

5.1. EVIDENCIA EN LAS INTERPRETACIONES DE LOS SISTEMAS DISCRETOS

Los sistemas discretos, a cuyos métodos de estudio se dedica este libro, existen en una diversidad exclusivamente grande. La apreciación del carácter discreto de tal o cual sistema se forma o bien debido a la comprensión de la existencia separada y discreta de los elementos del sistema dado, o bien se predetermina por el carácter discreto de la información recibida (por ejemplo, lecturas de los instrumentos). Correspondientemente, los métodos matemáticos de investigación de los sistemas semejantes se caracterizan por la misma diversidad.

Los modelos matemáticos de los sistemas discretos son, con mayor frecuencia, conjuntos de símbolos dispuestos en la forma discreta, habitualmente de puntos, unidos con líneas, si esto está determinado por las condiciones del problema.

A los modelos combinatorios se les atribuye una evidencia máxima posible no sólo para facilitar la comprensión de tal o cual problema. Los medios ilustrativos o didácticos del análisis combinatorio contribuyen frecuentemente a crear las posibilidades para resolver los problemas. De ejemplo puede servir el teorema de Ramsey (véase § 3.3).

Recordemos que el teorema de Ramsey generaliza la idea de partición de los conjuntos. Sea $P_r(S) = A_1 \cup A_2 \cup \dots \cup A_t$ una partición del conjunto de todos los r -subconjuntos del conjunto S . Supongamos, luego, que vienen dados números enteros q_i tales que $1 \leq r \leq q_i$; $i = 1, 2, \dots, t$. Si existe un q_i -subconjunto de S , cuyos r -subconjuntos se contienen todos en A_i , entonces se le llama (q_i, A_i) -subconjunto del conjunto S . El teorema de Ramsey afirma que existe un número mínimo natural $N(q_1, q_2, \dots, q_t, r)$, a partir del cual (es decir, para $n \geq N$) el conjunto S contiene un (q_i, A_i) -subconjunto para cierto $i = \{1, 2, \dots, t\}$.

Veamos un caso particular del teorema de Ramsey en que $r = 2$. La interpretación geométrica en este caso se expresa así: n elementos corresponden a n puntos en un plano; a un par de puntos corresponde un segmento que une dichos puntos; la inclusión del segmento en el subconjunto A_i se marca pintándolo de i -ésimo color ($i = 1, 2, \dots, t$). El teorema de Ramsey garantiza la existencia, para $n \geq R(p, q; 2)$, de p puntos unidos entre sí sólo con segmentos del primer color, o bien de q puntos unidos sólo con segmentos del segundo color.

La misma idea sirve de base para la aplicación del teorema de Ramsey a los polígonos convexos. Se analizan los n -conjuntos de puntos en un plano tales que ninguna combinación de tres puntos se dispone en una

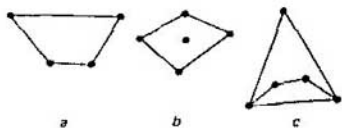


Fig. 5.1.

misma recta. Resulta que empleando los puntos de dicho conjunto como vértices, podemos construir cualquier polígono convexo de m lados, siempre que n sea suficientemente grande.

Teorema. Para un número entero dado m existe un número entero mínimo N_m tal que cualesquiera N_m puntos en el plano, de los cuales ninguna combinación de tres se dispone en una misma recta, contienen m puntos que forman un polígono convexo de m lados.

Demostración. Observemos, para iniciar, que entre cualesquiera 5 puntos se pueden elegir 4 que representen los vértices de un cuadrilátero convexo. La demostración se observa fácilmente en la fig. 5.1. Según se ve, incluso el caso "degenerado" (fig. 5.1, c) da un cuadrilátero convexo (en el interior de un triángulo).

Ahora, es lícita la afirmación de que si entre m puntos en consideración cualesquiera 4 son vértices de un cuadrilátero convexo, entonces dichos m puntos forman un polígono convexo de q lados. En efecto, supongamos que existe, a lo sumo, un polígono convexo de q lados ($q < m$). Sirviéndonos de las diagonales trazadas a partir de un mismo vértice, dividamos este polígono en triángulos y en este caso $(m - q)$ puntos caerán en el interior de los triángulos, lo que contradice la condición de convexidad de todos los cuadriláteros. Esto quiere decir que $q = m$, el polígono de m lados será convexo.

Analicemos ahora la afirmación general. Sea $m \geq 4$; $n \geq N(m, 5; 4)$. Partiremos los 4-conjuntos (hay en total $\binom{n}{4}$ de tales conjuntos) en convexos y no convexos. Entonces, según el teorema de Ramsey, o bien: a) existe un pentágono con todos los cuadriláteros no convexos, lo que contradice lo demostrado anteriormente; o bien b) existe un polígono de m lados cuyos cuadriláteros son todos convexos, y, por consiguiente, de conformidad con lo demostrado anteriormente, este polígono de m lados es convexo.

Las interpretaciones geométricas existen para casi todos los objetos combinatorios. Por ejemplo, para los cuadrados latinos la interpretación geométrica se construye del modo siguiente. Cada uno de los n^2 elementos de un cuadrado latino se considera como un punto sobre la diversidad de 2-coordenada. Por el conjunto de puntos que forman el cuadrado latino tracemos tres n -familias de líneas que pasan: 1) por las filas, 2) por las columnas, 3) por los elementos iguales. Estas familias forman una 3-red

con n^2 nodos. Inversamente, cada 3-red de esta índole puede interpretarse como cuadrado latino.

La condición de equivalencia de los cuadrados latinos en la interpretación que se considera aparece como conservación de las 3-redes al reenumerar las líneas de cada familia. A veces se admite la sustitución recíproca de las familias, con lo cual se amplía el concepto de equivalencia. Diremos, por fin, que la interpretación geométrica fue, evidentemente, la causa por la cual surgió el concepto de ortogonalidad en aplicación a los cuadrados latinos.

Una aplicación sucesiva y reiterada de las interpretaciones ilustrativas condujo a que se hayan separado clases de modelos matemáticos que poseen sus propios medios de investigación claramente distinguibles a saber, la axiomática, la terminología, los símbolos y los medios lógicos de demostración. En nuestro caso dichos medios se representan, con preferencia, por las geometrías y los grafos finitos. La tarea del capítulo presente consiste en introducir al lector en estas ramas extensas de las matemáticas, en interés del objetivo fundamental del libro.

5.2. IDEAS GEOMÉTRICAS FINITAS

Para los sistemas que se estudian en el análisis combinatorio se emplean geometrías finitas. Este término se usa para designar sistemas compuestos por un número finito de elementos, entre los cuales están establecidas las relaciones de incidencia. A los elementos, no definibles en general, se les asignan nombres geométricos: "puntos" P y "líneas" L . La relación de incidencia I se lee así: "un punto P se dispone en la línea L ", o bien "la línea L contiene el punto P ". Para dichas relaciones se introducen, además, axiomas del tipo que tienen los axiomas geométricos.

Las precisiones y explicaciones adicionales o bien las modificaciones de las enunciaciones iniciales conducen a diferentes tipos de geometrías finitas y a las partes de la geometría contiguas y próximas a las citadas geometrías. Así, por ejemplo, surgen las relaciones con los problemas de la topología combinatoria, la geometría discreta, la geometría proyectiva, la teoría geométrica de los números, la teoría de los grafos, la geometría combinatoria, etc.

Sobre los espacios proyectivos. Se conoce, por ejemplo, que la totalidad de todos los espacios de una estructura incidente $\{P, L, J\}$, que acabamos de mencionar, donde los elementos P se llaman puntos, los elementos L , líneas, y J , relaciones de incidencia, es un espacio proyectivo.

Existen diferentes métodos de construcción axiomática de un espacio proyectivo. El más empleado es la modificación del sistema de axiomas propuesta en 1899 por D. Hilbert con el fin de fundamentar la geometría elemental. Todo espacio proyectivo se considera como una colección de elementos de tres géneros: puntos, rectas y planos, entre los cuales se

establece la relación de incidencia, fundamental para la geometría proyectiva, que se caracteriza por axiomas adecuados. Estos axiomas se diferencian del grupo correspondiente de axiomas de la geometría elemental en que en los primeros se requiere que cada dos rectas dispuestas en un mismo plano tengan un punto común, y en cada recta se tengan al menos tres puntos distintos. Para obtener una geometría más sustancial, la totalidad citada de axiomas se completa con axiomas de orden y continuidad (para un espacio proyectivo real), con el axioma de Pappus (para la geometría proyectiva sobre un cuerpo conmutativo), con el postulado de Fano, etc.

Examinemos a continuación más detalladamente dos métodos de construir espacios proyectivos. En primer lugar puede tratarse de la realización de un espacio proyectivo como un conjunto de rectas en un espacio lineal. Sea L un espacio lineal sobre un campo K , y denotemos con $\dim L$ la dimensión de L . El conjunto $P(L)$ de rectas (es decir, de subespacios lineales unidimensionales) en L lleva el nombre de *espacio proyectivo* asociado con L , y las propias rectas en L se llaman puntos de $P(L)$. El número $\dim L - 1$ es la dimensión de $P(L)$ y se denota $\dim P(L)$. Los espacios proyectivos y bidimensionales se denominan *recta proyectiva* y *plano proyectivo*, respectivamente.

Un conjunto de la forma $P(A)$, donde $A \subseteq L$ se llama *subespacio proyectivo* en $P(L)$. Es evidente que $P(A_1 \cap A_2) = P(A_1) \cap P(A_2)$, es decir, una familia de espacios proyectivos está cerrada respecto de las intersecciones. Por eso, en el conjunto de subespacios proyectivos $P(A)$ que contienen el conjunto dado $S \subseteq P(L)$ hay un subespacio menor, que se representa por la intersección de todos los subespacios de este tipo. Este último lleva el nombre de *cápsula proyectiva* de S , se denota con \overline{S} , y coincide con $P(A)$, donde A es la cápsula lineal de todas las rectas correspondientes a los puntos $s \in S$ en L . Se puede mostrar, que si P_1 y P_2 son dos subespacios proyectivos finitos en el espacio proyectivo P , entonces

$$\dim(P_1 \cap P_2) + \dim(\overline{P_1 \cap P_2}) = \dim P_1 + \dim P_2.$$

En adelante nos harán falta dos espacios proyectivos concretos que se denominan configuraciones de Desargues y de Pappus.

Configuración de Desargues. Sea A una familia de puntos en un espacio proyectivo. El símbolo \overline{A} designará su cápsula proyectiva. En un espacio proyectivo tridimensional examinemos el seis ordenado de puntos $(a_1, a_2, a_3, b_1, b_2, b_3)$. Se supone que los puntos son distintos de dos en dos y que $a_1a_2a_3$ y $b_1b_2b_3$ son, en esencia, unos planos. Luego, supongamos que las rectas a_1b_1, a_2b_2 y a_3b_3 se intersecan en un solo punto p , distinto de $a_1, a_2, a_3, b_1, b_2, b_3$ (fig. 5.2). En este caso diremos que los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$ son perspectivas respecto del punto p y cada uno de ellos es una proyección del otro realizada desde el centro p , si ellos se disponen en diferentes planos. Entonces, para todo par de índices $\{i, j\} \subseteq \{1, 2, 3\}$ las rectas a_ia_j y b_ib_j no coinciden, pues en el caso contrario

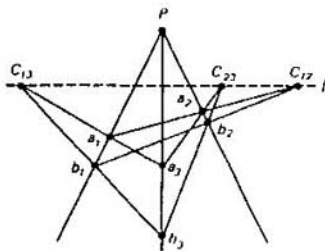


Fig. 5.2.

tendríamos $a_i = b_i$, puesto que a_i y b_i son puntos de intersección de estas rectas con la recta pa_ib_i . Además, las rectas a_1a_2 y b_1b_2 se disponen en el plano común pa_1a_2 . Por eso, ellas se intersecan en un punto que se designará con c_{ij} . Los puntos c_{12} , c_{13} , c_{23} son puntos de intersección de las prolongaciones de los pares de los lados correspondientes de los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$.

Diremos que los triángulos $(a_1a_2a_3)$ y $(b_1b_2b_3)$ son perspectivas respecto de la recta l , siempre que los puntos c_{12} , c_{13} , c_{23} se disponen en la recta l .

Teorema de Desargues. Si dos triángulos son perspectivas respecto de cierto punto, son perspectivas respecto de una recta.

Configuración de Pappus. Examinemos en un plano proyectivo dos rectas diferentes l_1 , l_2 y dos ternas de puntos distintos a_1, a_2, a_3 y b_1, b_2, b_3 (fig. 5.3) dispuestos en las rectas de dos en dos. Para cualquier par de índices $\{i, j\} \in \{1, 2, 3\}$ tales que $i < j$ designemos con c_{ij} el punto de intersección de las rectas a_ib_j y a_jb_i .

Teorema de Pappus. Los puntos c_{12} , c_{13} , c_{23} se disponen en una misma recta.

El lector puede encontrar la demostración de los teoremas de Desargues y de Pappus, como también otros pormenores, en el libro [69].

En segundo lugar, los espacios proyectivos pueden construirse por adición a un espacio afín de los "puntos infinitos alejados". Merced a este procedimiento se logra la uniformidad en las enunciaciones y demostraciones de toda una serie de teoremas de la geometría analítica y de algunos

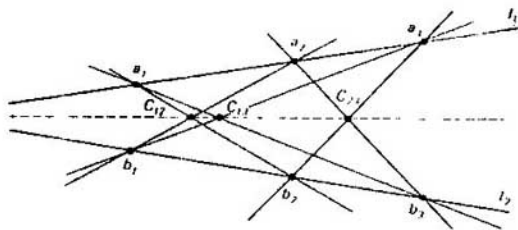


Fig. 5.3.

otros resultados geométricos. Este método condujo a la creación de la geometría proyectiva clásica. A medida que se desarrollaba el método axiomático fueron creadas diferentes descripciones axiomáticas de los espacios proyectivos. Las investigaciones ulteriores han mostrado que los sistemas geométricos que obedecen a dichos axiomas tienen su propio significado y la teoría de ellos resulta ser muy profunda (véanse [4, 21, 76 y 77]). Desde entonces por *espacio proyectivo clásico* se entiende, por lo común precisamente tal sistema geométrico. Por ejemplo, un espacio tridimensional se define como un conjunto, cuyos elementos se llaman *puntos*, provisto de dos sistemas de subconjuntos cuyos elementos se llaman *rectas* y *planos*, respectivamente. En este caso han de cumplirse los siguientes axiomas de incidencia:

1. Dos puntos distintos pertenecen a una recta única.
2. Tres puntos distintos, no dispuestos en una recta, pertenecen a un plano único.
3. Una recta y un plano tienen un punto común.
4. La intersección de dos planos contiene una recta.
5. Existen cuatro puntos que no se disponen en un plano y que son tales que cualesquiera tres de ellos no se disponen en una recta.
6. Cada recta consta no menos que de tres puntos.

En lo que sigue examinaremos sólo las geometrías proyectivas clásicas, denominándolas, para simplificar, *geometrías proyectivas*.

Un *plano proyectivo clásico* se define como un conjunto, cuyos elementos se llaman *puntos*, provisto de un sistema de subconjuntos cuyos elementos se llaman *rectas*. En este caso han de cumplirse los siguientes axiomas:

1. Dos puntos diferentes pertenecen a una recta única.
2. La intersección de dos rectas es no vacía.
3. Existen tres puntos que no se disponen en una recta.
4. Cada recta consta por lo menos de tres puntos.

El conjunto $P(L)$, donde L es un espacio lineal sobre el campo K de dimensión 4 ó 3, junto con los sistemas de planos proyectivos y rectas en ellos, tal como fueron definidos más arriba, satisfacen los axiomas del espacio proyectivo clásico tridimensional y del plano proyectivo clásico, respectivamente. No obstante, no todo espacio proyectivo clásico o plano es isomorfo a uno de nuestros espacios $P(L)$. Más aún, resulta que existen planos proyectivos clásicos no isomorfos incluso a ningún plano del tipo $P(L)$, donde L es un espacio proyectivo tridimensional sobre algún cuerpo. La razón de ello radica en que en los planos proyectivos del tipo $P(L)$ el teorema de Desargues sigue siendo válido, mientras que existen planos que no son de Desargues, donde el teorema no se cumple. Entencemos sin demostración el resultado siguiente:

Teorema 1. Tres propiedades del plano proyectivo clásico son equivalentes:

- a) en dicho plano se cumple el teorema plano de Desargues;

b) dicho plano puede encajarse en un espacio proyectivo clásico;

c) existe un espacio lineal tridimensional L sobre cierto campo K , definido de un modo unívoco, exactamente hasta un isomorfismo tal, que nuestro plano es isomorfo a $P(L)$.

El teorema de Pappus puede fallar incluso en los planos de Desargues. Al denominar una afirmación correspondiente axioma de Pappus, podemos formular el siguiente teorema que también se da sin demostración.

Teorema 2. Si en un plano proyectivo clásico se cumple el axioma de Pappus, el plano es de Desargues. El plano clásico de Desargues satisface el axioma de Pappus, cuando y sólo cuando el cuerpo asociado con el plano es conmutativo, es decir, el plano citado es isomorfo a $P(L)$, donde L es un espacio lineal tridimensional sobre el campo.

En todo espacio proyectivo clásico de dimensión superior a dos queda válido el teorema de Desargues. El teorema de Desargues permite la coordinatización de un espacio proyectivo clásico con ayuda de cierto cuerpo asociativo K . De este modo, el estudio de los espacios proyectivos clásicos, cuya dimensión es superior a dos, se reduce a un problema algebraico, a saber, a la teoría de los cuerpos asociativos. En los planos proyectivos el teorema de Desargues puede no tener lugar y la reducción al álgebra mencionada no se consigue en el caso general. Con este motivo el estudio de los planos proyectivos clásicos es de interés especial y la teoría de ellos se desarrolla exitosamente desde el principio de nuestro siglo.

Planos proyectivos finitos. Es fácil comprobar que la definición de plano proyectivo clásico propuesta más arriba es equivalente a la siguiente.

Se denomina plano proyectivo a un conjunto de puntos y rectas, entre los cuales queda establecida una relación de incidencia que obedece a los siguientes requisitos.

1. Dos puntos distintos son incidentes, con una y sólo una recta.
2. Dos rectas distintas son incidentes al menos con un punto.
3. Existen cuatro puntos en una posición general, es decir, tales que cualesquiera tres de ellos no son incidentes con una recta.

Efectivamente, para convencerse de esta afirmación, basta demostrar el siguiente teorema.

Teorema 3. Cada recta contiene no menos de tres puntos.

