)$

como $\alpha > 0$ y $w(e) - w(e') < 0$, utilizando el mismo razonamiento que se utilizó en la fórmula (4.14) se puede concluir que

$$\sum_{a \in E} w(a)\Big(x(a) + \alpha\big(\chi_e(a) - \chi_{e'}(a)\big)\Big) < \sum_{a \in E} w(a)x(a)$$

lo que contradice la optimalidad de x. De aquí que (4.15) se tiene.

El anterior razonamiento nos permite escribir $A_i = \bigcup\{dep(x, e) \mid e \in A_i\}$, pero recordemos que por el lema 3.1.2, el conjunto $D(x) = \{X \in \mathcal{D} \mid x(X) = f(X)\}$ es un retículo distributivo, luego se debe tener que $x(A_i) = f(A_i)$ para todo i. Por el lema 3.2.2, concluimos que $x \in B(f_1) \oplus B(f_2) \oplus \cdots \oplus B(f_p)$.

□

Teorema 4.2.7. *Una base $x \in B(f)$ es una solución óptima del problema $\mathbf{P}_w$ sii para cada $e, e' \in E$ tales que $e' \in dep(x, e)$ se tiene $w(e) \geq w(e')$.*

Prueba: ($\Rightarrow$) Sea $x \in B(f)$ solución óptima y sea i tal que $w(e) = w_i$. De la prueba del teorema 4.2.6 tenemos que $dep(x, e) \subseteq A_i$, luego si $e' \in dep(x, e)$, entonces $e' \in A_i$, esto es, $w(e) \leq w_i = w(e)$.

($\Leftarrow$) Sea $i \in \{1, 2, ..., p\}$ y $e \in A_i$ tal que $w(e) = w_i$, si $e' \in dep(x, e)$ entonces $w_i = w(e) \geq w(e')$, es decir $e' \in A_i$, o en otras palabras, $dep(x, e) \subseteq A_i$, así, $A_i = \bigcup\{dep(x, e) \mid e \in A_i\}$. De nuevo como

$$D(x) = \{X \in \mathcal{D} \mid x(X) = f(X)\}$$

es un retículo distributivo, entonces $x(A_i) = f(A_i)$ para todo $i = 1, 2, ..., p$, es decir, $x \in B(f_1) \oplus B(f_2) \oplus \cdots \oplus B(f_p)$ y por el teorema 4.2.6, x es solución óptima.

□

La prueba del teorema 4.2.5 nos conduce a un algoritmo, llamado el *Algoritmo Greedy* para resolver el problema $\mathbf{P}_w$. Este algoritmo fue implementado por Edmonds para Polimatroides y por Lovász para sistemas submodulares en $\mathcal{D} = 2^E$. Antes de presentar el algoritmo necesitamos algunos conceptos previos.

Definición 4.2.1. *Una sucesión $(e_1, e_2, ...e_n)$ que contiene a todos los elementos de E se denomina una extensión lineal de un poset $\mathcal{P} = (E, \preceq)$ si $e_i \preceq e_j$ implica $i \leq j$ $(i, j = 1, 2, ..., n)$. Una extensión lineal $(e_1, e_2, ...e_n)$ de $\mathcal{P} = (E, \preceq)$ es llamada monótona no decreciente respecto a $w : E \longrightarrow \mathbb{R}$, si $w(e_1) \leq w(e_2) \leq \cdots \leq w(e_n)$.*

La existencia de una extensión lineal de un poset $\mathcal{P} = (E, \preceq)$ está garantizada por un importante resultado llamado el principio de extensión de orden (*Order-extension principle*) [9]. Cuando el cardinal del conjunto E es pequeño, encontrar una extensión lineal no es tarea difícil, no obstante, en [3] se pueden encontrar diversos algoritmos utilizados para encontrar extensiones lineales.