Demostración. Del axioma 3 se deduce que existen cuatro puntos a_1, a_2, a_3, a_4 , de los cuales ninguna combinación de tres se dispone en una misma recta. A estos puntos les corresponden seis rectas distintas que los unen de dos en dos a saber, $l_1: a_1a_2b_1$; $l_2: a_1a_3b_2$; $l_3: a_1a_4b_3$; $l_4: a_2a_3b_3$; $l_5: a_2a_4b_2$; $l_6: a_3a_4b_1$. Aquí, b_1, b_2 y b_3 son puntos de intersección de estas rectas. Partiendo de que todas las seis rectas son distintas se obtiene con facilidad que todos los puntos b_1, b_2, b_3 son distintos y todos ellos se diferencian de los puntos a_i ; $i = 1, 2, 3, 4$. Cada una de las rectas l_i ($i = 1, 2, \dots, 6$) contiene por lo menos tres puntos. Si una recta arbitraria l no

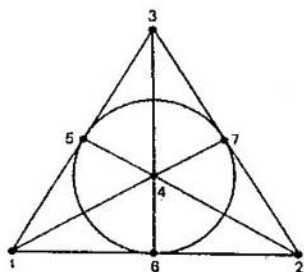


Fig. 5.4.

contiene a_1 , se interseca con las rectas l_1, l_2, l_3 en tres puntos diferentes. Si l no contiene a_2 , corta las rectas l_1, l_4, l_5 en tres puntos distintos. En cambio, si la recta l contiene tanto a_1 , como a_2 , entonces $l = l_1$, mas la recta l_1 contiene en todo caso tres puntos diferentes a_1, a_2 y b_1 . El teorema queda demostrado.

Observemos que del axioma 1 se desprende que dos rectas diferentes son incidentes no más que con un punto. En la fig. 5.4 se expone un plano proyectivo finito que contiene siete puntos y el mismo número de rectas, con la particularidad de que además de los lados y las medianas del triángulo equilátero tenemos que considerar como "recta" la circunferencia inscrita. Este plano proyectivo concreto se denomina, a menudo, configuración de Fano. Otro ejemplo de plano proyectivo ya ha sido examinado, precisamente el $P(L)$, cuando L es un espacio lineal sobre el campo de dimensión 3.

El axioma 3 en la definición de plano proyectivo sirve, principalmente, para desechar los planos "degenerados" tales, por ejemplo, como un plano vacío, una serie rectilínea de puntos (en particular, una recta), un haz de rectas (en particular, un punto), una serie rectangular de puntos por uno de los cuales están trazadas varias rectas y un haz de rectas, cortado por una recta.

Introduzcamos unas definiciones más. Un subconjunto π' de puntos y de rectas del plano proyectivo π se llama *subplano*, si, junto con cada dos puntos (rectas), π' contiene una recta (un punto) incidente con ellos. Dos planos se denominan *isomorfos*, si entre sus puntos y sus rectas puede establecerse una correspondencia biunívoca que conserva la incidencia. Específico para los planos proyectivos resulta ser el concepto de dualidad. Dos planos se llaman *duales*, si entre los puntos (rectas) de un plano y las rectas (puntos) del otro puede establecerse una correspondencia biunívoca que conserva la incidencia. Un plano dual al plano dado puede obtenerse al declarar "puntos" las rectas del plano dado, y "rectas", los puntos de él, conservando las incidencias existentes. No es difícil convencerse de que un plano isomorfo a un plano que es dual del plano π ,

será dual con relación a π . Un plano, dual respecto del dual, es isomorfo al plano de partida. La aplicación isomorfa de un plano sobre sí mismo lleva el nombre de *colineación*. La aplicación unívoca (no forzosamente biunívoca) de los puntos y de las rectas del plano π sobre los puntos y las rectas del plano π' con la conservación de la incidencia recibe el nombre de homomorfismo.

Teorema 4. Sea $n \geq 2$ un número entero arbitrario. Entonces son equivalentes las siguientes propiedades del plano proyectivo:

- a) cierta recta contiene exactamente $n + 1$ puntos;
- b) cierto punto pertenece exactamente a $n + 1$ rectas;
- c) cada recta contiene exactamente $n + 1$ puntos;
- d) cada punto se dispone exactamente en $n + 1$ rectas;
- e) en el plano π hay exactamente $n^2 + n + 1$ puntos;
- f) en el plano π hay exactamente $n^2 + n + 1$ rectas.

Demostración. Sean a_1, a_2, a_3, a_4 cuatro puntos, de los cuales ninguna combinación de tres se dispone en una misma recta. A estos puntos les corresponden seis diferentes rectas que los unen dos a dos (véase la demostración del teorema 3). Supongamos ahora que se cumple la propiedad a), es decir, en π existe una recta l que contiene exactamente $n + 1$ puntos, digamos, para concretar, c_1, c_2, \dots, c_{n+1} . Si b es un punto que no está dispuesto en l , entonces las rectas $bc_i, i = 1, 2, \dots, n + 1$, son distintas, pues, si $\overline{bc_i} = \overline{bc_j}$ para ciertos $i \neq j$, entonces b está situado en l , lo que contradice la hipótesis. Luego, cada recta, que pasa por b , corta l , y, por consiguiente, ha de ser una de las $n + 1$ rectas $bc_i, i = 1, 2, \dots, n + 1$. Por lo menos dos puntos de los cuatro a_1, a_2, a_3, a_4 no se disponen en la recta l , y, por eso, tal punto b existe. Sea, ahora, b un punto que yace exactamente en $n + 1$ rectas: m_1, m_2, \dots, m_{n+1} . Si l^* es una recta que no pasa por b , entonces l^* corta m_1, \dots, m_{n+1} en los puntos d_1, \dots, d_{n+1} , que son todos diferentes, por cuanto b es el único punto que yace más que en una de las rectas m_1, \dots, m_{n+1} . Si en l^* existiera un punto más d_{n+2} , entonces existiría una recta bd_{n+2} que no coincide con ninguna recta $m_j, j = 1, 2, \dots, n + 1$, puesto que de lo contrario la recta bd_{n+2} contendría cierto punto $d_j, j \neq n + 2$, y, entonces, $\overline{bd_{n+2}} = \overline{bd_{n+2}d_j} = \overline{d_{n+2}d_j} = l^*$, lo que contradice la suposición de que b no está situado en la recta l^* .

Nuestra recta inicial l contenía exactamente $n + 1$ puntos; por consiguiente, cada punto que no yace en l , está situado exactamente en $n + 1$ rectas; a los puntos de esta índole se relacionan por lo menos dos puntos de cuatro: a_1, a_2, a_3, a_4 , por ejemplo, a_1 y a_2 . Por eso, cada recta que no pasa por a_1 ó a_2 contiene exactamente $n + 1$ puntos, es decir, cada recta, salvo, quizás, la recta $l_1: a_1a_2b_1$ contiene exactamente $n + 1$ puntos. Entonces, $l_2: a_1a_3b_2$ contiene exactamente $n + 1$ puntos, y el punto b_3 que no yace en l_2 , está situado exactamente en $n + 1$ rectas; por consiguiente,

l_1 , que no contiene b_3 , también debe contener $n + 1$ puntos. Así pues, la propiedad a) lleva consigo la propiedad c).

Pero, para cualquier punto b puede encontrarse una recta que no pasa por este punto, y por esta razón existen (igual que anteriormente) exactamente $n + 1$ rectas que pasan por b , con lo que quedan demostradas las propiedades b) y d).

Demostremos ahora la propiedad e), partiendo de a). Sea b_0 un cierto punto del plano π , y supongamos que l_1, l_2, \dots, l_{n+1} representan $n + 1$ rectas que pasan por b_0 . Estas rectas contienen todos los puntos del plano π , con la particularidad de que cada una de ellas contiene otros b_0 y n puntos. El punto b_0 es el único, perteneciente a cualesquiera dos de las rectas l_1, l_2, \dots, l_{n+1} . Por consiguiente, el plano π contiene $1 + (n + 1)n = n^2 + n + 1$ puntos. Confirmemos ahora la propiedad f). Sea l_0 una recta del plano π , y supongamos que b_1, \dots, b_{n+1} son $n + 1$ puntos de la recta l_0 . Cada uno de estos puntos yace en l_0 y en otras n rectas. De este modo, obtenemos todas las rectas del plano π ; en total tenemos, por consiguiente, $1 + (n + 1)n = n^2 + n + 1$ puntos. Así pues, de la propiedad a) se deducen todas las demás propiedades mencionadas en el teorema.

En virtud de la dualidad, de b) también se deducen todas las propiedades restantes. Es evidente que de c) se deduce a), y de d) se deduce b). Si es cierta la propiedad e) y cierta recta consta de $m + 1$ puntos, donde m es un número natural, entonces el plano π se compone de $m^2 + m + 1 = n^2 + n + 1$ puntos, de donde $m = n$, es decir, de e) se desprende a). Análogamente, de f) se deduce la propiedad b). El teorema está demostrado.

Convengamos en llamar *orden* de un plano proyectivo finito al número n , si cada recta de este plano contiene $n + 1$ puntos. Observemos que el plano es de orden n , si posee cualquiera de las seis propiedades citadas en el teorema 4. El orden mínimo de los planos proyectivos finitos es igual a 2. El plano finito "más pequeño" es la configuración de Fano, cada recta de la cual contiene exactamente 3 puntos. La configuración se expone en la fig. 5.4. En total dicho plano contiene 7 ($= 2^2 + 2 + 1$) puntos y el mismo número de rectas. Las rectas en este plano son: $l_1: 1, 2, 6$; $l_2: 1, 3, 5$; $l_3: 2, 3, 7$; $l_4: 1, 4, 7$; $l_5: 3, 4, 6$; $l_6: 2, 4, 5$; $l_7: 5, 6, 7$.

Ya en este ejemplo más sencillo se revelan las posibilidades de interpretar los planos proyectivos finitos. A saber, un plano proyectivo finito de orden 2 resulta ser un sistema de ternas de Steiner de orden 7. Hemos de notar que este sistema es un bloque-esquema con los parámetros $v = b = 7$, $k = r = 3$, $\lambda = 1$. En el caso general los planos de orden n representan bloque-esquemas, cuyos parámetros tienen la forma: $v = b = n^2 + n + 1$, $k = r = n + 1$, $\lambda = 1$, es decir, bloque-esquemas incompletos equilibrados simétricos. Viceversa, los bloque-esquemas con tales parámetros son planos

proyectivos simétricos finitos, puesto que, evidentemente, se cumplen todos los axiomas. Establezcamos una conexión útil más entre los planos proyectivos finitos y las familias completas de cuadrados latinos ortogonales.

Teorema 5. Sea $n \geq 3$ un número entero arbitrario. Un plano proyectivo de orden n existe, cuando y sólo cuando puede construirse una familia completa de $n - 1$ cuadrados latinos ortogonales de orden n .

Demostración. Sea dado un plano proyectivo finito π de orden n . Veamos en él una recta l . En esta recta se tienen, como sabemos, $n + 1$ puntos: a_1, a_2, \dots, a_{n+1} . Supongamos que b_1, b_2, \dots, b_n son los demás n^2 puntos del plano π que no están situados en l . Cada punto a_j ($j = 1, 2, \dots, n + 1$) yace en n rectas, sin tener en cuenta la recta l . Numeremos arbitrariamente con números naturales de 1 hasta n todas las n rectas que pasan por el punto a_j . Procedamos de este modo para todo $j = 1, 2, \dots, n + 1$. Supongamos que la recta $b_i a_j$ está numerada con un número natural a_{ij} . Entonces, $A = [a_{ij}]$, donde $i = 1, 2, \dots, n^2$; $j = 1, 2, \dots, n + 1$, es una $n^2 \times (n + 1)$ -tabla de elementos $1, 2, \dots, n$.

Las filas de cualquier $(n^2 \times 2)$ -subtabla elegida de A representan n^2 pares de los elementos $1, 2, \dots, n$. Si suponemos que $a_{ij} = a_{i'j}$ y $a_{ik} = a_{i'k}$, donde $i \neq i'$; $j \neq k$, entonces $b_i a_j = b_{i'} a_j$, y $b_i a_k = b_{i'} a_k$. Mas, en este caso la recta $b_i b_{i'}$ contiene tanto el punto a_j , como el a_k , es decir, la recta $b_i b_{i'}$ coincide con la recta l , lo que contradice nuestra suposición. Esto quiere decir que la matriz $A = [a_{ij}]$ es una $n^2 \times (n + 1)$ -tabla ortogonal de fuerza 2, la cual, según lo demostrado en el capítulo 4, es equivalente al conjunto completo de $n - 1$ cuadrados latinos ortogonales de orden n .

Demostremos la afirmación inversa. Sea dado un conjunto completo de $n - 1$ cuadrados latinos ortogonales de orden n . Pasemos de éste, como lo hicimos en el cap. 4, a una tabla

$$A = [a_{ij}]; \quad i = 1, 2, \dots, n^2; \quad j = 1, 2, \dots, n + 1.$$

A cada una de n^2 filas de la tabla A pongamos en correspondencia los puntos b_1, b_2, \dots, b_n , y a las columnas, los puntos a_1, a_2, \dots, a_{n+1} . Definamos la recta l_i como compuesta por los puntos a_1, a_2, \dots, a_{n+1} . Por todo punto a_j tracemos n rectas l_{ij} que contienen aquellos puntos b_i , donde en la j -ésima columna figura el número i . Se obtiene un plano proyectivo finito de orden n , en el cual se tienen $n^2 + n + 1$ rectas y el mismo número de puntos, con la particularidad de que cada recta contiene exactamente $n + 1$ puntos y cada punto yace en $n + 1$ rectas. Efectivamente, sean l_{ij} y $l_{i'j}$ dos rectas, y, además $i \neq i'$. Estas rectas son incidentes con uno y sólo con un punto a_j . Las rectas l_{ij} y l también son incidentes con uno y sólo con un punto. Esto demuestra el requisito 3 en la definición de plano proyectivo. Los cuatro puntos: $(1, 1), (1, 2), (2, 1), (2, 2)$ satisfacen el requisito 3 de la definición de plano proyectivo, lo que quiere decir que

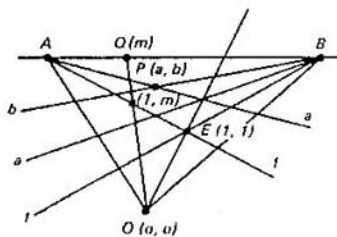


Fig. 5.5.

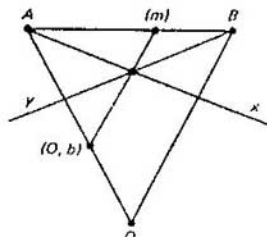


Fig. 5.6.

el plano construido π es realmente un plano proyectivo de orden n . El teorema queda completamente demostrado.

Introducción de las coordenadas. Elijamos en un plano proyectivo π cuatro puntos A, B, O, E en la posición general (véase fig. 5.5). Llamemos haz A a un conjunto de rectas que pasan por el punto a . Pongamos a cada recta del haz A en correspondencia aquella recta del haz B con la que ella se encuentra en la recta OE . Atribuyamos a cada una de las rectas del haz A (a excepción de AB) un símbolo especial a y asignemos el mismo símbolo a la recta correspondiente del haz B . Agreguemos a la recta AO símbolo 0 y a la recta AE , el símbolo 1 . Las rectas BO y BE recibirán los mismos símbolos. De este modo se define cierto conjunto de símbolos $M = \{0, 1, \{a\}\}$.

Supongamos que el punto P no pertenece a la recta AB (fig. 5.5). Por dicho punto pasa una recta de cada haz A y B , a saber, las rectas AP y BP . Supongamos que a la recta AP se le asigna el símbolo a , y a BP , el símbolo b . Entonces, al punto P le ponemos en correspondencia un par ordenado de símbolos (a, b) de M , que se denominarán coordenadas de este punto. Es fácil ver que la correspondencia entre los puntos que no están situados en AB y los pares (a, b) será biunívoca. Además, los puntos dispuestos en OE tendrán iguales ambas coordenadas. Asignemos ahora las coordenadas a los puntos de la recta AB distintos de A . Con este fin unamos el punto Q (véase fig. 5.5) con O y veamos donde la recta QO se cortará con AE . La segunda coordenada del punto de intersección se tomará por la coordenada del punto Q .

Observemos que mediante la misma construcción pueden obtenerse coordenadas cartesianas en un plano real habitual. El punto A debe considerarse como punto impropio del eje y , el punto B , como punto impropio del eje x , O se toma por origen de coordenadas y E , como el punto $(1, 1)$.

Definamos sobre el conjunto M , obtenido como resultado de la coordenatización, una operación ternaria, es decir, a cada terna ordenada de elementos x, m, b de M le ponemos en correspondencia un cierto elemento

$$y = x \cdot m \circ b$$

dispuesto en M . Este elemento y se define a base de los elementos x, m, b , como segunda coordenada del punto de intersección de la recta x del haz A con la recta $(m), (0, b)$ (fig. 5.6). Es evidente que por medio de este procedimiento y se define de un modo unívoco.

Teorema 6. Cualquier forma de prefijar cuatro puntos A, B, O, E en una posición general determina una operación ternaria con las siguientes propiedades:

1. $0 \cdot m \circ c = a \cdot 0 \circ c = 0$.
2. $1 \cdot m \circ 0 = m \cdot 1 \circ 0 = m$.
3. $a \cdot m \circ z = C$ resoluble unívocamente respecto de z .
4. $x \cdot m_1 \circ b_1 = x \cdot m_2 \circ b_2$ resoluble unívocamente respecto de x .
5. El sistema

$$a_1 = m \circ b - c_1; a_2 \cdot m \circ b = c_2; a_1 = a_2$$

define unívocamente un par (m, b) .

Demostración. Al elegir cuatro puntos A, B, O, E , de los cuales ninguna combinación de tres yace en una recta del plano π , formamos un anillo ternario con la operación $x \cdot m \circ b$, al igual que lo hicimos más arriba. Las propiedades 1 y 2 se deducen inmediatamente de la definición. La propiedad 3 significa que una recta que liga los puntos (m) y (a, c) corta OB en un punto bien determinado $(0, z)$. El significado de la exigencia 4 consiste en que dos rectas: $y = x \cdot m_1 \circ b_1$ e $y = x \cdot m_2 \circ b_2$ con diferentes direcciones m_1 y m_2 se intersecan en el único punto finito. La exigencia 5 dice que si (a_1, c_1) y (a_2, c_2) son dos puntos finitos para $a_1 \neq a_2$, entonces existe una recta única $y = x \cdot m \circ b$, que pasa por estos dos puntos.

El conjunto M con la operación ternaria que posee las propiedades 1... 5 del teorema 6 lleva el nombre de *terna*. Si una terna dada está construida sobre el plano π según la regla descrita anteriormente, se llamará terna del plano π .

Aduzcamos el resultado siguiente sin demostrarlo.

Teorema 7. Dada una terna M , puede determinarse con su ayuda un plano proyectivo con los puntos $(a, c), (m), A$, y con las rectas $y = x \cdot m \circ b, x = a, l_\infty$, donde a, b, c, m representan toda clase de elementos de M , mientras que A y l_∞ son ciertos símbolos auxiliares; la incidencia se define del modo siguiente: (a, c) es incidente con $y = x \cdot m \circ b$, cuando y sólo cuando $c = a \cdot m \circ b$; (a, c) es incidente con $x = a$ para todo c ; (m) es incidente con $y = x \cdot m \circ b$ para todo b ; (m) es incidente con l_∞ , cualquiera que sea m ; A es incidente con $x = a$ para cualesquiera a , y, además, con l_∞ .