Observamos que si $w : E \longrightarrow \mathbb{R}$ es una función monótona no decreciete de $\mathcal{P} = (E, \preceq)$ a $(\mathbb{R}, \leq)$, entonces existe una extensión lineal monótona no decreciente de $\mathcal{P}$ respecto a w. Basta definir $B_i := \{e \in E \ \mid w(e) = w_i\}$ para cada i, donde $w_1 < w_2 < \cdots < w_p$ son los diferentes valores que toma w en E, en seguida se encuentra una extensión lineal $(e_{i_1}, e_{i_1}, ..., e_{i_{B_i}})$ de $\mathcal{P}$ en B_i para cada $i = 1, 2, ...p$. La extensión linea monótona no decreciente será

$$(e_{1_1}, e_{1_2}, ..., e_{1_{B_1}}, e_{2_1}, e_{2_2}, ..., e_{2_{B_2}} ..., e_{p_1}, e_{p_2}, ..., e_{p_{B_p}})$$

Teorema 4.2.8. *(Algoritmo Greedy) Considere el problema $\mathbf{P}_w$ y suponga que la función de peso w es monótona no decreciente de $\mathcal{P} = (E, \preceq)$ a $(\mathbb{R}, \leq)$.*

Paso 1: *Encontremos una extensión lineal monótona no decreciente de $\mathcal{P}$ respecto a w.*

Paso 2: *Definamos el vector* $x \in \mathbb{R}^E$ *por*

$$x(e_i) = f(S_i) - f(S_{i-1}) \quad \textit{para } i = 1, 2, ..., n \tag{4.16}$$

donde para cada $i = 1, 2, ..., n$, $S_i := \{e_1, e_2, ...e_i\}$ *y* $S_0 := \varnothing$.

Entonces x *es una base de mínimo peso de* $(\mathcal{D}, f)$ *respecto a* w.

Prueba: Veamos que $x \in B(f)$, es decir $x(X) \leq f(X)$ para todo $X \in \mathcal{D}$ y $x(E) = f(E)$. Tomemos entonces $X \in \mathcal{D}$, para entender mejor la prueba supondremos primero que X tiene un elemento, después supondremos que tiene dos y tres elementos para probar después el caso general. Si X tiene un elemento, es decir, $X = \{e_k\}$ para algún $k \in \{1, 2, ..., n\}$

$$x(e_k) = f(S_k) - f(S_{k-1}) = f(S_{k-1} \cup \{e_k\}) - f(S_{k-1}) \tag{4.17}$$

por el teorema 1.1.1 (éste teorema es cláramente válido no sólo para funciones submodulares en 2^E sino en cuaquier conjunto $\mathcal{D} \subseteq 2^E$, siempre y cuando los conjuntos involucrados sean elementos de $\mathcal{D}$) tenemos:

$$f(S_{k-1} \cup \{e_k\}) - f(S_{k-1}) \leq f(\varnothing \cup \{e_k\}) - f(\varnothing) = f(\{e_k\}) \tag{4.18}$$

de (4.17) y (4.18) concluímos que $x(X) \leq f(X)$.

Supongamos ahora que X tiene dos elementos, es decir, supongamos que $X = \{a, b\}$ donde a aparece primero que b en la extensión lineal. Existen $A, B \in \{S_1, S_2, ...S_n\}$ con $A \subseteq B$ tales que

$$x(a) = f(A \cup \{a\}) - f(A) \quad \text{y} \quad x(b) = f(B \cup \{b\}) - f(B) \tag{4.19}$$

hagamos $S := A \cup \{a, b\} = A \cup X$. Como f es submodular, se debe tener $f(S \cap B) + f(S \cup B) \leq f(S) + f(B)$, es decir,

$$f(A \cup \{a\}) + f(B \cup \{b\}) \leq f(A \cup \{a, b\}) + f(B) \tag{4.20}$$

de aquí,

$$\begin{aligned} x(X) = x(a) + x(b) &= f(A \cup \{a\}) - f(A) + f(B \cup \{b\}) - f(B) \\ &\leq f(A \cup \{a, b\}) - f(A) \end{aligned}$$

y de nuevo por el teorema 1.1.1, $f(A \cup \{a, b\}) - f(A) \leq f(\{a, b\})$, en otras palabras $x(X) \leq f(X)$.

Si X tiene tres elementos se sigue un procedimiento parecido: Supongamos que $X = \{a, b, c\}$ en donde a está primero que b y b primero que c en la extensión lineal, existen $A, B, C \in \{S_1, S_2, ...S_n\}$ con $A \subseteq B \subseteq C$ tales que

$$x(a) = f(A \cup \{a\}) - f(A) \ , \quad x(b) = f(B \cup \{b\}) - f(B) \quad \text{y}$$

$$x(c) = f(C \cup \{c\}) - f(C)$$

Hagamos $X_B := B \cup \{b, c\} = B \cup X$ y $X_A := A \cup \{a, b, c\} = A \cup X$, se debe tener:

$$\begin{aligned} &f(X_B \cap C) + f(X_B \cup C) \leq f(X_B) + f(C) \quad \text{y} \\ &f(X_A \cap B) + f(X_A \cup B) \leq f(X_A) + f(B) \end{aligned}$$

es decir,

$$\begin{aligned} &f(B \cup \{b\}) + f(C \cup \{c\}) \leq f(B \cup \{b, c\}) + f(C) \quad \text{y} \\ &f(A \cup \{a\}) + f(B \cup \{b, c\}) \leq f(A \cup \{a, b, c\}) + f(B) \end{aligned}$$

luego de lo anterior, junto con el teorema 1.1.1, obtenemos:

$$\begin{aligned} x(X) &= x(\{a, b, c\}) \\ &= f(A \cup \{a\}) - f(A) + f(B \cup \{b\}) - f(B) + f(C \cup \{c\}) - f(C) \\ &\leq f(A \cup \{a\}) - f(A) + f(B \cup \{b, c\}) - f(B) \\ &\leq f(A \cup \{a, b, c\}) - f(A) \\ &\leq f(\{a, b, c\}) = f(X) \end{aligned}$$

Ahora supongamos que $X = \{e_{i_1}, e_{i_2}, ..., e_{i_m}\}$ donde $(e_{i_1}, e_{i_2}, ..., e_{i_m})$ es una subsucesión de $(e_1, e_2, ..., e_n)$. Hagamos

$$X_k = S_{(i_k - 1)} \cup \{e_{i_k}, e_{i_{(k+1)}}, \dots, e_{i_m}\} = S_{(i_k - 1)} \cup X$$

para $k = 1, 2, ..., m-1$. Obsérvese que $X_{i_k} \in \mathcal{D}$ para todo k, luego como f es submodular, para $k = 1, 2, ..., m-1$ tenemos:

$$f(X_k \cap S_{\{i_{(k+1)}-1\}}) + f(X_k \cup S_{\{i_{(k+1)}-1\}}) \leq f(X_k) + f(S_{\{i_{(k+1)}-1\}})$$

es decir,

$$f(S_{i_k}) + f(X_{k+1}) \leq f(X_k) + f(S_{\{i_{(k+1)}-1\}})$$

y a partir de esto tenemos:

$$\begin{aligned}
x(X) &= \sum_{k=1}^{m} x(e_{i_k}) = \sum_{k=1}^{m} f(S_{i_k}) - f(S_{i_k-1}) \\
&= f(S_{i_1}) - f(S_{(i_1-1)}) + ... + f(S_{i_{(m-1)}}) - f(S_{(i_{(m-1)}-1)}) + f(S_{i_m}) - f(S_{(i_m-1)}) \\
&\leq f(S_{(i_1-1)} \cup X) - f(S_{(i_1-1)}) \\
&\leq f(X)
\end{aligned}$$

Por otro lado, como $(e_1, e_2, ..., e_n)$ es una extensión lineal monótona no decreciente de $\mathcal{P}$ respecto a w, entonces cada conjunto A_i definido en la fórmula (4.11) se puede escribir como $A_i = \{e_1, e_2, ...e_r\}$ para algún $r \in \{1, 2, ..., n\}$, entonces

$$x(A_i) = \sum_{k=1}^{m} x(e_k) = \sum_{k=1}^{m} f(S_k) - f(S_{k-1}) = f(S_m) = f(A_i),$$

en particular, $x(E) = f(E)$. Por el teorema 4.2.6, x es una solución óptima del problema $\mathbf{P}_w$.

□

Observamos que siendo $(e_1, e_2, ..., e_n)$ una extensión lineal monótona no decreciente, entonces cada S_i es un ideal de $\mathcal{P}$, o equivalentemente, $S_i \in \mathcal{D}$ para cada i, luego $\varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E$ es una cadena maximal en $\mathcal{D}$ que además contiene a los elementos A_i definidos en (4.11). Recíprocamente, cada cadena maximal $\mathcal{C} : \varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E$ de $\mathcal{D}$ proporciona una extensión lineal monótona no decreciente $(e_1, e_2, ..., e_n)$ de $\mathcal{P}$, definiendo $\{e_i\} = S_i - S_{i-1}$ para $i = 1, ..., n$, pues

siendo S_j un ideal de $\mathcal{P}$ para cada $j \in \{1, ..., n\}$, entonces $e_i \preceq e_j$ implica $e_i \in S_j$, es decir, $i \leq j$.