Introduzcamos, para el plano dado π , las operaciones naturales: adición y multiplicación, que se definen partiendo de una de las ternas de este plano, según las reglas siguientes:

$$\begin{aligned} a + b &= a \cdot 1 \circ b; \\ ab &= a \cdot b \circ 0. \end{aligned}$$

El conjunto M con la adición y multiplicación definidas del modo citado recibe el nombre de *cuerpo natural* del plano π .

Ha de notarse que un cuerpo natural no define, hablando en general, la terna que lo engendra, y el teorema, análogo al teorema 7, no tiene lugar en este caso.

Así pues, los planos proyectivos finitos admiten diferentes interpretaciones: mediante ternas, bloque-esquemas y familias completas de cuadrados latinos ortogonales. Sin embargo, la teoría de los planos finitos todavía está lejos de ser elaborada. Ni siquiera existe la respuesta completa a la pregunta sobre el volumen de los objetos combinatorios de esta clase, como también sobre las condiciones, bajo las cuales los planos finitos pueden existir y no existir.

Se conoce que un plano proyectivo finito existe, si su orden n tiene la forma p^α , donde p es un número primo, y α , un número natural. En efecto, se ha demostrado (véase cap. 4) que si $n = p^\alpha$, $n \geq 3$, entonces existe una familia completa de $n - 1$ cuadrados latinos ortogonales. Mientras tanto, la existencia de tal conjunto es equivalente a la existencia del plano proyectivo finito que se busca.

De los demás resultados señalaremos los siguientes: a) el plano finito no existe, si $n \equiv 1, 2 \pmod{4}$ y si la parte de n , libre del cuadrado, tiene por lo menos un factor $p \equiv 3 \pmod{4}$; b) la condición necesaria de existencia de un plano proyectivo finito cuyo orden tiene la forma $n \equiv 1, 2 \pmod{4}$ consiste en la existencia de los números enteros x y y tales, que $n = x^2 + y^2$ (teorema de Bruck—Ryser). No damos aquí las demostraciones por ser éstas demasiado complejas. De los resultados citados se desprende, en particular, que los planos para $n = 6, 12, 14$ están ausentes. La cuestión referente a $n = 10, 18$ y 20 , queda por ahora abierta. Para que el lector perciba lo difícil del problema, indiquemos, como ejemplo, que la construcción de un plano de orden 10 requiere la construcción de una familia de 9 cuadrados latinos ortogonales dos a dos del mismo orden.

Una dirección importante en el desarrollo de la teoría de los planos proyectivos finitos la representan el cálculo del número de planos de un orden dado y el estudio de sus tipos diferentes. Para el año 1956 quedó demostrado que cuando $n < 9$ ($n = 2, 3, 4, 5, 7, 8$) existe el único plano proyectivo de orden n , a saber, el de Desargues. Tres planos construidos por Veblen y Wedderburn en 1907, junto con el plano de Desargues construido un año antes por Veblen y Bassi son los únicos que hoy día se conocen para el orden 9. No se han obtenido planos nuevos, como tampoco se ha demostrado que ellos no existen en general. En esta dirección se emprenden varias investigaciones. Parece que estamos a punto de terminar la resolución de la cuestión sobre la existencia de los planos proyectivos de orden 10.

5.3. SOBRE LOS GRAFOS

En el lenguaje de la teoría de los grafos se describen con éxito muchos tipos de los problemas combinatorios. Las representaciones gráficas son en este caso no simplemente ilustraciones, sino que también permiten obtener nuevos resultados. En este párrafo daremos una información indispensable de la teoría de los grafos que se emplea en la obra dada.

Se denomina *grafo orientado* una terna $G = (X, A, \varphi)$ compuesta de un conjunto no vacío X , cuyos elementos llevan el nombre de *vértices* de un conjunto A de *arcos* (llamados también), a veces, flechas y de una función $\varphi: A \rightarrow X \times X$, la cual a todo arco $a \in A$ se le hace corresponder un par ordenado (p, q) de vértices denominados *finales* de dicho arco. Un arco, cuyos finales (p, p) se encuentran en un mismo vértice, lleva el nombre de *lazo*. Un grafo privado de tales lazos se llama *grafo sin lazo*.

En la fig. 5.7. están expuestos cuatro grafos orientados; cada uno con cuatro vértices. Todos estos grafos no tienen lazos. Además, son *simples*: lo último significa por definición que cualquier par de vértices p, q se une mediante un arco a lo sumo. En los grafos orientados *simples* el único arco (si existe) con las finales, p, q se designará mediante pq .

Describamos dos familias importantes de los grafos orientados simples, cada uno de los cuales depende del parámetro n . Un camino *simple* de longitud n consta de $n + 1$ vértices: x_1, \dots, x_{n+1} y n arcos que unen los vértices vecinos: $x_k x_{k+1}$. El *contorno simple* de longitud n consta de n diferentes vértices: x_1, \dots, x_n y n arcos del tipo $x_k x_{k+1}$ para $k < n$, y, además, $x_n x_1$ (en particular, un contorno simple de longitud 1 es un lazo con vértice x_1). Todos los caminos de longitud n son isomorfos, como lo son también todos los contornos simples de longitud n (véase la definición más abajo). En la fig. 5.7, *a* está expuesto un trayecto simple de longitud 3, y en la fig. 5.7, *b*, un ciclo simple de longitud 4.

Se denomina *isomorfismo* de los grafos orientados $G = (X, A, \varphi)$ y $G^* = (X^*, A^*, \varphi^*)$ a un par de biyecciones $\beta: X \rightarrow X^*$ y $j: A \rightarrow A^*$ tal que en G el arco a va del vértice p al vértice q , cuando y sólo cuando en G^* el arco $j(a)$ va del vértice $\beta(p)$ al vértice $\beta(q)$. En otras palabras, $\varphi(a) = (p, q)$ en G es equivalente a la condición $\varphi^*(j(a)) = (\beta(p), \beta(q))$ en G^* . Dos grafos orientados se llaman isomorfos, si entre ellos existe un isomorfismo. Es natural identificar los grafos orientados isomorfos.

Todo grafo orientado $G = (X, A, \varphi)$ define en el conjunto de sus vértices una relación binaria de *sucesiones* σ . Por definición, $p\sigma q$ significa que $\varphi(a) = (p, q)$ para cierto $a \in A$. Viceversa, toda relación binaria ρ en el conjunto X define un grafo orientado simple $G(\rho) = (X, A(\rho), \psi)$, en el que $A(\rho) = \{(p, q) | p \in X, q \in X, p\rho q\}$ y $\psi(pq) = (p, q)$. De aquí se deduce inmediatamente que la clase de conjuntos con una relación binaria (x, ρ) se encuentra en una correspondencia biyectiva natural con la clase de grafos orientados simples.

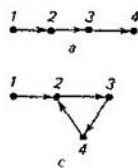


Fig. 5.7.

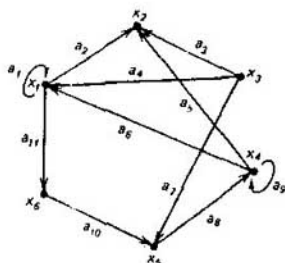
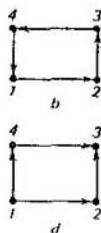


Fig. 5.8.

Se llama *grafo no orientado* (o, simplemente, *grafo*) la terna $G = (X, E, \Theta)$ compuesta de un conjunto no vacío X (cuyos elementos llevan el nombre de *vértices*), el conjunto E de aristas y la función Θ que a cada arista $a \in E$ le pone en correspondencia un par no ordenado de vértices $(p, q) = (q, p)$ que se denominan extremos de dicha arista. La arista (p, q) se llama *lazo*.

Los conceptos, introducidos para los grafos orientados, pueden ser extendidos a los grafos no orientados, si consideramos que una arista no orientada (p, q) corresponde a un par de arcos pq y qp .

Una *cadena simple* de longitud n consta de $n + 1$ diferentes vértices x_1, \dots, x_{n+1} y de n aristas que unen los vértices vecinos. Un *ciclo simple* de longitud n consta de n diferentes vértices: x_1, \dots, x_n y n aristas del tipo (x_k, x_{k+1}) para $k < n$, y, además, de una arista (x_n, x_1) .

El grafo orientado G puede también considerarse como un par $G = (X, \Gamma)$, formado por el conjunto X y la aplicación multiforme del conjunto X en sí mismo (es decir, por la aplicación de X en el conjunto $P(X)$ de todos los subconjuntos del conjunto X). Para el grafo expuesto en la fig. 5.8 tenemos: $\Gamma(x_3) = \{x_2, x_5\}$; $\Gamma(x_2) = \emptyset$. Con $\Gamma^{-1}(x_k)$ se designará el conjunto de vértices x_k del grafo, para los cuales en G existe un arco (x_k, x_i) . Sea $X_q = \{x_1, \dots, x_q\}$. Entonces, por $\Gamma(X_q)$ se entiende la unión $\Gamma(x_1) \cup \dots \cup \Gamma(x_q)$. La aplicación $\Gamma(\Gamma(x_i))$ se escribe como $\Gamma^2(x_i)$. Análogamente, una aplicación "triple" $\Gamma(\Gamma(\Gamma(x_i)))$ se escribe como $\Gamma^3(x_i)$, etc. Por ejemplo, para nuestro grafo orientado tenemos

$$\Gamma^2(x_1) = \Gamma(\Gamma(x_1)) = \Gamma(\{x_1, x_2, x_6\}) = \{x_1, x_2, x_5, x_6\};$$

$$\Gamma^3(x_1) = \Gamma(\Gamma(\Gamma(x_1))) = \Gamma(\{x_1, x_2, x_5, x_6\}) = \{x_1, x_2, x_4, x_5, x_6\}.$$

Un grafo (X', A') se llama *subgrafo* del grafo (X, A) , si X' y A' están contenidos en X y A , respectivamente. Se denomina *subgrafo de esqueleto* (*grafo parcial*) G_p del grafo $G = (X, A)$ al grafo (X, A_p) , para el cual $A_p \subseteq A$. De este modo, el subgrafo de esqueleto tiene también el mismo conjunto de vértices que el grafo G , pero el conjunto de arcos del

subgrafo G_p es un subconjunto del conjunto de arcos del grafo inicial. Sea un grafo $G = (X, \Gamma)$. Se llama *subgrafo generado* G_s al grafo (X_s, Γ_s) , para el cual $X_s \subseteq X$ y $\Gamma_s(x_i) = \Gamma(x_i) \cap X_s$ para cada vértice $x_i \in X_s$. Así pues, un subgrafo generado consta de un subconjunto de vértices X_s , de un conjunto de vértices del grafo inicial y de todos aquellos arcos del grafo G , cuyas vértices extremales y originales pertenecen al subconjunto X_s . En la fig. 5.9, *a, b* se exponen un subgrafo generado, un subgrafo de esqueleto y un subgrafo, respectivamente, con la particularidad de que este último no es ni generado ni subgrafo de esqueleto del grafo G , expresado en la fig. 5.8.

Sea $G = (X, A)$ un grafo orientado con un conjunto de vértices X y un conjunto de arcos A . Si queremos menospreciar el carácter orientado de los arcos del conjunto A , entonces un grafo no orientado, correspondiente a G , se denotará con $\bar{G} = (X, A)$ y se llamará *sosia (doble) no orientado* del grafo G . Un grafo no orientado G se denomina *conexo*, si cada par de sus vértices puede ser unido con una cadena. Un grafo finito, que no es conexo, puede partirse en un número finito de subgrafos finitos llamados *componentes*. Un grafo no orientado G se llama *completo*, si para cada par de sus vértices existe una arista que los une. Un grafo completo sobre n vértices suele designarse con K_n . Un grafo no orientado se denomina *bipartido*, si el conjunto de sus vértices puede ser partido en tales dos subconjuntos X_1 y X_2 que cada arista tenga uno de sus extremos en X_1 , y el otro en X_2 ; aquí $X_1 \cap X_2 = \emptyset$. Un grafo bipartido $G = (X_1 \cup X_2, A)$ se llama *completo*, si para cualesquiera dos vértices $x_i \in X_1$ y $x_j \in X_2$ existe una arista $(x_i, x_j) \in A$. Si $|X_1| = n$, y $|X_2| = m$, el grafo completo no orientado $G = (X_1 \cup X_2, A)$ se denota con $K_{n,m}$. Un grafo orientado G se llama *conexo (bipartido, respectivamente)*, si lo es su sosia no orientada \bar{G} .

Se llama *grado* del vértice en un grafo no orientado al número de aquellas aristas de cuyo extremo sirve el vértice.

Examinemos un grafo no orientado $G = (X, A)$ sin lazos. Elijamos una arista $a = (x, y) \in A$. Sustituyamos el vértice y en todas las aristas $b \in (A \setminus \{a\})$ por el vértice x y denotemos con A' el conjunto obtenido de aristas. Diremos que el grafo $G' = (X \setminus \{y\}, A')$ se ha obtenido del grafo G como resultado de contraer la arista a al vértice x . Un grafo G se denomina *planario*, si puede ser dibujado en un plano de un modo tal que dos aristas arbitrarias del grafo no se intersequen. En la fig. 5.10 están expuestos el grafo completo K_5 y el grafo completo bipartido $K_{3,3}$, los cuales, como se sabe, no son planarios. Estos dos grafos tienen gran importancia en la teoría de los grafos planarios y se conocen como grafos de Kuratovski, por cuanto fue él el primero en establecer que un grafo no orientado G no es planario, si y sólo si contiene un subgrafo contraído al grafo K_5 ó al $K_{3,3}$.

Matriz de adyacencia. Sea dado un grafo orientado simple G con los vértices x_1, x_2, \dots, x_n . Se llama *matriz de adyacencia* de dicho grafo a

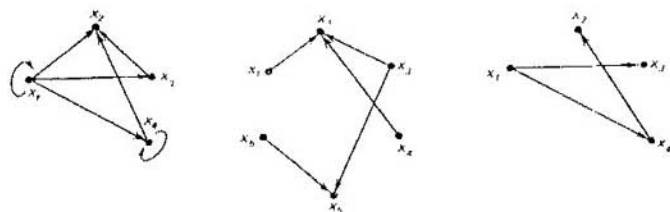


Fig. 5.9.

una matriz cuadrada $B = [b_{ij}]$ de orden n , donde

$$b_{ij} = \begin{cases} 1, & \text{si en } G \text{ existe el arco } (x_i, x_j); \\ 0, & \text{si en } G \text{ no hay arco } (x_i, x_j). \end{cases}$$

Así pues, la matriz de adyacencia del grafo expuesto en la fig. 5.8. tiene la forma siguiente

$$B = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ x_1 & 1 & 1 & 1 & 1 & 0 & 1 \\ x_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ x_3 & 0 & 1 & 0 & 0 & 1 & 0 \\ x_4 & 0 & 1 & 0 & 1 & 0 & 0 \\ x_5 & 0 & 0 & 0 & 1 & 0 & 0 \\ x_6 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

La matriz de adyacencia define completamente la estructura del grafo. Por ejemplo, la suma de todos los elementos de la fila x_i de la matriz B da el número de arcos que tienen el vértice x_i como su vértice original, y la suma de elementos de la columna x_i da el número de arcos en los que x_i figura como vértice final. El conjunto de columnas que tienen 1 en la fila x_i es un conjunto $\Gamma(x_i)$, y el conjunto de filas que tienen 1 en la columna x_i coincide con el conjunto $\Gamma^{-1}(x_i)$.

Elevemos al cuadrado la matriz de adyacencia. Sea $b_{ik}^{(2)}$ un elemento de la matriz B^2 , entonces, $b_{ik}^{(2)} = \sum_{j=1}^n b_{ij}b_{jk}$. El sumando $b_{ij}b_{jk}$ es igual a 1, cuando y sólo cuando ambos números b_{ij} y b_{jk} son iguales a 1; de lo contrario, es igual a 0. De la igualdad $b_{ij} = b_{jk}$ se desprende la existencia del

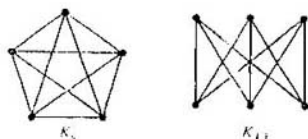


Fig. 5.10

camino de longitud 2 del vértice x_i al vértice x_k , pasando por el vértice x_j . Por consiguiente, $b_{ik}^{(2)}$ en la matriz B^2 es igual al número de caminos de longitud 2 que van de x_i a x_k .

Análogamente, el elemento $b_{ik}^{(p)}$ de la matriz B^p es igual al número de caminos (no forzosamente simples) de longitud p que van de x_i a x_k .

La matriz de adyacencia de un grafo no orientado se construye de un modo análogo.

Matriz de incidencias. Sea dado un grafo orientado G con los vértices x_1, x_2, \dots, x_n y los arcos a_1, a_2, \dots, a_m . Se llama matriz de incidencias del grafo G a una matriz $A = [a_{ij}]$ de dimensión $n \times m$, cuyas filas corresponden a los vértices, y las columnas, a los arcos, y

$$a_{ij} = \begin{cases} 1 & \text{si el vértice } x_i \text{ es el origen del arco } a_j; \\ -1, & \text{si el vértice } x_i \text{ es el extremo del arco } a_j; \\ 0, & \text{si el vértice } x_i \text{ no es incidente con el arco } \\ & a_j, \text{ o si el arco } a_j \text{ no es un lazo.} \end{cases}$$

Por ejemplo, para el grafo que se expone en la fig. 58 la matriz de incidencias es de la forma

$$A = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{matrix} & \left\{ \begin{array}{cccccccccccc} 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{array} \right. \end{matrix}$$

Por cuanto cada arco es incidente con dos diferentes vértices, a excepción del caso en que un arco forma un lazo, cada columna contiene o bien un solo elemento igual a 1, y un elemento igual a -1 , o bien todos los elementos iguales a 0.

Si G es un grafo no orientado, su matriz de incidencias se determina igual que antes, salvo que todos los elementos iguales a -1 se sustituyen por $+1$.

Matriz de ciclos (de contornos). Sea G un grafo no orientado. Demos a sus aristas los números a_1, a_2, \dots, a_m , y a los ciclos simples, los números c_1, c_2, \dots, c_p . Se llama matriz de ciclos del grafo G a una matriz $C = [c_{ij}]$ de dimensión $p \times m$, donde

$$c_{ij} = \begin{cases} 1, & \text{si el ciclo simple } c_i \text{ del grafo } G \text{ contiene la arista } a_j; \\ 0, & \text{en el caso contrario.} \end{cases}$$

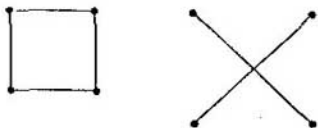


Fig. 5.11.

Así pues, las filas de una matriz de ciclos corresponden a los ciclos simples, y las columnas, a las aristas.