Corolario 4.2.9. *El problema* $\mathbf{P}_w$ *tiene única solución óptima si* $w : E \longrightarrow \mathbb{R}$ *es una función monótona no decreciente uno a uno de* $\mathcal{P} = (E, \preceq)$ *a* $(\mathbb{R}, \leq)$.

Teorema 4.2.10. *Sea* $f : \mathcal{D} \longrightarrow \mathbb{R}$ *una función definida en un retículo distributivo simple* $\mathcal{D} \subseteq 2^E$ *tal que* $\varnothing, E \in \mathcal{D}$ *y* $f(\varnothing) = 0$. *Definimos*

$$B(f) := \{x \in \mathbb{R}^E \mid \forall X \in \mathcal{D}, x(X) \leq f(X), x(E) = f(E)\} \tag{4.21}$$

entonces el algoritmo greedy descrito anteriormente funciona para $B(f)$ *definida en (4.21) sii* f *es una función submodular.*

Prueba: Es suficiente mostrar la parte " $\Rightarrow$ ". Suponga que el algoritmo funciona para $B(f)$, para cada cadena maximal

$$\mathcal{C} : \varnothing = S_0 \subset S_1 \subset \cdots \subset S_n = E \tag{4.22}$$

de $\mathcal{D}$, el vector $x \in \mathbb{R}^E$ definido por (4.16) pertenece a $B(f)$ (por hipótesis), además, $x(S_i) = f(S_i)$ para cada $i = 1, ..., n$.

Dados $X, Y \in \mathcal{D}$ escoja una cadena maximal $\mathcal{C}$ de $\mathcal{D}$ que contenga a los elementos $X \cup Y$ y $X \cap Y$,y defina x por (4.16), como $x \in B(f)$,

$$x(X) \leq f(X), \ x(Y) \leq f(Y),$$

$$x(X \cup Y) = f(X \cup Y) \ \text{ y } \ x(X \cap Y) = f(X \cap Y)$$

luego

$$\begin{aligned} f(X) + f(Y) &\geq x(X) + x(Y) \\ &= x(X \cup Y) + x(X \cap Y) \\ &= f(X \cup Y) + f(X \cap Y), \end{aligned}$$

concluímos que f es submodular

$\square$

Ejemplo 4.2.1. *(Solución de un problema de optmización lineal utilizando el algoritmo greedy)*

Para un mayor entendimiento de los resultados de esta sección, solucionaremos un problema de optimización en particular utilizando los procedimientos descritos aquí, así como el algoritmo greedy del teorema 4.2.8.

Consideremos el sistema submodular $(\mathcal{D}, f)$ sobre el conjunto $E = \{1, 2, 3, 4, 5\}$ en donde $\mathcal{D} = \{\varnothing, E, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 4, 5\}\}$ y $f : \mathcal{D} \longrightarrow \mathbb{R}$ está definida como:

$$f(\varnothing) = 0, \quad f(E) = 4, \quad f(\{1, 2\}) = 1, \quad f(\{1, 2, 3\}) = 3 \text{ y } f(\{1, 2, 4, 5\}) = 3$$

Se puede verificar fácilmente qe f es una función submodular en $\mathcal{D}$. El grafo asociado a $\mathcal{D}$ junto con su descomposición se observa en la siguiente figura:

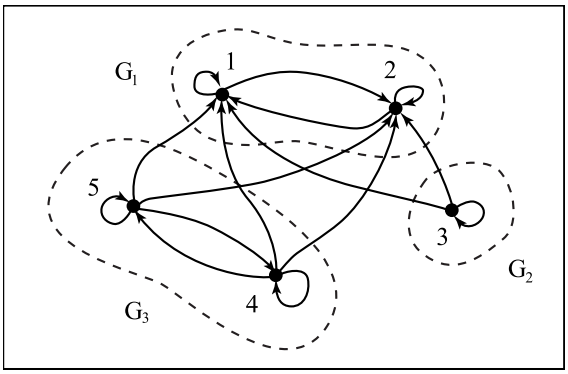

Figura 4.2: Descomposición del grafo asociado a $\mathcal{D}$.

Se tiene entonces $\Pi(\mathcal{D}) = \{\{1, 2\}, \{3\}, \{4, 5\}\}$. Sean $a = \{1, 2\}, b = \{3\}$ y $c = \{4, 5\}$, entonces la relación de orden $\preceq_{\mathcal{D}}$ en $\Pi(\mathcal{D})$ está determinada por:

$$a \preceq_{\mathcal{D}} a, \quad b \preceq_{\mathcal{D}} b, \quad c \preceq_{\mathcal{D}} c, \quad a \preceq_{\mathcal{D}} b \quad \text{y} \quad a \preceq_{\mathcal{D}} c$$

definimos la función de peso $w : \mathcal{D} \longrightarrow \mathbb{R}$ por:

$$w(1) = 2, \quad w(2) = 2, \quad w(3) = 5, \quad w(4) = 7 \quad \text{y} \quad w(5) = 7$$

Consideremos el problema $\mathbf{P}_w$ de la ecuación (4.1) para el sistema submodular $(\mathcal{D}, f)$ y la función de peso w aquí definidos, es decir, consideremos el problema:

$$\mathbf{P}_w : \text{ Minimizar} \sum_{i=1}^{5} w(i)x(i) \qquad \text{sujeto a } \ x \in B(f) \tag{4.23}$$

si escribimos $x(i) = x_i$ para $i = 1, 2, 3, 4, 5$ entonces este problema es equivalente a:

$$\text{Minimizar } 2x_1 + 2x_2 + 5x_3 + 7x_4 + 7x_5$$

sujeto a:

$$\begin{gathered} x_1 + x_2 \leq 1 \\ x_1 + x_2 + x_3 \leq 3 \\ x_1 + x_2 + x_4 + x_5 \leq 3 \\ x_1 + x_2 + x_3 + x_4 + x_5 = 4. \end{gathered}$$

Siendo $(\mathcal{D}, f)$ un sistema submodular no simple, consideremos entonces la simplificación $(\widehat{\mathcal{D}}, \widehat{f})$ de $(\mathcal{D}, f)$ para así utilizar el teorema 4.2.4 y su demostración.

Si hacemos $X_0 = \varnothing, X_1 = E, X_2 = \{1, 2\}, X_3 = \{1, 2, 3\}$ y $X_4 = \{1, 2, 4, 5\}$, entonces los elementos de $\widehat{\mathcal{D}}$ (obsérvese el teorema 4.1.4) serán:

$$\widehat{X}_0 = \varnothing, \quad \widehat{X}_1 = \{a, b, c\}, \quad \widehat{X}_2 = \{a\}, \quad \widehat{X}_3 = \{a, b\}, \quad \widehat{X}_4 = \{a, c\}$$

y $\widehat{f} : \widehat{\mathcal{D}} \longrightarrow \mathbb{R}$ estará definida por:

$$\begin{gathered} \widehat{f}(\varnothing) = f(\varnothing) = 0, \quad \widehat{f}(\{a, b, c\}) = f(X_1) = 4, \quad \widehat{f}(\{a\}) = f(X_2) = 1, \\ \widehat{f}(\{a, b\}) = f(X_3) = 3, \quad \text{y} \quad \widehat{f}(\{a, c\}) = f(X_4) = 3, \end{gathered}$$

luego el problema $\widehat{\mathbf{P}}_w$ de la ecuación 4.7 para este caso en particular será:

$$\widehat{\mathbf{P}}_w : \text{ Minimizar } w(a)\widehat{x}(a) + w(b)\widehat{x}(b) + w(c)\widehat{x}(c) \quad \text{sujeto a } \ \widehat{x} \in B(\widehat{f}) \tag{4.24}$$

escribiendo $\widehat{x}(a) = \widehat{x}_a, \widehat{x}(b) = \widehat{x}_b$ y $\widehat{x}(c) = \widehat{x}_c$, el problema $\widehat{\mathbf{P}}_w$ es equivalente a:

$$\widehat{\mathbf{P}}_w : \text{Minimizar } 2\widehat{x}_a + 5\widehat{x}_b + 7\widehat{x}_c$$

sujeto a:

$$\begin{aligned} \widehat{x}_a &\leq 1 \\ \widehat{x}_a + \widehat{x}_b &\leq 3 \\ \widehat{x}_a + \widehat{x}_c &\leq 3 \\ \widehat{x}_a + \widehat{x}_b + \widehat{x}_c &= 4. \end{aligned}$$

Obsérvese que $w : \widehat{\mathcal{D}} \longrightarrow \mathbb{R}$ es monótona no decreciente de $\mathcal{P} = (\Pi(\mathcal{D}), \preceq_{\mathcal{D}})$ a $(\mathbb{R}, \leq)$.