Si G es un grafo orientado, numeremos sus arcos con a_1, a_2, \dots, a_m , y los contornos simples, con c_1, c_2, \dots, c_p ; la matriz de contornos $C = [c_{ij}]$ se determina de un modo análogo:

$$c_{ij} = \begin{cases} 1, & \text{si el contorno } c_i \text{ contiene un arco } a_j \text{ y los mismos estan igual-} \\ & \text{mente orientados;} \\ -1, & \text{si el contorno } c_i \text{ contiene un arco inverso de } a_j; \\ 0, & \text{en todos los demas casos.} \end{cases}$$

Observacion. La eleccion de sımbolos iguales para denotar las matrices de los grafos orientado y no orientado se debe a que tanto los grafos orientados, como los no orientados, poseen iguales propiedades en lo que se refiere a la estructura de estas matrices. Por eso, no hay confusion ninguna al introducir tales sımbolos. Ademas, las propiedades de las matrices de los grafos orientados y los metodos de su demostracion son los mismos que para las matrices de los grafos no orientados.

Ejercicios. Cerciorese de que son validas las siguientes afirmaciones.

1. En todo grafo no orientado finito el numero de vertices de grado impar es par.
2. Sea A una matriz de incidencias de un grafo no orientado simple sin lazos. Entonces, la matriz de adyacencia B se obtiene de AA^T (donde A^T es una matriz obtenida por transposicion de la matriz A) por sustitucion de todos los elementos en la diagonal por ceros.
3. Un grafo G es bipartido, cuando y solo cuando para todo numero impar n todos los elementos diagonales de la matriz A^n son nulos.
4. Si los grafos G_1 y G_2 son isomorfos, los valores propios de las matrices de adyacencia de estos grafos son iguales.

Observacion. La afirmacion recıproca no es valida en el caso general. Por ejemplo, dos grafos expuestos en la fig. 5.11 tienen iguales valores propios ($\lambda = \pm 2, 0, 0, 0$), sin embargo, evidentemente, ellos no son isomorfos.

5. Cualquier menor de la matriz de incidencias del grafo G es igual a $+1, -1, \delta$ o 0 .
6. El rango de una matriz de incidencias A de un grafo conexo (tanto orientado, como no orientado) G con n vertices es igual a $n - 1$.
7. El rango de una matriz de ciclos C de un grafo conexo G con m aristas y n vertices es igual a $m - n + 1$.
8. Una matriz de incidencias A es ortogonal respecto de la matriz C^T , transpuesta a la matriz de ciclos C y obtenida con el mismo orden de numeracion de las aristas que en la matriz de incidencias A , es decir, dicho de otro modo, $AC^T \equiv 0 \pmod{2}$.

Se llama *grafo euleriano* a un grafo no orientado que posee la siguiente propiedad: existe un ciclo que pasa exactamente una vez por cada una de las aristas del grafo. El ciclo en la definicion de grafo euleriano se denomina *euleriano*.

Teorema 8. Un grafo no orientado finito G es euleriano cuando y sólo cuando es conexo y todos los vértices suyos son de grado par.

Demostración. Es evidente que nuestras condiciones son necesarias, puesto que cada vez que el ciclo euleriano pasa por algún vértice, ha de entrar en él por una arista y salir por la otra.

A la inversa, supongamos ahora que G es conexo y todos los vértices suyos son de orden par. Convengamos en considerar un punto arbitrario x del grafo G como origen de la cadena P y prolonguémosla, cuanto sea posible, todo el tiempo a través de nuevas aristas. Por cuanto en cada vértice el número de aristas es par, el proceso puede terminar sólo en x . Si P contiene no todas las aristas del grafo G , eliminemos de G una parte de P compuesta por las aristas de este ciclo.

Los grafos P y G tienen vértices de grado par; lo mismo ha de ser válido también para el grafo restante \bar{P} . Por cuanto el grafo G es conexo, en \bar{P} debe encontrarse un vértice y que sea incidente con las aristas de \bar{P} . A partir de y puede construirse una cadena nueva P' que contenga sólo aristas de \bar{P} . De nuevo tal cadena quedará terminada sólo cuando vuelva a y . Pero en este caso de P y P' podemos hacer un ciclo nuevo

$$P_1 = P(x, y) \cup P' \cup P(y, x),$$

el cual vuelve a x y contiene más aristas que P . Si P_1 no es un ciclo euleriano, la construcción se repite. Finalizado este proceso, el ciclo euleriano será construido.

Se denomina *ciclo hamiltoniano (contorno hamiltoniano)* a un ciclo simple (contorno simple) que pasa por todos los vértices del grafo. Un grafo se llama *hamiltoniano*, si contiene un ciclo hamiltoniano.

En las aplicaciones de los grafos a los juegos los vértices corresponden a posiciones diferentes. Así pues, la existencia del ciclo hamiltoniano es equivalente a la de una sucesión cíclica de jugadas que contiene cada posición una sola vez. De ejemplo sirve el conocido problema del caballo ajedrecístico: ¿se podrá realizar una jugada con el caballo, comenzando con un escaque arbitrario en el tablero de ajedrez, en una sucesión tal que permita pasar por cada uno de los 64 escaques y regresar al escaque de partida. En la fig. 5.12 se muestra una de las soluciones posibles.

56	41	58	35	50	39	60	33
47	44	55	40	59	34	51	38
42	57	46	49	36	53	32	61
45	48	43	54	31	62	37	52
20	5	30	63	22	11	16	13
29	64	21	4	17	14	25	10
6	19	2	27	8	23	12	15
1	28	7	18	3	26	9	24

Fig. 5.12.

En la fig. 5.13 se aduce el ejemplo de un grafo euleriano que no es

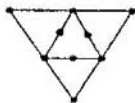


Fig 5.13.

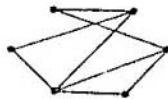


Fig.5.14

hamiltoniano, y en la fig. 5.14, de un grafo hamiltoniano, mas no euleriano.

Pese a la similitud de las definiciones para los ciclos euleriano y hamiltoniano, las teorías correspondientes para estos conceptos tienen poco de común. El criterio de existencia de los ciclos eulerianos es simple (véase el teorema 8); para los ciclos hamiltonianos no se conoce una regla general. Más aún, incluso para los grafos concretos resulta, a veces, difícil decidir si es posible hallar tal ciclo. Algunas condiciones, bajo las cuales en un grafo existe el ciclo hamiltoniano, sin embargo, han sido encontradas, y se dan en ciertas obras referentes a la teoría de los grafos. Para aquellos que desean familiarizarse con estos resultados les recomendamos los libros [24], [25].

Se denomina *árbol* un grafo conexo no orientado sin ciclos. Se denomina *bosque* un grafo no orientado en el que cada componente representa un árbol.

Teorema 9. Las siguientes afirmaciones se consideran equivalentes:

- un grafo no orientado G es el árbol;
- cualesquiera dos vértices en G están unidos por una cadena única;
- un grafo G es conexo, mas la supresión de cualquiera de sus aristas lo hace no conexo.
- la adición de cualquier arista nueva al grafo G conduce a que aparezca exactamente un ciclo.

La demostración de este teorema se omite por ser trivial.

El siguiente resultado sobre los árboles pertenece a los trabajos de Cayley que investigó estos grafos en relación con las fórmulas estructurales químicas.

Teorema 10. El número de diferentes árboles que pueden ser construidos sobre n vértices dados es igual a n^{n-2} .

Demostración. Numeremos todos los vértices de un árbol T con números naturales de 1 a n . Designemos con b_1 un vértice pendiente (es decir, un vértice de grado 1) con un número mínimo, y con $l_1 = (b_1, a_1)$, la correspondiente arista pendiente. Al suprimir en T la arista l_1 y el vértice b_1 , obtendremos un nuevo árbol T_1 . Hallemos para T_1 un vértice pendiente con un número mínimo; designémoslo con b_2 , y la arista correspondiente, con $l_2 = (b_2, a_2)$. Esta reducción se repite hasta que no quede, tras la eliminación de la arista $l_{n-2} = (b_{n-2}, a_{n-2})$, la única arista $l_{n-1} = (b_{n-1}, a_{n-1})$ que une los dos vértices restantes. Entonces, la siguiente colección entre paréntesis

$$(a_1, a_2, \dots, a_{n-2})$$

llamada *código de Prufer*, se define unívocamente por el árbol T , y a dos árboles distintos T y T' les corresponden, evidentemente, diferentes códigos de Prufer.

Además, los códigos de Prufer definen los árboles T con ayuda de una construcción inversa. Si está dado el código, entonces se halla el primer vértice b_1 no contenido en él. Esto define la arista $l_1 = (b_1, a_1)$. Luego, eliminamos los vértices a_1 del código y b_1 , de la sucesión $1, 2, \dots, n$, y continuamos la construcción para los números restantes. El grafo, obtenido como resultado de esta construcción será un árbol, lo que puede ser establecido, por ejemplo, con ayuda de la inducción. Después de eliminar a_1 , el código contendrá $n - 3$ números. Si éstos corresponden al árbol T_1 , entonces el grafo T , obtenido de T_1 por adición de la arista $l_1 = (b_1, a_1)$, es también un árbol, puesto que el vértice b_1 no pertenece a T_1 . Así pues, se ha establecido que entre los códigos de Prufer y los árboles hay una correspondencia biunívoca. Mas, en el código de Prufer cada elemento puede asumir cualquiera de n valores posibles. Todos ellos corresponden a diferentes árboles. En total pueden haber n^{n-2} códigos de Prufer, de lo que precisamente se deduce nuestro teorema.

Si G es un grafo no orientado con un conjunto de vértices X , $|X| = n$, entonces, se llamará *árbol de esqueleto* (o, simplemente, esqueleto) del grafo G a todo subgrafo de esqueleto del grafo G que sea un árbol. Las aristas del grafo, pertenecientes al esqueleto, llevan el nombre de *ramas* y todas las demás aristas, se denominan *cuerdas*.

Hay situaciones cuando surge la necesidad de confeccionar una lista completa de los esqueletos del grafo G (por ejemplo, en el caso cuando hay que elegir el "mejor" árbol, mientras que el criterio que permite realizar tal elección es muy complejo, así que la resolución inmediata del problema de optimización resulta ser ejecutable). En otras situaciones, por ejemplo al hallar las funciones de transferencia de un sistema o al calcular los determinantes de ciertas matrices en la teoría macroeconómica, podemos, con ayuda de todos los esqueletos del grafo correspondiente, conseguir la simplificación de los procedimientos de cálculo.

El número de diferentes esqueletos de un grafo marcado conexo no orientado completo sobre n vértices nos da el teorema 10. Las fórmulas para el número de esqueletos en los grafos más generales pueden encontrarse en el libro [14].

He aquí uno de los resultados.

Teorema 11. Sea G un grafo no orientado sin lazos sobre n vértices, y sea A_0 su matriz de incidencias con una fila eliminada (es decir, una matriz con $n - 1$ filas independientes). Supongamos que A_0' es una matriz transpuesta respecto de la matriz A_0 . Entonces, el determinante del producto $A_0 A_0'$ es igual al número de diferentes esqueletos del grafo G , un doble (sosa) no orientado de G .

La demostración de este teorema se aduce en el libro [78]. Los algorit-

mos que generan todos los árboles de esqueleto del grafo están detalladamente analizados en [24].

Todos los árboles de esqueleto del grafo expuesto en la fig. 5.15 (hay en total 21 árboles) se aducen en la fig. 5.16. La matriz de incidencias del grafo tiene en este caso por expresión.

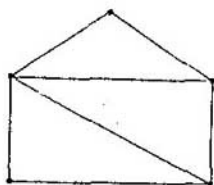


Fig. 5.15.

$$A = \begin{matrix} & \begin{matrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 \end{pmatrix} \end{matrix}$$

Aquí consideramos que cada arista está orientada de su vértice final de índice mínimo al vértice de índice máximo.

Eliminando, por ejemplo, la fila x_2 , obtenemos la matriz A_0 . El producto de las matrices $A_0 A_0^T$ tiene en este caso la forma

$$A_0 A_0^T = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

El determinante del producto es igual a 21. Por consiguiente, en la fig. 5.16 se da la lista completa de los esqueletos del grafo G , expuesto en la fig. 5.15.

Muchos esqueletos obtenidos son, evidentemente, isomorfos y surge la cuestión sobre el número que hay entre ellos no isomorfos. Este es un problema más complejo, sin embargo es muy importante para muchas aplica-

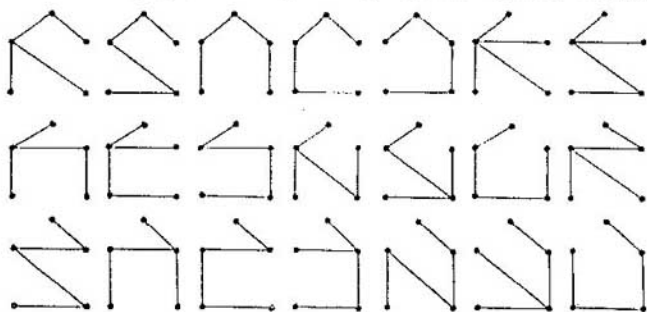


Fig. 5.16.

ciones. Es por eso que existe una literatura espaciosa también para las cuestiones análogas relacionadas con la enumeración de todos los grafos de tipos parciales. La mayoría de estas obras se apoya en la teoría de Redfield—Polya (véase § 25).

Se llama *árbol orientado* con raíz x_0 a un grafo orientado conexo, en el que del vértice x_0 a cada otro vértice conduce un camino y, además, único.

Sea G un grafo no orientado con n vértices, m aristas y p componentes conexas. El número $\varrho(G) = n - p$ se denomina *cociclomático*. El número $\nu(G) = m - \varrho(G) = m - n + p$ se denomina *ciclomático*.

En la teoría de los circuitos eléctricos los números $\varrho(G)$ y $\nu(G)$ tienen un significado físico directo. Así, por ejemplo, el número ciclomático es igual al número máximo de contornos independientes en el grafo del circuito eléctrico, es decir, al número máximo de corrientes circulares independientes que pueden fluir en el circuito. El número cociclomático es igual al número máximo de diferencias de potencial independientes entre los nudos del circuito eléctrico.

Ejercicios 9. Dense ejemplos de árboles orientados con raíces.

10. Demuéstrese que el número ciclomático de un árbol es igual a 0.

11. Demuéstrese que cada árbol es un grafo bipartido.

Sea T un esqueleto del grafo G . En el caso general, $\nu(G)$ ciclos que se obtienen por adición de cualquier cuerda de G a las ramas de T , llevan el nombre de *ciclos fundamentales*.

Cabe notar que aunque el número de ciclos fundamentales es igual a $\nu(G)$, estos propios ciclos están definidos multiformemente y dependen del esqueleto T elegido originalmente. Dicho de otro modo, el ciclo fundamental (respecto del esqueleto T) es un ciclo obtenido por adición de una cuerda al esqueleto T . Por ejemplo, para un grafo G y para su esqueleto T , representados en la fig. 5.17, los ciclos fundamentales serán:

$$\begin{aligned}\Phi_1 &= \{a_1, a_{10}\}; \Phi_2 = \{a_2, a_{12}, a_{10}\}; \Phi_3 = \{a_3, a_{11}, a_{13}, a_{12}, a_{10}\}; \\ \Phi_4 &= \{a_4, a_{14}, a_{11}\}; \Phi_5 = \{a_5, a_{11}\}; \Phi_6 = \{a_6, a_{15}, a_{13}, a_{14}\}; \\ \Phi_7 &= \{a_7, a_{13}, a_{14}\}; \Phi_8 = \{a_8, a_{12}, a_{15}\}; \Phi_9 = \{a_9, a_{12}\}.\end{aligned}$$

Si una familia de todos los ciclos fundamentales la designamos con Φ (en nuestro ejemplo $\Phi = \{\Phi_1, \dots, \Phi_9\}$), entonces cualquier otro ciclo del grafo que no pertenece a Φ puede ser expresado en forma de una combinación lineal de ciclos pertenecientes a Φ , siempre que convenimos en lo siguiente.

Supongamos que cada ciclo fundamental Φ_i ; $i = 1, 2, \dots, \nu(G)$, está representado por un vector m -dimensional en el que la j -ésima componente es igual a 1 ó a 0, lo que depende de si pertenece o no la j -ésima arista al ciclo dado. Entonces, todo ciclo C del grafo G puede representarse como suma respecto del módulo 2 de los ciclos fundamentales. Así por ejemplo, el ciclo $C_1 = \{a_3, a_{11}, a_{14}, a_6, a_8, a_1\}$ de nuestro grafo puede ser representado en la forma

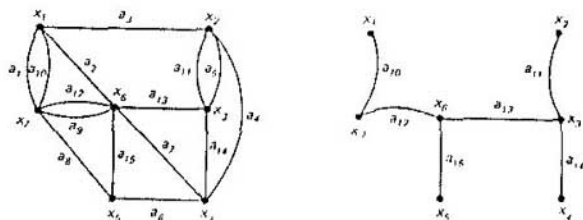


Fig. 5.17.

$$C = \Phi_1 + \Phi_3 + \Phi_6 + \Phi_8 \pmod{2} =$$

$$+ \left\{ \begin{array}{l} 10000000100000 \\ 00100000011100 \\ 000001000000111 \\ 000000010001001 \end{array} \right\} \pmod{2}.$$

$$= \left\{ \begin{array}{l} 101001010010010 \end{array} \right\}$$

Cabe notar que la inversión de la afirmación mencionada no es válida, a saber, cierta suma respecto del módulo 2 de los ciclos fundamentales no da obligatoriamente el único ciclo, pero puede representar dos y más ciclos. Por ejemplo, la suma $\Phi_2 + \Phi_3 + \Phi_6 + \Phi_7 \pmod{2}$ corresponde a dos ciclos simples $\{a_2, a_3, a_{11}, a_{13}\}$ y $\{a_6, a_7, a_{15}\}$. Así pues, para generar todos los ciclos simples del grafo G no es menester tomar todas las $2^{r(G)} - 1$ combinaciones de los ciclos fundamentales y sumarlas respecto del módulo 2; algunas de estas sumas no serán, de hecho, ciclos. Más aún, si la suma dada no genera un ciclo, no se pueden rechazar otras sumas que la contienen, puesto que, sumando respecto del módulo 2 con otra suma, podemos obtener un ciclo simple.

Subrayemos también que en el grafo G podemos encontrar un conjunto de $\nu(G)$ ciclos simples independientes, que no pueden obtenerse por adición de las aristas al árbol, como lo hacíamos antes. De tal conjunto no se debe decir que es fundamental. En la fig. 5.18 se muestra el conjunto de $\nu(G) = 4$ ciclos simples independientes del grafo G , que no puede obtenerse por adición de las aristas a ningún esqueleto del grafo G y el cual, por esta razón, no es un conjunto fundamental.

Se denomina *corte* del grafo un conjunto mínimo de aristas cuya eliminación aumenta el número de componentes del grafo.

El *corte fundamental* (respecto del esqueleto T) es aquél que contiene exactamente una arista del esqueleto T . De este modo, en un grafo conexo pueden distinguirse $n - 1$ cortes fundamentales.

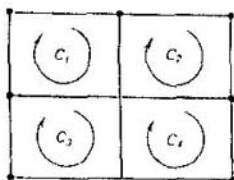


Fig. 5.18.

El teorema que sigue más abajo establece una relación entre los cortes fundamentales y los ciclos fundamentales, proporcionando el método de construcción de los cortes fundamentales.