Utilicemos ahora el algoritmo greedy del teorema 4.2.8. Siendo (a, b, c) una extensión lineal monótona no decreciente de $\mathcal{P} = (\Pi(\mathcal{D}), \preceq_{\mathcal{D}})$ respecto a w, entonces, siguiendo el paso 2 del algoritmo, obtenemos la base de mínimo peso $\widehat{x}_*$ de $(\widehat{\mathcal{D}}, \widehat{f})$ respecto a w definida por:

$\widehat{x}_*(a) = \widehat{f}(\{a\}) = 1$

$\widehat{x}_*(b) = \widehat{f}(\{a, b\}) - \widehat{f}(\{a\}) = 3 - 1 = 2$

$\widehat{x}_*(c) = \widehat{f}(\{a, b, c\}) - \widehat{f}(\{a, b\}) = 4 - 3 = 1,$

finalmente, siguiendo los pasos de la demostración del teorema 4.2.4, una base de mínimo peso x_* de $\mathbf{P}_w$ estará definida como:

$x_*(1) = \widehat{x}_*(a)/|a| = 1/2$

$x_*(2) = \widehat{x}_*(a)/|a| = 1/2$

$x_*(3) = \widehat{x}_*(b)/|b| = 2/1 = 2$

$x_*(4) = \widehat{x}_*(c)/|c| = 1/2$

$x_*(5) = \widehat{x}_*(c)/|c| = 1/2$

en otras palabras, una solución óptima x_* de $\mathbf{P}_w$ es:

$$(x_*(1), x_*(2), x_*(3), x_*(4), x_*(5)) = (\tfrac{1}{2}, \tfrac{1}{2}, 2, \tfrac{1}{2}, \tfrac{1}{2})$$

Aportes del Autor

Como se dijo anteriormente, este trabajo se basa en el estudio realizado por Satoru Fujishije en los capítulos I y II de [7]. Teniendo en cuenta que muchos de los resultados que allí aparecen o son mencionados, no todos están demostrados o detallados en su totalidad, entonces aquí se realizan con detalle las demostraciones de tales resultados. A continuación se menciona en su totalidad el trabajo realizado por el autor.

Se entiende por el autor, al estudiante Arley Ramsés Gómez Rios, persona quien escribió el presente trabajo.

Capítulo 1.

Las demostraciones de los teoremas 1.1.1 y 1.1.2 son hechas por el autor, las demás demostraciones se pueden encontrar utilizando la bibliografía.

Capítulo 2.

Sección 2.1: La demostración de la proposición 2.1.1 la realiza Satoru en [7], aunque es detallada por el autor, las demás demostraciones (Proposiciones 2.1.2, 2.1.3, 2.1.4, 2.1.5 y 2.1.6) son hechas en su totalidad por el autor.

Sección 2.2: Aunque en [7] se mencionan los ejemplos de esta sección, el autor muestra con detalle cada afirmación hecha por Satoru en cada uno de estos ejemplos.

Sección 2.3: Satoru en [7] afirma que existe una relación entre matroides y polimatroides, afirma además que la función de saturación es una generalización de la función clausura de un matroide y que la función de dependencia generaliza la noción

de circuito fundamental. Teniendo en cuenta que Satoru no muestra dichas relaciones, el autor las establece y demuestra en cada uno de los resultados y afirmaciones de esta sección.

Capítulo 3.

Sección 3.1: Las pruebas de los lemas 3.1.1, 3.1.2, 3.1.3, 3.1.7 y el teorema 3.1.5 se encuentran en [7], aunque son detalladas por el autor. Las demás pruebas (proposiciones 3.1.4 y 3.1.6) son hechas en su totalidad por el autor.

Sección 3.2: Todos los resultados de esta sección son demostrados en [7], aunque sus demostraciones fueron detalladas por el autor.