Teorema 12. Si T es el esqueleto de un grafo no orientado G , el corte fundamental definido por la rama a_i de T está formado por a_i y aquellas cuerdas de G , las cuales, siendo adicionadas a T , proporcionan ciclos fundamentales que contienen a_i .

Demostración. Si eliminamos la arista a_i del esqueleto T , este último se descompone en dos subárboles: T_1 y T_2 . Cualquier arista, uno de cuyos vértices extremos yace en T_1 , y el otro, en T_2 , debe pertenecer al corte fundamental, puesto que la adición de cualquiera de tales aristas a las aristas de T_1 y T_2 conduce a la formación del otro esqueleto del grafo G , y, por consiguiente, cualquier conjunto privado de tales aristas no será un corte. El conjunto de estas aristas junto con la arista a_i es un corte, puesto que su eliminación divide el grafo en dos subgrafos, uno de los cuales tiene como conjunto de sus vértices T_1 , y el otro, T_2 . Esto quiere decir que este corte es fundamental. Más aún, por cuanto la arista a_i es la única por la que pasan las cadenas del esqueleto T , que tienen por origen los vértices de T_1 y termina en los vértices de T_2 las únicas aristas que cierran los ciclos fundamentales con la arista a_i incluida, serán aquellas, uno de cuyos vértices extremos yace en T_1 , y el otro en T_2 . De esta forma queda demostrado el teorema.

Sea T un esqueleto del grafo no orientado G . Se denomina *matriz de los ciclos fundamentales* Φ a una submatriz de la matriz de los ciclos C , cuyas filas corresponden a los ciclos simples, definidos por las cuerdas que están unidas con el esqueleto dado T . Dicho de otro modo, si T es un esqueleto, entonces se llamará matriz de los ciclos fundamentales del grafo G a la matriz $\Phi = [\varphi_{ij}]$, compuesta de $\nu(G)$ filas y m columnas, en la cual

$$\varphi_{ij} = \begin{cases} 1, & \text{si la arista } a_j \text{ pertenece al ciclo fundamental } \Phi_i; \\ 0, & \text{en el caso contrario.} \end{cases}$$

Si numeramos las cuerdas del grafo G en una forma sucesiva desde 1 hasta $\nu(G)$, y las ramas, de $(\nu(G) + 1)$ a m , entonces la matriz de los ciclos fundamentales tendrá por expresión

$$\Phi = [E|\Phi_{12}],$$

donde E es una matriz unidad de orden $\nu(G)$, y Φ_{12} , una matriz de dimensión $\nu \times (m - \nu)$. Esto se explica por el hecho de que cada ciclo Φ_i contiene una y sólo una cuerda y los ciclos siempre pueden ser numerados según el número de cuerdas, a consecuencia de lo cual todas las unidades en la primera $(\nu \times \nu)$ -submatriz de la matriz Φ se hallan en la diagonal.

La *matriz de los cortes fundamentales* $K = [k_{ij}]$ se define como matriz con $n - 1$ filas y m columnas, donde

$$k_{ij} = \begin{cases} 1, & \text{si la arista } a_j \text{ pertenece al corte fundamental } k_i; \\ 0, & \text{en el caso contrario.} \end{cases}$$

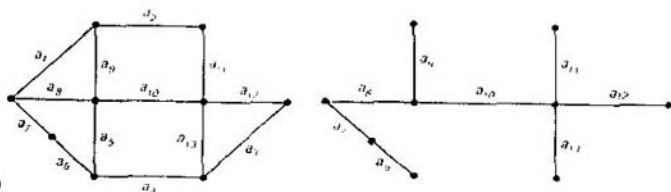


Fig. 5.19

Para la misma numeración de las aristas que en la matriz de los ciclos fundamentales Φ , la matriz K tendrá por expresión

$$K = [K_{ij}/E],$$

puesto que en este caso todo corte fundamental tiene una y sólo una rama de T .

Admitamos que el grafo G y su esqueleto T están representados en la fig. 5.19. Entonces la matriz de los ciclos fundamentales es igual a

$$\Phi = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} \\ \Phi_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ \Phi_2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \Phi_3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \Phi_4 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ \Phi_5 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

y la matriz de los cortes fundamentales K es igual a

$$K = \begin{pmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & a_{12} & a_{13} \\ k_1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ k_2 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ k_3 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ k_4 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ k_5 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ k_6 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ k_7 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ k_8 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

donde los cortes k_i , correspondientes a las aristas del esqueleto, tienen los números

$$\mu(G) + i = 5 + i; \quad i = 1, 2, \dots, 8.$$

Existen varias correlaciones interesantes entre las matrices de los ciclos fundamentales y de los cortes fundamentales y las incidencias de los grafos no orientados sin lazos.

Teorema 13. La matriz de incidencias A_0 (con una fila suprimida) y la matriz transpuesta de los ciclos fundamentales Φ^T son ortogonales, es decir, $A_0 \Phi^T \equiv 0 \pmod{2}$.

Teorema 14. La matriz de los ciclos fundamentales Φ y la matriz transpuesta de los cortes fundamentales K^t son ortogonales, es decir, $\Phi \cdot K^t \equiv 0 \pmod{2}$.

Los teoremas 13 y 14 son consecuencia de dos hechos evidentes:

1) cada vértice en un ciclo es incidente al número par de aristas de este ciclo, y en el caso de un ciclo simple, a dos aristas;

2) cada corte de un ciclo, inducido por cierto corte, tiene un número par de aristas, comunes con dicho corte.

El teorema 13 se deduce del hecho 1), y el 14, del hecho 2), si recordamos que todas las operaciones se estudian con relación al módulo 2.

En lo que se refiere al teorema 14, podemos escribir

$$\Phi \cdot K^t = [E] \Phi_{12} \begin{Bmatrix} K_{11} \\ E \end{Bmatrix} = K_{11}^t + \Phi_{12} = 0 \pmod{2}.$$

Por eso, $K_{11}^t = -\Phi_{12} = \Phi_{12} \pmod{2}$. En otras palabras, la matriz de los cortes fundamentales puede ser obtenida en cuanto se conozca la matriz de los ciclos fundamentales, y viceversa.

Ejercicios. Cierciórese de la validez de las siguientes afirmaciones

12. Una submatriz cuadrada de dimensión $(n-1) \times (n-1)$ de la matriz de incidencias I de un grafo conexo es regular, si las columnas de la submatriz citada corresponden a las ramas de cierto árbol.

13. Supongamos que la matriz de ciclos C de un grafo conexo tiene el rango $n - m + 1$. Su submatriz cuadrada B de orden $n - m + 1$ es regular cuando y sólo cuando las columnas de la submatriz citada corresponden al conjunto de cuerdas respecto de cierto árbol de esqueleto del grafo dado.

14. Una submatriz regular de dimensión $(n-1) \times (n-1)$ de la matriz de incidencias I corresponde biunívocamente a los espáculos del grafo.

15. Una matriz regular de dimensión $(n-m+1) \times (n-m+1)$ de la matriz de ciclos C corresponde biunívocamente a los complementos de los espáculos del grafo.

16. Si las columnas de las matrices A_0 , Φ y K las escribimos valiéndonos de un mismo orden de aristas con relación a cierto árbol de esqueleto, respecto del cual están formados los ciclos y cortes fundamentales, y, a continuación, las representamos en forma de $A_0 = [A_0]A_{12}$, donde A_{12} es una matriz cuadrada regular de orden $n-1$, $\Phi = [E]\Phi_{12}$ y $K = \{K_{11}\}E$, entonces

$$K_{11} = \Phi_{12}^t A_{12}^{-1} A_{11}.$$

Con esto hemos de acabar con la introducción a la teoría de los grafos. Esta teoría ha acumulado un enorme material de hechos reales. A la exposición de esta teoría están dedicados toda una serie de libros. Algunos de ellos ya han sido mencionados en el texto. Sin embargo, debemos atraer la atención del lector a la obra [79] (probablemente la única de su género) sobre el origen y desarrollo de la teoría de los grafos.

CAPÍTULO 6 MÉTODOS DE RESOLUCIÓN DE LOS PROBLEMAS EXTREMALES

6.1. PLANTEAMIENTO DE LOS PROBLEMAS COMBINATORIOS EXTREMALES Y ACCESOS A SU RESOLUCIÓN

Según lo dicho en el capítulo 1, existen tres tipos de problemas combinatorios: en unos se resuelve la cuestión de existencia o no existencia de las soluciones; en otros se calcula el número de soluciones del problema; por fin, en los terceros de una colección de todas las soluciones se elige aquella que posee cierta propiedad en un grado máximo o mínimo. Los problemas de este último tipo se denominan precisamente extremales.

Demos a conocer unos cuantos ejemplos de problemas combinatorios extremales.

1. Problema de nombramientos. Se tienen n plazas vacantes y el mismo número de candidatos para ocuparlas. El nombramiento del i -ésimo candidato a la j -ésima plaza lleva consigo el pago del salario c_{ij} ; $i, j = 1, 2, \dots, n$. Para cada candidato puede ser asignada sólo una plaza, y cada plaza puede ser ocupada sólo por un candidato. En otras palabras, no se admiten el ejercer un cargo simultáneamente con otro y el fraccionamiento de los cargos. El problema consiste en llevar a cabo los nombramientos de un modo tal que la economía de los fondos de salario sea máxima, es decir, en encontrar

$$\min \sum_{i=1}^n c_{ij_i}$$

donde j_i es el cargo que obtiene el i -ésimo candidato, y (j_1, j_2, \dots, j_n) es una sustitución correspondiente a los nombramientos: $i \rightarrow j_i$. Si interpretamos c_{ij} como rendimiento, el problema consistirá en la búsqueda del máximo y no del mínimo. Los problemas de nombramientos tienen muchas formas diferentes en función de las condiciones. Por ejemplo, para ejecutar n operaciones independientes están elegidos m obreros; se da una matriz $\{t_{ij}\}$, donde t_{ij} es el tiempo que gasta el i -ésimo obrero para realizar la j -ésima operación. La magnitud, cuyo mínimo se busca, es el tiempo de realización de todo el trabajo:

$$F = \min_i E_i; \quad E_i = \sum_{j \in R_i} t_{ij}$$

donde R_i es el conjunto de operaciones ejecutadas por el i -ésimo obrero.

En un caso particular, puede prefijarse el número de operaciones n_i para el i -ésimo obrero: $\sum_i n_i = n$. Cuando $n_i = 1$ ($i = 1, 2, \dots, n$; $n = m$), obtenemos el problema de nombramientos con el criterio de mini-máx. Por último, demos a conocer tal formulación del problema de nombramientos

cuando se toma en consideración el tiempo que gasta un obrero para pasar de una operación a la otra. Sean dadas n operaciones y m obreros. Se hacen las matrices: $T = [t_{ij}]$, donde t_{ij} es el tiempo durante el cual el i -ésimo obrero cumple la j -ésima operación, y $L = [l_{ij}]$, donde l_{ij} es el tiempo que gasta cualquier obrero para pasar de la i -ésima operación a la j -ésima. Se requiere distribuir las operaciones entre los obreros y de tal manera considerar el tiempo de paso, que el tiempo de ejecución de todo el complejo de operaciones sea mínimo.

2. Problema de un viajante de comercio. Un viajante de comercio tiene que visitar varias ciudades. Ha de elegir la ruta más corta para que empezando a viajar desde su ciudad, pueda visitar otras ciudades una sola vez y regresar. Las distancias entre las ciudades calculadas de dos en dos vienen dadas en forma de una matriz $C = [c_{ij}]$, $i, j = 1, 2, \dots, n$, donde n es el número de ciudades.

3. Problema de una mochila. Se tienen n objetos cuyos pesos son a_1, a_2, \dots, a_n , y el valor s_1, s_2, \dots, s_n , respectivamente. Se pide llenar la mochila, capaz de contener un peso no superior a R , con un surtido de objetos que posea un valor máximo.

Construyamos un vector $X = (x_1, x_2, \dots, x_n)$, donde $x_i = 1$, si el i -ésimo objeto debe colocarse en la mochila, y $x_i = 0$; en el caso contrario; $i = 1, 2, \dots, n$. El problema se formulará como problema de búsqueda de

$$\max \sum_{i=1}^n s_i x_i$$

(del valor máximo de los objetos colocados) a condición de que

$$\sum_{i=1}^n x_i a_i \leq R.$$

4. Minimización del tiempo de ejecución de un juego de operaciones. Supongamos que se pide ejecutar n operaciones. A cada operación se le hace corresponder un vértice del grafo G , donde está trazado un arco desde el vértice i hacia el j , si la operación i precede, según las condiciones tecnológicas, a la operación j . Todas las operaciones están divididas en m clases, con la particularidad de que en un momento dado puede ejecutarse sólo una operación de cada clase. Se conoce el tiempo de ejecución de las operaciones (t_j para la j -ésima operación). Se requiere determinar el orden de ejecución de todas las operaciones que asegure la realización del trabajo en el tiempo mínimo T . Por ejemplo, sea una red (esquema) de 6 operaciones: A_1, A_2, A_3, A_4 son operaciones de primera clase y B_1, B_2 , de segunda; el tiempo de ejecución de cada operación viene indicado dentro de los cuadrados en la fig. 6.1. Sean R_1 y R_2 sucesiones de ejecución de las operaciones para cada clase, respectivamente. Por ejemplo, para $R_1 = (2, 1, 3, 4)$; $R_2 = (2, 1)$ obtendremos: $T = 36$. Es evidente que el número de combinaciones admisibles depende considerablemente de la estructura de la red.

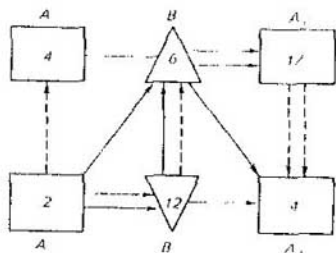


Fig. 6.1

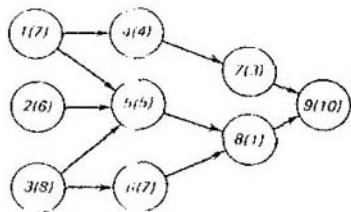


Fig. 6.2.

5. Minimización del tiempo medio de mecanizado de un lote de piezas.

Se analiza un grafo orientado G con los vértices $1, 2, \dots, n$, donde el vértice i corresponde al i -ésimo tipo de las piezas a maquinar. Para cada arco (i, j) viene dada su longitud $l_{i,j}$, la cual se interpreta como período entre el mecanizado de las piezas del i -ésimo tipo y el mecanizado de las piezas del j -ésimo tipo. La sucesión de vértices $i_1, i_2, \dots, i_k, i_{k+1}$, donde $i_{k+1} = i_1$ define el ciclo del mecanizado. Se pide minimizar el tiempo medio de este proceso

$$T_{\text{m}} = \frac{1}{k} \sum_{j=1}^k l_{i_j, i_{j+1}}$$

es decir, hallar en el grafo orientado G el contorno que tenga una longitud media mínima.

6. Minimización de la parada de una línea de montaje. Supongamos que n operaciones se efectúan en un transportador de m tramos. Para que el transportador se mueva continuamente, el tiempo de trabajo del operario no debe sobrepasar la duración del ciclo. Admitamos que en el i -ésimo tramo se realizan un conjunto de operaciones R_i . Entonces, el tiempo activo en el i -ésimo tramo es $t_i = \sum_{\mu \in R_i} t_\mu$ y la parada (interrupción) en él es igual a

$d_i = C - \sum_{\mu \in R_i} t_\mu$, donde C es la duración total del ciclo de operaciones. La

magnitud sumaria de la parada $d = \sum_{i=1}^m d_i$ debe hacerse mínima a cuenta de la distribución de las operaciones por tramos. Además, pueden plantearse las cuestiones: a) sobre el mínimo de duración del ciclo $C = \max_i T_i$ para un número dado m de tramos; b) sobre el mínimo del número m para una longitud dada del ciclo. En la fig. 6.2 vienen introducidas 9 operaciones (entre paréntesis se indica el número de unidades de tiempo que se requieren para la ejecución de la operación correspon-

diente) y 3 tramos:

$$R = (R_1, R_2, R_3) = (1, 2, 3; 4, 5, 6; 7, 8, 9),$$

tenemos

$$F_1 = 21; F_2 = 16; F_3 = 14,$$
$$C = \max_i F_i = 21; T = \sum_i t_i = 51.$$

El tiempo total de parada es

$$d = \sum_i d_i = mC - T = 3 \cdot 21 - 51 = 12.$$

7. Problema de recubrimiento. Hállese, para un grafo no orientado dado G , el recubrimiento mínimo, es decir, un surtido mínimo de aristas tal que cualquier vértice resulte ser incidente a cierta arista perteneciente al surtido mencionado. Formalicemos el planteamiento de este problema. Numeremos los vértices del grafo con los números $1, 2, \dots, m$, y las aristas, con los números $1, 2, \dots, n$. Sea $[a_{ij}]$ la matriz de incidencias del grafo G . Unamos con las aristas las variables x_j ($j = 1, 2, \dots, n$) tales que $x_j = 1$, si la arista j integra el recubrimiento; $x_j = 0$, en el caso contrario. El problema de hallar el recubrimiento mínimo es equivalente a la búsqueda del $\min \sum_{j=1}^n x_j$ con requerimientos adicionales:

$$\sum_{j=1}^n a_{ij} x_j \geq 1; i = 1, 2, \dots, m.$$

La forma general de un problema combinatorio extremal es la siguiente. Si tiene un n -conjunto de elementos, en el que se da un conjunto (finito, por regla general) de combinaciones

$$P = (\pi_1, \pi_2, \dots, \pi_r).$$

Por combinaciones $\pi_1, \pi_2, \dots, \pi_r$ pueden entenderse las permutaciones, combinaciones, diferentes sucesiones, etc. En el conjunto P se define una función F . Se necesita hallar el extremo de F (el máximo o el mínimo), o bien los elementos del conjunto P que aseguran dicho extremo.

La propia formulación de los problemas combinatorios extremales predetermina la elección de las operaciones que se aplican para su resolución. En primer lugar, hay que disponer de un conjunto, de valores de la función F y saber seleccionarlas de un modo adecuado. En segundo lugar es necesario desarrollar el método de comparar estos valores y distinguir entre ellos el valor máximo o mínimo.

La primera operación de la selección raras veces, prácticamente, resulta ser realizable, puesto que el número de toda clase de combinaciones posibles puede ser demasiado grande. Efectivamente, en el primer ejemplo el número de todas las combinaciones es igual a $n!$, en el segundo a

$(n-1)!$, en el tercer ejemplo a 2^n , en el cuarto a $(n!)^m$, en el quinto a $\sum_{k=2}^n \binom{n}{k} (k-1)$, etc.

No es más fácil la operación de comparación. En el caso general es difícil juzgar, sin realizar cálculos inmediatos, si tiene lugar $F(\pi_i) > F(\pi_j)$, o, viceversa, $F(\pi_i) \leq F(\pi_j)$. Además, la propia determinación del valor de la función $F(\pi_i)$ representa, como regla, un problema que no es nada fácil.

Las dificultades relacionadas con la elección de las variantes y la comparación de los valores son considerables. Precisamente ellas constituyen un obstáculo para el progreso de esta parte del análisis combinatorio, a pesar de su actualidad evidente. Solamente con la introducción (hace 30 años) en la práctica matemática de los ordenadores se hizo posible la resolución de toda una serie de problemas extremales.

El esquema general que caracteriza la conexión de los problemas combinatorios extremales con los métodos de programación lineal puede ser representado aproximadamente así: los elementos π_i se interpretan como puntos de un espacio euclídeo para que la función de "especial" F se haga una forma lineal. Se examina el problema de encontrar el extremo de esta función en la cápsula convexa de los puntos dados (es decir, en un poliedro convexo). En efecto, el extremo de una forma lineal en un poliedro se consigue en uno de los vértices que integran el conjunto de los elementos en consideración. Mientras tanto, el problema de hallar el extremo de una forma lineal es precisamente un problema de programación lineal. La peculiaridad de los problemas combinatorios, al realizarse tal esquema, consiste en que en el proceso de búsqueda de la solución hemos de limitarnos a los puntos cuyas coordenadas son números enteros.

Mostremos con ejemplos cómo se plantean los problemas de programación para los problemas combinatorios extremales. La solución del problema de nombramientos (ejemplo 1) representa una permutación (p_1, p_2, \dots, p_n) de los números $1, 2, \dots, n$, obtenida como resultado de los nombramientos del tipo $i \rightarrow p_i$; $i = 1, 2, \dots, n$. El objetivo consiste en encontrar $\min \sum_{i=1}^n c_{ij}$ en el conjunto finito de las permutaciones mencionadas. Cada permutación puede interpretarse como un punto en el espacio euclídeo n^2 -dimensional; en este caso ésta puede representarse con mayor comodidad en forma de una $(n \times n)$ -matriz $X = \|x_{ij}\|$, donde $x_{ij} = 1$, si el i -ésimo candidato está designado para ocupar el j -ésimo cargo, y $x_{ij} = 0$, en el caso contrario. Sustituyamos esta condición por la otra: $x_{ij} \geq 0$, para que no se impida la aplicación de la programación. Los requerimientos referentes a la inadmisibilidad de ejercer un cargo simultáneamente con otro y al fraccionamiento de los cargos se escribirán como condiciones:

$$\sum_{j=1}^n x_{ij} = 1; \quad i = 1, 2, \dots, n;$$

$$\sum_{i=1}^n x_{ij} = 1; \quad j = 1, 2, \dots, n;$$

Los gastos sumarios, cuyo mínimo se busca, se escribirán así:

$$\sum_{i=1}^n \sum_{j=1}^n c_{ij} x_{ij}.$$

En el problema de un viajante (ejemplo 2) se busca, como sabemos,

$$\min \sum_{i=0}^n \sum_{j=0}^n c_{ij} x_{ij}$$

bajo las condiciones de que

$$\sum_{i=1}^n x_{ij} = 1; \quad j = 1, 2, \dots, n;$$

(de cada ciudad el viajante se va sólo una vez),

$$\sum_{j=1}^n x_{ij} = 1, \quad i = 1, 2, \dots, n;$$

(en cada ciudad el viajante llega una sola vez) y

$$u_i - u_j + n x_{ij} \leq n - 1; \quad i, j = 1, 2, \dots, n; \quad i \neq j.$$

La última condición está introducida con el fin de asegurar la unicidad del ciclo en el camino del viajante. Efectivamente, si existiera una solución que contenga dos o más ciclos, se encontraría un subciclo τ con k eslabones el cual no pase por el punto de partida. La sumación de las últimas condiciones para τ llevaría a una contradicción:

$$nk \leq (n - 1)k.$$

En la última condición no se han determinado u_i y u_j . Pueden hallarse para cualquier ciclo que tiene por origen el punto de partida. Si el punto i se visita en la p -ésima etapa del recorrido ($p = 1, 2, \dots, n$), entonces pongamos $u_i = p$. Para todos los i y j de aquí se deduce que $u_i - u_j \leq n - 1$. La última condición se cumple para cualesquiera $x_{ij} = 0$. En cambio, si $x_{ij} = 1$, las citadas condiciones se escriben como igualdades:

$$u_i - u_j + n x_{ij} = p - (p + 1) + n = n - 1.$$

El problema de recubrimiento (ejemplo 7) ya está formulado de tal modo que permite plantear el problema correspondiente de la programación lineal. Su planteamiento más general consiste en hallar, para un conjunto finito dado $S = \{s_1, s_2, \dots, s_n\}$ y cierta totalidad finita de sus subconjuntos S_j ($j = 1, 2, \dots, m$), un recubrimiento mínimo del conjunto S , es decir, un juego mínimo de subconjuntos S_j , en el cual todo elemento $s \in S$ pertenezca por lo menos a uno de los subconjuntos.

Para este problema general se aplica un acceso análogo: se hace una matriz de incidencias $|a_{ij}|$, en la cual $a_{ij} = 1$, si $s_i \in S_j$, y $a_{ij} = 0$, en el caso contrario. Hallemos x_{ij} que son iguales a uno, si S_j integra el recubrimiento, y a 0, si no lo integra. Por consiguiente, resulta aplicable al problema la programación lineal.

Para la resolución de los problemas combinatorios extremales han sido elaborados diferentes métodos de cálculo. Los más perspectivas de ellos han formado una rama aparte de la programación combinatoria (véase [80]). La idea general de dichos métodos consiste en la sustitución de la selección completa de todas las variantes por las selecciones parciales de menor volumen. Para poder realizar tal idea se hallan ciertos subconjuntos que a ciencia cierta no contienen el extremo buscado, estrechando el dominio de las variantes posibles. En esta ocasión los métodos resultan ser los más diversos y se definen por la estructura de los conjuntos finitos correspondientes. Más abajo se describen varios métodos ampliamente usados hoy día.

6.2. MÉTODO DE RAMIFICACIONES Y RESTRICCIONES¹⁾

De acuerdo con este método, el conjunto de todas las combinaciones admisibles Q se parte en subconjuntos Q_1, Q_2, \dots, Q_{s_1} . Cada uno de los últimos se parte, a continuación, en subconjuntos $Q_{i1}, Q_{i2}, \dots, Q_{is_2}$ ($i = 1, 2, \dots, s_1$), etc. hasta que se obtengan combinaciones separadas. Este proceso de ramificación resulta cómodo expresarlo en forma de un árbol. En cada subconjunto Q_{i_1, i_2, \dots, i_n} se define una función de preferencia $\varphi(Q_{i_1, i_2, \dots, i_n})$. Comenzando por el vértice Q , el proceso se realiza una etapa a otra, eligiéndose cada vez un subconjunto que tenga el valor mínimo de la citada función (o el valor máximo, si lo exigen las condiciones del problema). El árbol obtenido será un árbol de soluciones.

El problema consiste en la elección acertada de la función de preferencia para que se pueda obtener una buena solución. A veces, la elección de un subconjunto se realiza al azar: en este caso la probabilidad de que se elija un conjunto dado es tanto mayor cuanto menor es el valor de su función de preferencia (funciones randomizadas de preferencia).

El problema principal en este método consiste en elegir el procedimiento de determinación de la frontera inferior (o superior). No es siempre fácil obtener el valor bastante exacto de ella, pero, conseguido este valor, el número de ramas que se examinan en el árbol de soluciones se reduce en el caso general. Esto se debe a que, si el valor de la frontera inferior para tal o cual subconjunto es superior o igual (en el caso de minimización;

¹⁾ Se conoce también como método de particiones progresivas y estimaciones, método de ramas y fronteras

inferior o igual, en el caso contrario) al valor de la función minimizada de una de las soluciones ya obtenidas, entonces, la rama correspondiente del árbol de soluciones se excluye del análisis.

Sea Q un conjunto de todas las soluciones admisibles de cierto problema: $Q = \{q_1, q_2, \dots, q_n\}$; en el conjunto de dichas soluciones está dada una función f . Se pide hallar un subconjunto $Q_n \subset Q$, en el cual la función f alcance su mínimo (o su máximo).

Supongamos que conocemos cierta frontera inferior b_0 para la función f en el conjunto Q . Admitase que por medio de un procedimiento se ha logrado partir Q en un conjunto A y un complemento de éste \bar{A} , y además, precisar, después, las fronteras inferiores de f en los conjuntos A y \bar{A} , las cuales son iguales a b_1 y b_1' , respectivamente, con la particularidad de que $b_1 \geq b_0$; $b_1' \geq b_0$. Supongamos ahora que con ayuda de otros procedimientos logramos partir Q en dos partes: B y \bar{B} ; C y \bar{C} , Examinemos los conjuntos $A \cap B$; $A \cap \bar{B}$; $\bar{A} \cap B$; $\bar{A} \cap \bar{B}$ y las fronteras inferiores precisadas de f en dichos conjuntos b_2 , b_2' , b_2'' y b_2''' , respectivamente. En este caso, además,

$$\begin{aligned} b_2 &\geq b_1 \geq b_0; & b_2'' &\geq b_1' \geq b_0; \\ b_2' &\geq b_1 \geq b_0; & b_2''' &\geq b_1' \geq b_0. \end{aligned}$$

Así pues, podemos construir un árbol orientado con la raíz Q (bisárbol) (fig. 6.3), con la particularidad de que no hay necesidad de hacerlo hasta el fin. Supongamos que hemos construido una parte del árbol partiendo el conjunto Q en ciertos subconjuntos y encontramos las fronteras inferiores para los vértices correspondientes a los subconjuntos citados. Elegimos de todos los vértices pendientes aquél que tiene una frontera mínima, y, a continuación, partiendo los subconjuntos correspondientes, obtenemos dos vértices nuevos, en los cuales precisamos las fronteras inferiores de la función f .

Enunciemos la siguiente afirmación sencilla: en cualquier etapa la unión de los subconjuntos, correspondientes a los vértices pendientes, da todo el Q . Por eso, si obtenemos, como resultado del proceso dado, un vértice compuesto por un solo conjunto $\{q_i\}$, y si la frontera inferior de la función f en el citado vértice es menor que en los restantes vértices pendientes entonces f toma en q_i su valor mínimo. El proceso se da por terminado.

Con ayuda de los razonamientos análogos se puede describir el algoritmo para hallar la solución máxima, si ésta existe.

Desgraciadamente, el método en consideración no da una respuesta constructiva al problema planteado. En cada problema concreto hay que idear algoritmos de partición del conjunto Q y los mejores procedimientos para precisar las fronteras inferiores b en los vértices pendientes.

Veamos cómo se aplica el método al problema del viajante. En el lenguaje de la teoría de los grafos en este problema se trata de la búsqueda de los contornos hamiltonianos óptimos.

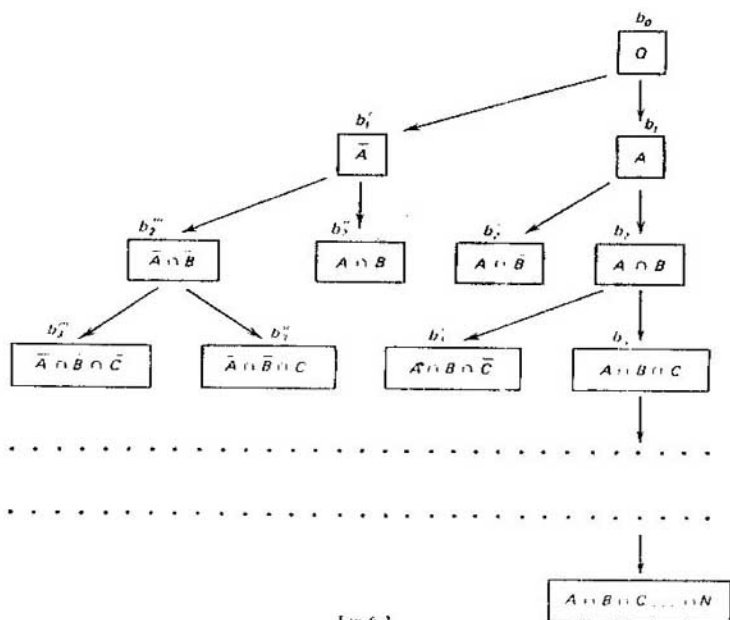


Fig 6.3.

Sin embargo, no todos los grafos contienen el contorno de Hamilton. Por consiguiente, antes de proceder a buscar el contorno hamiltoniano óptimo, hemos de tratar, por lo menos, de establecer si tal contorno existe en el grafo dado. Por otra parte, existen grafos que disponen de una gran cantidad de contornos hamiltonianos, lo que hace imposible la selección completa incluso en los ordenadores. Por ejemplo, para un grafo completo en n vértices existen $(n - 1)!$ diferentes contornos hamiltonianos. En 1963 Little encontró un método estricto de optimización para los problemas del tipo dado con gran número de contornos hamiltonianos.

Sea dado un grafo orientado $G = (X, U)$, donde $|X| = n$, y a todo arco $(x_i, x_j) \in U$ le está asignado el valor $c_{ij} = c(x_i, x_j)$. Vamos a suponer que cualesquiera $c_{ij} \geq 0$. Si es que el arco $(x_i, x_j) \notin U$, entonces $c_{ij} = \infty$.

El algoritmo para encontrar el contorno hamiltoniano mínimo consistirá en la construcción de un bisárbol descrito en el método de ramificaciones y restricciones.

A) Q (la raíz del bisárbol) es el conjunto de todos los contornos hamiltonianos. La frontera inferior de E , es, obviamente, un cero, pero encontraremos la frontera mayor en el proceso de búsqueda de los conjuntos

A y \bar{A} del bisárbol. Procedamos de la manera siguiente. En cada fila de la matriz $[c_{ij}]$ encontramos el elemento mínimo y lo restamos de todos los elementos de dicha fila. Si en la matriz obtenida hay columnas que no tienen elementos nulos, encontramos en cada una de ellas el elemento mínimo y lo restamos de todos los elementos de la columna. De resultas, tendremos una matriz nueva $[c'_{ij}]$ en la que en cada fila y en cada columna se contiene por lo menos un solo cero.

B) Sumemos todos los elementos que hemos sustraído en A). La suma de ellos b será precisamente la frontera inferior para Q . Efectivamente, tomemos un contorno arbitrario L , en el cual la suma de los valores de los arcos es igual a l . En el punto A), al transformar la matriz $[c_{ij}]$ en $[c'_{ij}]$, hemos sustraído ciertos números de todos los elementos de las filas correspondientes y de ciertas columnas. Con ello, cada vez de la magnitud l estos números se restaban también y, al fin y al cabo, la longitud del contorno L en la matriz $[c'_{ij}]$ es también no negativa. Por consiguiente, $l \geq b$.

C) Pasemos a la construcción de los vértices del bisárbol del siguiente nivel. Sea $c'_{ij} = 0$. Definamos γ_{ij} como suma del elemento mínimo de la i -ésima fila y del elemento mínimo de la j -ésima columna (excluyendo c_{ij} que es igual a cero) y hallemos

$$\theta = \gamma_{kl} - \max_{i, j} \gamma_{ij}$$

Veamos la propiedad p_{ij} : "el contorno no contiene el arco (x_i, x_j) " que se aplicará a los arcos con $c'_{ij} = 0$.

Si un contorno hamiltoniano no contiene el arco (x_i, x_j) , entonces obligatoriamente empleará ciertos dos arcos (x_i, x_s) ; $s \neq j$, y (x_r, x_j) ; $r \neq i$.

D) Demostremos que $b'_i = b + \gamma_{kl}$ es la frontera inferior para el conjunto de contornos que satisfacen la propiedad p_{kl} . En efecto, analicemos un contorno hamiltoniano arbitrario L_1 privado del arco (x_k, x_l) . Entonces, realizada la transformación A), la longitud de este contorno en la matriz renovada será no inferior a cero. Con eso queda construido el vértice del bisárbol E_{kl} con la frontera inferior $b' = b + \theta$.

E) Construimos el vértice E_{kl} que se define por la propiedad p_{kl} : "el contorno emplea el arco (x_k, x_l) ". Suprimimos de la matriz la k -ésima fila y la l -ésima columna, sustituyéndolas por ∞ :

A) o bien el valor de c'_{ik} , si el contorno no empleaba hasta este momento los arcos que entran en el vértice x_k y salen del vértice x_i , puesto que la conexión del arco (x_i, x_k) con los arcos del contorno ya elegidos conduce a la aparición de un contorno no hamiltoniano; b) o bien el valor de c'_{lm} , si el contorno contiene, además del arco (x_k, x_l) , todos los arcos de cierto camino simple del vértice x_m al vértice x_k , puesto que la adición del arco (x_l, x_m) a los arcos del contorno ya elegidos también conduce a la aparición de un contorno no hamiltoniano; c) o bien el valor de c'_{pk} , si el contorno contiene, además del arco (x_k, x_l) , todos los arcos del camino simple que

	x_1	x_2	x_3	x_4	x_5
x_1	∞	2	3	2	7
x_2	3	∞	7	6	7
x_3	1	7	∞	3	7
x_4	6	7	3	∞	7
x_5	7	6	7	6	∞

Fig. 6.4.

llevan del vértice x_i al vértice x_j , puesto que la conexión del arco (x_i, x_j) con los arcos del contorno ya elegidos conduce a la aparición de un contorno no hamiltoniano.

F) Actuamos igual que en A) con la matriz obtenida como resultado de E).

G) Actuamos igual que en B) con la matriz obtenida como resultado de F). Al agregar la suma obtenida a la frontera para el vértice antecedente (en el primer caso, para Q), obtenemos la frontera para el vértice de E).

H) Si, como resultado de F), se obtiene una matriz de orden 1, el proceso se da por terminado. Si no, pasamos a I).

I) Entre todos los vértices pendientes del bisárbol ya construido elegimos un vértice de frontera mínima.

J) Si el vértice elegido en I) correspondía en el proceso de construcción a la propiedad p_{ij} , pasamos a C). Si no, pasamos a K).

K) Supongamos que el vértice elegido correspondía en la construcción a la propiedad p_{ij} . Sustituimos el valor en la célula (i, j) de la matriz correspondiente por ∞ . En la i -ésima fila, igual que en la j -ésima columna, encontramos un elemento mínimo y lo restamos de todos los elementos de dicha fila (columna). Pasamos luego a C).

Examinemos ahora un ejemplo numérico que muestra cómo se aplica el algoritmo de Little a la búsqueda del contorno hamiltoniano mínimo de un grafo. Sea G un grafo con 5 vértices x_1, x_2, x_3, x_4, x_5 , y a cada arco de este grafo (x_i, x_j) se le asigna un número c_{ij} : $i, j = 1, 2, \dots, 5$. Estos valores están escritos en forma de una matriz de valores en la fig. 6.4.

Construyamos un bisárbol, descrito por el método de ramificaciones y restricciones, para este ejemplo numérico.

A) De los elementos contenidos en las filas x_1, x_2, x_3, x_4, x_5 sustraemos respectivamente sus elementos mínimos 2, 3, 1, 2, 4; además, de la columna x_5 sustraemos 1 (véase fig. 6.6 a).

B) Calculamos la suma $2 + 3 + 1 + 2 + 4 + 1 = 13$. De este modo, para el vértice Q (el conjunto de todos los contornos hamiltonianos) obtenemos una frontera: $b = 13$.