Sección 3.3: Cada ejemplo de esta sección es mencionado por Satoru en [7] aunque muchas afirmaciones que allí no son demostradas, aquí son mostradas en detalle por el autor.

Capítulo 4.

Sección 4.1: La prueba del teorema 4.1.1 se encuentra en [7] aunque fué detallada por el autor. Los teoremas 4.1.2, 4.1.3 y 4.1.4 se mencioan en [7] aunque allí no son probados, las pruebas de dichos teoremas fueron hechas por el autor.

Sección 4.2: Las pruebas de los teoremas 4.2.1, 4.2.2, 4.2.5, 4.2.6, 4.2.7 y 4.2.10 se encuentran en [7] aunque son detalladas por el autor. Los teoremas 4.2.3 y 4.2.4 son mencionados en [7] aunque allí no son demostrados. Éstas pruebas son hechas por el autor.

El funcionamiernto del Algoritmo Greedy (Teorema 4.2.8) también fue probado por el autor ya que en [7] no está demostrado.

El ejemplo 4.2.1 fue desarrollado por el autor.

Conclusiones

El hecho de que se asocie el concepto de matroide con conjuntos poliedrales a través de la submodularidad de la función rango de un matroide hizo emerger las aplicaciones de las funciones submodulares, es por ello que este hecho es bastante importante.

Las características del retículo distributivo asociado a un sistema submodular, así como su grafo asociado, juegan un papel muy importante en la determinación de las condiciones para que el problema de optimización lineal del capítulo 4 tenga solución óptima.

Es interesante el hecho de que siendo el concepto de sistema submodular una generalización del concepto de matroide pueda ser utilizado para solucionar problemas de optimización lineal a través de la función submodular asociada.

Trabajo Futuro

Encontrar aplicaciones de la función de saturación y la función de dependencia, ya sea en teoría de matroides o en otras ramas de la matemática como teoría de grafos, redes, etc.

Encontrar el papel que juegan los conceptos de capacidad de saturación y capacidad de intercambio ya sea en teoría de matroides o en optimización.

Encontrar más ejemplos de sistemas submodulares y polimatroides.

Encontrar condiciones para que un problema de optimización lineal arbitrario pueda ser asociado a un sistema submodular y así poder ser solucionado por el algoritmo greedy. En tal caso sería interesante compararlo con el método simplex.

Especular sobre qué tipo de problemas reales pueden ser solucionados por el algoritmo greedy.

Discutir la voracidad del algoritmo greedy presentado aquí.

Bibliografía

[1] Gilles Brassard y Paul Bratley. *Algoritmos voraces*. Fundamentos de Algoritmia. Prentice Hall, Madrid, 1997.

[2] G. Choquet. *Theory of capacities*. Annals de l′institut Fourier 5. 1955.

[3] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, y Clifford Stein. *Introduction to Algorithms*. MIT press and McGraw-Hill, second ed$^{\underline{\text{ón}}}$., 2001.

[4] W.H. Cunningham. *Submodular optimization*. Presented at the Bielefeld Workshop on Matroids. August 1983.

[5] J. Edmonds. *Submodular functions, matroids, and certain polyhedra*. Combinatorial Structures and Their Applications, New York, 1970.

[6] A. Frank. *An algorithm for submodular functions on graphs*. Annals of Discrete Mathematics 16. 1982.

[7] Satoru Fujishige. *Submodular functions and optimization*. Annals of discrete mathematics. Second ed$^{\underline{\text{ón}}}$., 2005.

[8] L. Lováz. *Submodular functions and convexity*. In: Mathematical Programming - The State of the Art, A. Bachem, M. Grötschel and B. Korte, eds. Springer, Berlin, 1983.

[9] Edward Marczewski. *Sur léxtension de lórdre partiel*. Fundamenta Mathematicae. 1930.

[10] James G Oxley. *Matroid theory*. Oxford graduate text in mathematics, 1992.

[11] R. T. Rockafellar. *Convex analysis*. Princeton University Press, N.J., 1970.

[12] A. Schrijver. *Combinatorial Optimization*. Polyhedra and Efficiency. Springer, Heidelberg, 2003.

[13] M. Shubik. *Game theory in the social sciences*. Concepts and Solutions. The MIT Press, Cambridge, MA, 1982.