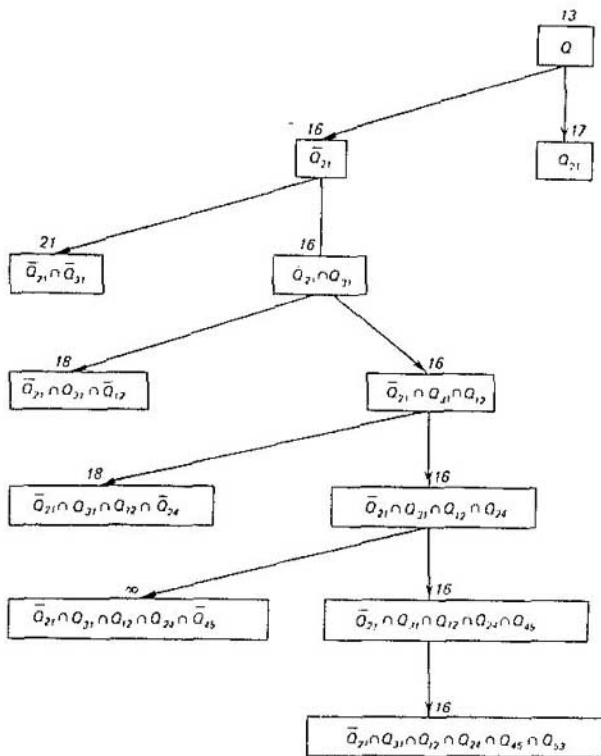


Fig. 6.5.

C) Pasamos al examen de todos los elementos nulos de la matriz:

$$\begin{aligned} \gamma(x_1, x_2) = 2; \quad \gamma(x_1, x_4) = 2; \quad \gamma(x_2, x_1) = 3; \quad \gamma(x_3, x_1) = 2; \\ \gamma(x_4, x_3) = 0; \quad \gamma(x_4, x_5) = 3; \quad \gamma(x_5, x_3) = 2. \end{aligned}$$

Por consiguiente, $\theta = \gamma(x_2, x_1) = 3$.

D) Expresamos el vértice \bar{Q}_{21} del bisárbol con la frontera inferior igual a $13 + 3 = 16$.

E) Expresamos el vértice Q_{21} definido por la propiedad "el contorno contiene el arco (x_2, x_1) ". Suprimimos la segunda fila y la primera columna. En la célula (1, 2) colocamos ∞ (véase fig. 6.6 b).

F) Al principio sustraemos 2 de todos los elementos de la fila x_3 , y después, de la primera columna sustraemos 2 (véase fig. 6.6 c).

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	0	∞	4	3	3
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	2	0	2	∞

	x_2	x_3	x_4	x_5
x_1	∞	1	0	4
x_3	2	∞	0	3
x_4	3	0	∞	0
x_5	0	0	2	∞

	x_1	x_3	x_4	x_5
x_1	∞	1	0	4
x_4	6	∞	2	5
x_5	5	0	∞	0
x_3	2	0	2	∞

Fig. 6.6

G) La suma de los elementos sustraídos de las filas y columnas en F), es igual a 4. Por consiguiente, para el vértice \bar{Q}_{21} tenemos una frontera 17 (=13 + 4).

H) Se ha obtenido aquí la matriz de orden 4. Pasamos a la siguiente operación I).

I) El vértice pendiente con el valor mínimo igual a 16 es \bar{Q}_{21} .

J) El vértice \bar{Q}_{21} se ha obtenido con ayuda de la propiedad p_{21} . Pasamos a K).

K) En la célula (x_2, x_1) de la matriz expuesta en la fig. 6.6 a) disponemos ∞ . Obtenemos una matriz que se muestra en la fig. 6.7 a). A continuación sustraemos 3 de todos los elementos de la fila x_2 . Véanse en la fig. 6.7 b) los resultados.

C) Para la matriz expuesta en la fig. 6.7 b) calculamos:

$$\gamma(x_1, x_2) = 2, \quad \gamma(x_1, x_4) = 0, \quad \gamma(x_2, x_4) = 0, \quad \gamma(x_2, x_5) = 0,$$

$$\gamma(x_3, x_1) = 5, \quad \gamma(x_4, x_3) = 0, \quad \gamma(x_4, x_5) = 0, \quad \gamma(x_5, x_3) = 2.$$

$$\text{Obtenemos } \Theta = \gamma(x_3, x_1) = 5.$$

D) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31}$, al que corresponde la frontera: 16 + 5 = 21;

E) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31}$. Suprimimos la fila x_3 y la columna x_1 . En la célula (x_1, x_3) de la matriz expuesta en la fig. 6.7 b) colocamos ∞ . Véase el resultado en la fig. 6.7 c).

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	∞	∞	4	3	1
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	4	0	2	∞

	x_1	x_2	x_3	x_4	x_5
x_1	∞	0	1	0	4
x_2	∞	∞	1	0	0
x_3	0	6	∞	2	5
x_4	3	5	0	∞	0
x_5	3	2	0	2	∞

	x_2	x_3	x_4	x_5
x_1	0	∞	0	4
x_2	∞	1	0	0
x_3	5	0	∞	0
x_4	2	0	2	∞

Fig.6.7

F) Todas las filas y todas las columnas de la matriz en la fig. 6.7 c contienen ceros.

G) Por consiguiente, para el vértice $Q_{21} \cap Q_{31}$ la frontera es igual a 16.

H) Se ha obtenido la matriz de orden 4. Pasamos a D).

I) El vértice pendiente de valor mínimo igual a 16 es $Q_{21} \cap Q_{31}$.

J) El vértice $Q_{21} \cap Q_{31}$ se ha obtenido con ayuda de la propiedad p_{31} . Pasamos a C).

C) Calculamos γ para todos los elementos nulos de la matriz expuesta en la fig. 6.7 c:

$$\begin{aligned} \gamma(x_1, x_2) &= 2; \quad \gamma(x_1, x_4) = 0; \quad \gamma(x_2, x_4) = 0; \\ \gamma(x_2, x_5) &= 0; \quad \gamma(x_4, x_1) = 0; \quad \gamma(x_4, x_5) = 0; \quad \gamma(x_5, x_3) = 2; \\ \theta &= \gamma(x_1, x_2) = 2. \end{aligned}$$

D) Construimos el vértice $Q_{21} \cap Q_{11} \cap \bar{Q}_{12}$ cuya frontera es 18 (=16 + 2).

F) Construimos el vértice $Q_{21} \cap Q_{11} \cap Q_{12}$. Suprimimos la fila x_1 y la columna x_2 de la matriz expuesta en la fig. 6.7 c. Obtenemos una matriz que se muestra en la fig. 6.8 a. En la célula (2, 3) colocamos el signo ∞ .

F) En cada fila y en cada columna de la matriz expuesta en la fig. 6.8 a se contienen ceros.

G) La frontera inferior para el vértice $\bar{Q}_{21} \cap Q_{31} \cap Q_{12}$ es igual a 26.

H) Se ha obtenido una matriz de orden 3. Pasamos a I).

	x_1	x_2	x_3
x_1	∞	0	0
x_2	0	∞	0
x_3	0	2	∞

	x_1	x_2
x_4	∞	0
x_5	0	∞

	x_1
x_5	0

Fig. 6.8.

- I) El vértice pendiente de valor mínimo 16 es $Q_{21} \cap Q_{31} \cap Q_{12}$.
 J) El vértice $Q_{21} \cap Q_{11} \cap Q_{13}$ se ha obtenido con ayuda de la propiedad p_{12} . Por eso, pasamos a C).
 C) Calculamos γ para todos los elementos nulos de la matriz en la fig. 6.8 a:

$$\begin{aligned} \gamma(x_2, x_1) &= 2; \quad \gamma(x_2, x_3) = 0; \quad \gamma(x_4, x_1) = 0; \\ \gamma(x_2, x_5) &= 0; \quad \gamma(x_5, x_3) = 2; \\ \Omega &= \gamma(x_2, x_1) = 2. \end{aligned}$$

D) Construimos el vértice $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$, cuya frontera será igual a 18 ($= 16 + 2$).

E) Construimos el vértice $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$. Suprimimos la fila x_2 y la columna x_3 de la matriz en la fig. 6.8 a; en la célula (4, 3) colocamos ∞ . Obtenemos una matriz que se muestra en la fig. 6.8 b).

F) En la matriz de la fig. 6.8 cada fila y cada columna contiene ceros.

G) La frontera para el vértice $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$ será igual a 16.

H) Tenemos una matriz de orden 2, por eso pasamos a I).

I) El vértice pendiente de valor mínimo 16 es $Q_{21} \cap Q_{31} \cap Q_{12} \cap Q_{24}$.

J) Este vértice se ha obtenido con ayuda de la propiedad p_{24} . Pasamos a C).

C) Calculamos γ para los elementos nulos de la matriz en la fig. 6.8 b:

$$\begin{aligned} \gamma(x_4, x_1) &= \infty; \quad \gamma(x_5, x_1) = \infty; \\ \Omega &= \gamma(x_4, x_5) = \infty. \end{aligned}$$

D) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31} \cap \bar{Q}_{12} \cap \bar{Q}_{45}$. La frontera para el es igual a ∞ .

E) Construimos el vértice $\bar{Q}_{21} \cap \bar{Q}_{31} \cap \bar{Q}_{12} \cap \bar{Q}_{24} \cap \bar{Q}_{45}$. Suprimimos la fila x_4 y la columna x_5 . Obtenemos una matriz que se muestra en la fig. 6.8 c).

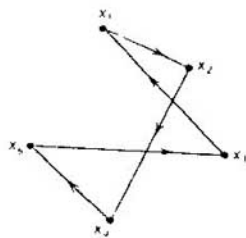
F) Es obvio que en la matriz obtenida no se requiere ninguna sustracción.

G) La frontera inferior para el vértice en E) es igual a $16 + 0 = 16$.

H) Tenemos la matriz de orden 1 con frontera mínima entre todos los vértices pendientes. Al agregar el arco (x_5, x_3) , obtenemos el contorno hamiltoniano buscado de valor 16: $(x_3, x_1, x_2, x_4, x_5)$. La solución hallada se representa en las figs. 6.9 a y 6.9 b.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	2	3	2	7
x_2	3	∞	7	6	/
x_3	1	7	8	3	7
x_4	5	7	2	∞	3
x_5	7	6	4	6	∞

a



b

Fig. 6.9

El algoritmo obtenido permitió hallar la solución del problema planteado. Sin embargo, pueden existir otros contornos de Hamilton con el mismo valor. Esto resulta de que la elección de los vértices en algunos casos no era unívoca.

He aquí un ejemplo más que ilustra la aplicación del algoritmo de Little para la búsqueda en un grafo de 5 vértices del contorno hamiltoniano mínimo, si la matriz de valores está representada en la fig. 6.10 a. Construyamos el bisárbol.

A) De los elementos contenidos en las filas x_1, x_2, x_3, x_4, x_5 sustraemos respectivamente sus elementos mínimos: 3, 0, 5, 3, 16, y, a continuación, de la matriz obtenida (véase fig. 6.10 b) restamos de las columnas x_1, x_2, x_3 1, 2, 1, respectivamente (véase fig. 6.10 c).

B) Calculamos la suma $3 + 0 + 5 + 3 + 16 + 1 + 2 + 1 = 31$. Por consiguiente, para el vértice Q obtenemos la frontera inferior: $b = 31$.

C) Pasamos al examen de todos los elementos nulos de la matriz:

$$\gamma(x_1, x_3) = 17; \gamma(x_2, x_5) = 1; \gamma(x_3, x_4) = 13; \gamma(x_4, x_5) = 16;$$

$$\gamma(x_5, x_1) = 7; \gamma(x_5, x_2) = 13; \gamma(x_5, x_3) = 16; \gamma(x_5, x_4) = 1.$$

Así pues, $\Theta = \gamma(x_1, x_3) = 17$.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	31	47	20	3
x_2	8	∞	17	1	0
x_3	26	20	∞	40	5
x_4	25	21	29	∞	3
x_5	17	18	17	16	∞

a

b

c

Fig. 6.10.

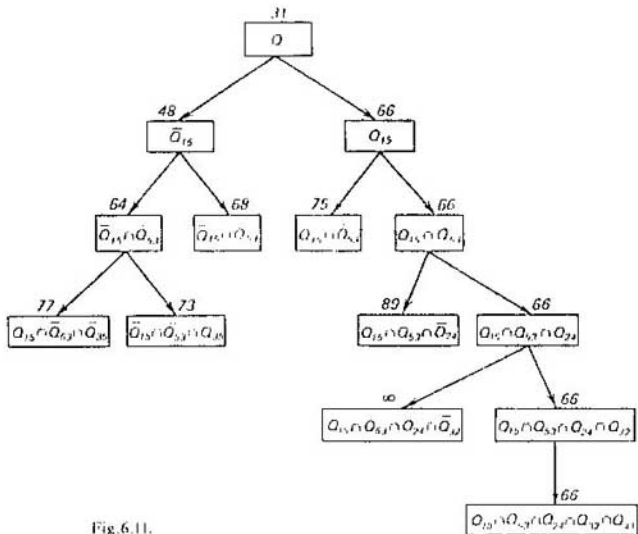


Fig. 6.11.

D) Expresamos el vértice Q_{15} del bisímbol cuya frontera inferior es igual a $31 + 17 = 48$.

E) Expresamos el vértice Q_{15} que se define por la propiedad p_{15} "el contorno hamiltoniano contiene el arco (x_1, x_3) ". Suprimimos en la matriz expuesta en la fig. 6.10 c la primera fila y la quinta columna, y en la célula $(5, 1)$ ponemos ∞ (véase fig. 6.12 a).

F) Sustraigamos de todos los elementos de las filas x_2, x_3, x_4 de la matriz obtenida, 1, 3, 16, respectivamente, y luego, de la primera columna sustraigamos además 5 (véase fig. 6.12 b).

G) La suma de los elementos sustraídos de las filas y columnas de la matriz en el punto anterior es igual a 35. Por consiguiente, para el vértice Q_{15} la frontera inferior es igual a $31 + 35 = 66$.

H) En la última etapa se ha obtenido una matriz de cuarto orden. Por eso pasamos a la operación siguiente.

	x_1	x_2	x_3	x_4
x_1	7	∞	16	1
x_2	20	13	∞	35
x_3	21	16	25	∞
x_4	∞	0	0	0

a

	x_1	x_2	x_3	x_4
x_1	1	∞	15	0
x_2	2	0	∞	22
x_3	0	0	9	∞
x_4	∞	0	0	0

b

Fig. 6.12.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	20	13	17	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	0	0	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	26	0	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	0	0	∞

b

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	0	∞	
x_2	7	∞	1	0	
x_3	20	13	35	∞	
x_4	21	16	∞	0	

c

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	0	∞	
x_2	0	∞	1	0	
x_3	0	0	22	∞	
x_4	14	16	∞	0	

d

Fig. 6.13

1) El vértice pendiente de frontera mínima 48 es Q_{15} .

1) El vértice se ha obtenido con ayuda de la propiedad p_{15} . Pasamos a K).

K) En la célula (1, 5) de la matriz expuesta en la fig. 6.10 e colocamos ∞ . Obtenemos una matriz que se expone en la fig. 6.13 a. A continuación sustraemos 17 de todos los elementos de la fila x_1 . El resultado está representado en la fig. 6.13 b.

C) Para la matriz expuesta en la fig. 13 b calculamos:

$$\begin{aligned} \gamma(x_1, x_4) &= 9; \quad \gamma(x_2, x_3) = 1; \quad \gamma(x_3, x_5) = 13; \quad \gamma(x_4, x_5) = 16; \\ \gamma(x_5, x_1) &= 7; \quad \gamma(x_5, x_2) = 9; \quad \gamma(x_5, x_3) = 16; \quad \gamma(x_5, x_4) = 0, \end{aligned}$$

por lo cual tenemos, por ejemplo, $\theta = \gamma(x_5, x_1) = 16$.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	26	0	∞
x_2	7	∞	16	1	0
x_3	20	13	∞	35	0
x_4	21	16	25	∞	0
x_5	0	0	∞	0	∞

a

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	∞
x_2	7	∞	0	1	0
x_3	20	13	∞	35	0
x_4	21	16	9	∞	0
x_5	0	0	∞	0	∞

b

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	
x_2	7	∞	0	1	
x_3	12	7	0	∞	
x_4	0	0	∞	0	

c

	x_1	x_2	x_3	x_4	x_5
x_1	∞	9	10	0	
x_2	7	∞	0	1	
x_3	12	7	0	∞	
x_4	0	0	∞	0	

d

Fig. 6.14

D) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$, al cual corresponde la frontera inferior $48 + 16 = 64$.

E) Construimos el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$. Suprimimos la quinta fila y la tercera columna en la matriz expuesta en la fig. 6.13 b, y en la célula (3, 5) colocamos ∞ . El resultado se muestra en la fig. 6.13 c).

F) De todos los elementos de la fila x_3 sustraemos primero 13, y luego de la primera columna sustraemos 7 (véase fig. 6.13 d)

G) La suma de los elementos sustraídos de las filas y columnas en F) es igual a $13 + 7 = 20$. Por consiguiente, para el vértice $\overline{Q}_{15} \cap \overline{Q}_{53}$ tenemos la frontera igual a $48 + 20 = 68$.

H) Pasamos a I).

I) El vértice pendiente $\overline{Q}_{15} \cap \overline{Q}_{53}$ tiene el valor mínimo de la frontera inferior.

J) El vértice elegido correspondía, al construir el bisárbol, a la propiedad \overline{p}_{53} . El valor dentro de la célula (5, 3) en la matriz expuesta en la fig. 6.13 b se sustituye por ∞ , y de todos los elementos de la columna x_3 se resta el número 16. El resultado se muestra en la fig. 6.14 b). Pasamos a C).

C) Para todos los elementos nulos de la matriz en la fig. 6.14 b calculamos:

$$\gamma(x_1, x_4) = 9; \gamma(x_2, x_3) = 9; \gamma(x_2, x_5) = 0; \gamma(x_3, x_5) = 13;$$

$$\gamma(x_4, x_5) = 9; \gamma(x_5, x_1) = 7; \gamma(x_5, x_2) = 9; \gamma(x_5, x_4) = 0.$$

De aquí, $\Theta = \gamma(x_3, x_5) = 13$.

D) Construimos el vértice $Q_{15} \cap \overline{Q}_{51} \cap \overline{Q}_{35}$, al cual corresponde la frontera inferior $64 + 13 = 77$.

E) Para construir el vértice $Q_{15} \cap Q_{53} \cap Q_{35}$ suprimimos la fila x_3 y la columna x_5 . Observemos que aquí no hay necesidad de colocar ∞ en la célula (5, 3), puesto que ella ya se encuentra allí. El resultado se muestra en la fig. 6.14 c).

F) De todos los elementos de la fila x_4 sustraemos el número 9 y obtendremos una matriz expuesta en la fig. 6.14 d).

G) Para el vértice $\overline{Q}_{15} \cap Q_{53} \cap \overline{Q}_{35}$ tenemos una frontera inferior igual a $64 + 9 = 73$.

H) Se ha obtenido la matriz de orden 4. Por eso pasamos a I).

I) Entre los vértices pendientes del bisárbol ya construido el Q_{15} tiene frontera inferior mínima.

J) Por cuanto el vértice elegido Q_{15} correspondía durante la construcción a la propiedad p_{15} , pasamos a C).

C) Para la matriz expuesta en la fig. 6.12 b calculamos:

$$\gamma(x_5, x_4) = 0; \gamma(x_2, x_4) = 1; \gamma(x_3, x_2) = 2; \gamma(x_4, x_1) = 1;$$

$$\gamma(x_4, x_2) = 0; \gamma(x_5, x_2) = 0; \gamma(x_5, x_3) = 9. \text{ De aquí, } \Theta = \gamma(x_5, x_3) = 9.$$

D) Construimos el vértice $Q_{15} \cap \overline{Q}_{53}$, al cual corresponde la frontera inferior $66 + 9 = 75$.

	x_1	x_2	x_3
x_1	1	∞	0
x_2	∞	0	22
x_3	0	0	∞

a

	x_1	x_2
x_1	∞	0
x_2	0	∞

b

	x_1
x_1	0

c

Fig. 6.15

E) Construimos el vértice $Q_{15} \cap Q_{53}$. Con este fin suprimimos la fila x_3 y la columna x_3 en la matriz expuesta en la fig. 6.12 b, y colocamos ∞ en la célula (3, 1), puesto que la inclusión del arco (x_3, x_1) conduce a la aparición del contorno x_1, x_5, x_3, x_1 , que no es hamiltoniano. El resultado se presenta en la fig. 6.15 a.

F) Todas las filas y columnas de la matriz expuesta en la fig. 6.15 a contienen ceros.

G) Por consiguiente, para el vértice $Q_{15} \cap Q_{53}$ tenemos una frontera inferior igual a 66.

H) Se ha obtenido la matriz de orden 3. Pasamos a I).

I) El vértice pendiente de frontera inferior mínima es $Q_{15} \cap Q_{53}$.

J) El vértice $Q_{15} \cap Q_{53}$ se ha obtenido con ayuda de la propiedad p_{53} . Pasamos a C).

C) Calculamos γ para todos los elementos nulos de la matriz expuesta en la fig. 6.15 a: $\gamma(x_2, x_4) = 23$; $\gamma(x_1, x_2) = 22$; $\gamma(x_3, x_1) = 1$; $\gamma(x_4, x_2) = 0$. De aquí $\Theta = \gamma(x_2, x_4) = 23$.

D) Construimos el vértice $Q_{15} \cap Q_{53} \cap \bar{Q}_{24}$, cuya frontera inferior es igual a $66 + 23 = 89$.

E) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24}$. Suprimimos en la matriz expuesta en la fig. 6.15 a la fila x_2 y la columna x_4 , sustituyendo el valor en la célula (4, 2) por ∞ . Obtenemos una matriz que se expone en la fig. 6.15 b.

F) Cada fila y cada columna de la matriz obtenida contiene ceros.

G) La frontera inferior para el vértice $Q_{15} \cap Q_{53} \cap Q_{24}$ es igual a 66.

H) Se ha obtenido la matriz de orden 2. Pasamos a I).

I) La frontera inferior mínima la tiene el vértice pendiente $Q_{15} \cap Q_{53} \cap Q_{24}$.

J) El vértice $Q_{15} \cap Q_{53} \cap Q_{24}$ se ha obtenido con ayuda de la propiedad p_{24} . Por eso, pasamos a C).

C) Calculamos: $\gamma(x_3, x_2) = \infty$; $\gamma(x_4, x_1) = \infty$. Quiere decir, $\Theta = \gamma(x_3, x_2) = \infty$.

D) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap \bar{Q}_{32}$. La frontera para dicho vértice es igual al infinito (∞).

E) Construimos el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap Q_{32}$. Suprimimos la fila x_3 y la columna x_2 . Obtenemos la matriz expuesta en la fig. 6.15 c.

	x_1	x_2	x_3	x_4	x_5
x_1	∞	31	47	20	3
x_2	8	∞	17	1	0
x_3	26	20	∞	40	5
x_4	25	21	29	∞	3
x_5	17	18	17	16	∞

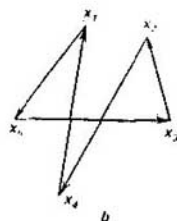


Fig. 6.16.

F) Es obvio que en la matriz obtenida no hay necesidad de sustraer nada.

G) La frontera inferior para el vértice en E) es igual a $66 + 0 = 66$.

H) La matriz obtenida es de orden 1, y el vértice $Q_{15} \cap Q_{53} \cap Q_{24} \cap Q_{32}$ tiene una frontera inferior mínima entre todos los vértices pendientes del bisárbol. Agregando el arco (x_4, x_1) , obtenemos el contorno hamiltoniano buscado de longitud 66: $(x_1, x_5, x_3, x_2, x_4, x_1)$. La solución hallada está representada en la fig. 6.16 a y b.

No es difícil modificar el algoritmo de Little para resolver los problemas referentes a la búsqueda del contorno hamiltoniano de valor máximo. No obstante, al cambiar un tanto las condiciones, reduzcamos el problema al tipo ya analizado.

Agreguemos el símbolo $(-\infty)$ a todos los pares de vértices (x_i, x_j) , para los cuales en el grafo $G = (X, U)$ no hay arco (x_i, x_j) . Sea $[c_{ij}]$ una matriz de valores del grafo $G = (X, U)$, $|X| = n$. Construyamos una matriz nueva cuyos elementos los hallamos por la fórmula

$$c'_{ij} = C - c_{ij} \quad (i, j = 1, 2, \dots, n), \text{ donde } C = \max_{i,j} c_{ij}.$$

Apliquemos ahora a la matriz $[c'_{ij}]$ el algoritmo de Little para determinar los contornos hamiltonianos mínimos. La solución mínima obtenida nos ofrece la solución máxima para la matriz $[c_{ij}]$.

6.3. MÉTODOS HEURÍSTICOS

Los métodos exactos de optimización mencionados más arriba requieren, por regla general, unos cálculos muy voluminosos. Por eso, tiene gran importancia práctica la creación de métodos sencillos que aseguran la obtención de soluciones, que son bastante próximas a las óptimas. Los métodos que no prevén la estimación de la proximidad de las soluciones obtenidas a las óptimas suelen llamarse *heurísticos*. Un método se llama *aproximado*, si admite la posibilidad de estimar la desviación de una solución de la óptima. Naturalmente, el método heurístico puede pasar al orden de aproximado tras un exitoso análisis teórico.

En los métodos heurísticos con mayor frecuencia se emplea la llamada optimización local.

Sea un conjunto de combinaciones $P = \{\pi_1, \pi_2, \dots, \pi_N\}$; se plantea el problema de calcular el mínimo de la función F , definida en este conjunto, y buscar las combinaciones sobre las cuales dicho mínimo se alcanza. El método de optimización local consiste en lo siguiente.

Para cada combinación $\pi_i \in P$ definamos un conjunto Q_i de combinaciones que se llamarán vecinas de π_i . En el lenguaje de la teoría de los grafos esto significará lo siguiente: a cada combinación π_i se le hace corresponder cierto vértice del grafo G . Los vértices vecinos π_i y π_j del grafo G se unen con arco (i, j) ($\pi_j \in Q_i$). El grafo obtenido se denominará grafo de vecindad de las combinaciones. En el grafo G puede haber una totalidad Z de sus vértices, en la cual

$$F(\pi_i) = F(\pi_j), \text{ siempre que } \pi_i, \pi_j \in Z;$$

$$F(\pi_i) < F(\pi_j), \text{ siempre que } \pi_i \in Z; \pi_j \notin Z; (i, j) \in U.$$

donde U es el conjunto de arcos del grafo de vecindad. Tal conjunto Z se llama aislado. Este conjunto puede estar compuesto, por supuesto, de un solo vértice.

El primer paso del acceso local, del cual se trata aquí, consiste en que, al elegir una combinación arbitraria π_i , hallar para ella el grafo de vecindad, después de lo cual determinar en este grafo los valores de $F(\pi_j)$ para cualquier $\pi_j \in Q_i$.

La segunda etapa consiste en una operación a la que se acostumbra llamar descenso. Encontramos $\pi_{j_0} \in Q_i$ de tal manera que

$$F(\pi_{j_0}) = \min_{\pi_j \in Q_i} F(\pi_j),$$

y si $F(\pi_{j_0}) < F(\pi_i)$, pasamos a la combinación π_{j_0} .

Procediendo de esta manera, llegaremos a un conjunto aislado tras un número finito de pasos. Sin embargo, ya encontrándonos en un conjunto aislado, no podemos estar seguros de que se ha alcanzado precisamente un extremo local (mínimo, en el caso dado), y no el general. Por eso tenemos que continuar las "pruebas" locales. Uno de los métodos de continuación de la búsqueda puede ser la elección de una nueva combinación admisible y la repetición de la operación de descenso en el antiguo grafo de vecindad. Son posibles también otros métodos. Definamos, por ejemplo, una sucesión de grafos de vecindad G_1, G_2, \dots, G_s (grafos de 1-, 2-, ..., s -vecindad). Después de caer en G_1 sobre un conjunto aislado de vértices Z , pasamos al grafo G_2 partiendo de un vértice arbitrario $\pi_i \in Z$. Si Z queda aislado también en G_2 , pasamos a G_3 , etc. Si resulta posible el descenso del vértice π_i al conjunto aislado Z_1 , entonces, al pasar a algún vértice $\pi_j \in Z_1$, volveremos al grafo G_1 y repetiremos el proceso, partiendo esta vez del vértice π_j .

En el procedimiento descrito hay muchas cosas indeterminadas. Sobre

todo no está claro cómo se deben construir los grafos de vecindad y por qué dicha construcción es mejor que la elección aleatoria de las combinaciones. La experiencia muestra, no obstante, que la obtención de las combinaciones vecinas es, a veces, más fácil que construir las sucesiones aleatorias. El análisis de un problema concreto lleva frecuentemente al método de determinación de las combinaciones vecinas que asegura el descenso eficaz hacia una solución óptima o, en todo caso, bastante buena.

Veamos algunos ejemplos de determinación de la vecindad.

1. Si las combinaciones son simplemente permutaciones, entonces las combinaciones vecinas se definen como obtenidas por conmutación de los elementos próximos. Llamemos combinaciones 1-vecinas aquellas que se obtienen por conmutación de dos elementos vecinos; combinaciones 2-vecinas, aquellas que se obtienen por conmutación de dos elementos que ocupan los lugares pares o impares, etc. Por ejemplo, una permutación $\pi_1 = (1, 2, 3, 4)$ tiene cuatro combinaciones 1-vecinas: $\pi_2 = (2, 1, 3, 4)$, $\pi_3 = (1, 3, 2, 4)$, $\pi_4 = (1, 2, 4, 3)$ y $\pi_5 = (4, 2, 3, 1)$ (los elementos 1 y 4 también se consideran vecinos) y dos combinaciones 2-vecinas: $\pi_6 = (3, 2, 1, 4)$, $\pi_7 = (1, 4, 3, 2)$. Con tal definición las combinaciones 3-vecinas y 1-vecinas coinciden.

En el caso general, las combinaciones 2-vecinas se definen del modo siguiente: π_i es una combinación 2-vecina para π_j , si existe π_k tal que π_i es una combinación 1-vecina para π_k , y π_k es una combinación 1-vecina para π_j . Análogamente se introducen las combinaciones 3-vecinas, etc.

2. En los problemas 1 y 6 del § 6.1 se ha estudiado la cuestión sobre el nombramiento óptimo y sobre la parada mínima de una línea de montaje. En dichos problemas el valor de F estaba definido como $\max F_j$ para cada tramo del transportador. Los tramos, para los cuales $F_i = \max F_j$ se llamarán críticos. En tal caso resulta natural definir las combinaciones vecinas del modo siguiente: la combinación π_i es 1-vecina para π_j si puede obtenerse de π_j por conmutación de dos operaciones, una de las cuales está designada para el tramo crítico y la otra, no.

Veamos, por ejemplo, el esquema del problema 6 (fig. 6.2) una vez más. Para la solución $R^0 = (1, 2, 3; 4, 5, 6; 7, 8, 9)$ obtendremos: $F_1 = 21$; $F_2 = 16$; $F_3 = 14$; $F = 21$. Resultó ser crítico el primer tramo. Las permutaciones son admisibles sólo entre las operaciones de los tramos primero y segundo. Tenemos:

$$\begin{aligned}
 R^1 &= (1, 2, 4; 3, 5, 6; 7, 8, 9); & F_1 &= 17; & F_2 &= 20; & F_3 &= 14; & F &= 20; \\
 R^2 &= (1, 3, 4; 2, 5, 6; 7, 8, 9); & F_1 &= 19; & F_2 &= 18; & F_3 &= 14; & F &= 19; \\
 R^3 &= (1, 3, 6; 2, 4, 5; 7, 8, 9); & F_1 &= 22; & F_2 &= 18; & F_3 &= 14; & F &= 22; \\
 R^4 &= (2, 3, 6; 1, 4, 5; 7, 8, 9); & F_1 &= 21; & F_2 &= 16; & F_3 &= 14; & F &= 21.
 \end{aligned}$$

En la solución R^2 el primer tramo quedó crítico pero el análisis de las soluciones vecinas no mejora la situación, pues son admisibles sólo las permuta-

ciones de los dos primeros tramos, mientras que el segundo tramo tiene el tiempo de trabajo $F_2 = 18$, próximo al tiempo crítico $F = 19$. Para disminuir F_2 , estudiemos las permutaciones admisibles de las operaciones destinadas para los tramos segundo y tercero.

Se tienen dos permutaciones admisibles de las operaciones 6, 7 y de las 5, 7. Las combinaciones correspondientes son:

$$R^5 = (1, 3, 4; 2, 5, 7; 6, 8, 9); \quad F_1 = 19; \quad F_2 = 14; \quad F_3 = 18; \quad F = 19;$$

$$R^6 = (1, 3, 4; 2, 6, 7; 5, 8, 9); \quad F_1 = 19; \quad F_2 = 16; \quad F_3 = 16; \quad F = 19.$$

En ambas soluciones, R^5 y R^6 , el valor de F_2 disminuyó. Realicemos la siguiente etapa del descenso, al tomar, por ejemplo, la solución R^6 . Las soluciones vecinas de la misma son:

$$R^7 = (1, 2, 3; 4, 6, 7; 5, 8, 9); \quad F_1 = 21; \quad F_2 = 14; \quad F_3 = 16; \quad F = 21;$$

$$R^8 = (1, 2, 4; 3, 6, 7; 5, 8, 9); \quad F_1 = 17; \quad F_2 = 18; \quad F_3 = 16; \quad F = 18;$$

$$R^9 = (1, 3, 6; 2, 4, 7; 5, 8, 9); \quad F_1 = 22; \quad F_2 = 13; \quad F_3 = 16; \quad F = 22.$$

Demostremos que la solución R^6 es óptima. Con este fin veamos la aplicación del método de ramificaciones y restricciones al problema 6b del § 6.1. Definamos el proceso de ramificaciones del modo siguiente. Supongamos predestinadas las operaciones para k tramos, es decir, están definidos R_1, R_2, \dots, R_k conjuntos. Llamemos el conjunto R_{k+1} admisible, si: 1) la designación del conjunto de operaciones R_{k+1} en el $(k+1)$ -ésimo tramo no perturba la sucesión del montaje; 2) $\sum_{i \in R_{k+1}} t_i \leq C$ (C es la duración prefijada del ciclo); 3) no existe un conjunto R' tal, que $R_{k+1} \subseteq R'$,

que satisfaga las condiciones 1 y 2. La estimación desde abajo del número de tramos para cualquier subconjunto de soluciones tiene la forma: $\varphi(R_1, \dots,$

$$R_k) = k + \frac{1}{C} \sum_{i \in R_k} t_i, \text{ donde } R(k) \text{ es el conjunto de operaciones por ahora}$$

no designadas.

Tomemos en nuestro ejemplo $C = 17$. Para la designación de las operaciones en el primer tramo hay cinco posibilidades:

$$R_1^1 = (1, 4, 7); \quad \varphi(R_1^1) = 1 + \frac{37}{17} = 3 \frac{3}{17};$$

$$R_1^2 = (1, 2, 4); \quad \varphi(R_1^2) = 1 + \frac{34}{17} = 3;$$

$$R_1^3 = (1, 3); \quad \varphi(R_1^3) = 1 + \frac{36}{17} = 3 \frac{2}{17};$$

$$R_1^4 = (3, 6); \quad \varphi(R_1^4) = 1 + \frac{36}{17} = 3 \frac{2}{17};$$

$$R_1^5 = (2, 3); \quad \varphi(R_1^5) = 1 + \frac{37}{17} = 3 \frac{3}{17};$$

Elegimos $R_1^2 = (1, 2, 4)$. Admitamos las siguientes designaciones para el segundo tramo:

$$R_2^1 = (3, 5; 7); \varphi(R_1^2, R_2^1) = 2 + \frac{18}{17} = 3 \frac{1}{17};$$

$$R_2^2 = (3, 6); \varphi(R_1^2, R_2^2) = 2 + \frac{19}{17} = 3 \frac{2}{17}.$$

De aquí concluimos que para $C = 17$ no son suficientes tres tramos, lo que confirma el carácter óptimo de la solución elegida R^B del problema 6a.

Ejercicio. Muéstrase que si $C = 17$, en el problema 6b es óptima la designación de las operaciones para 4 tramos. Hállense dichas designaciones.

Veamos cómo se aplica la optimización local al problema del viajante de comercio. Tenemos un conjunto $A = \{a_1, \dots, a_n\}$ de ciudades y la matriz de valores $C = [c_{ij}]$ del traslado de una ciudad a_i a otra ciudad a_j . En cada paso del algoritmo construiremos una ruta $A_r = (a_{i_1}, \dots, a_{i_r})$ a base del r -subconjunto del conjunto A ; $r = 1, 2, \dots, n$. Además consideramos que tiene lugar el traslado de la ciudad a_{i_k} a la ciudad $a_{i_{k+1}}$, $k = 1, 2, \dots, r-1$, y de la ciudad a_{i_r} a la a_{i_1} . Tomamos para $r = 1$: $A_1 = (a_i)$ (la ruta consta de una ciudad arbitrariamente elegida a_i). Si A_r está construido, buscamos en el siguiente paso la ciudad $a_i \in (A \setminus A_r)$ tal que $\varrho(a_i, A_r) = \min_{a_j \in (A \setminus A_r)} \varrho(a_j, A_r)$, donde con $\varrho(a_i, A_r)$ se denota la "distancia" desde a_i hasta A_r : $\varrho(a_i, A_r) = \min\{c_{ij} | a_j \in A_r\}$. En este caso, si $\varrho(a_i, a_2) = C_{i_1, i_2}$ para $a_{i_1} \in A_r$, entonces se toma $A_{r+1} = (a_{i_1}, \dots, a_{i_r}, a_i, a_{i_1}, \dots, a_{i_r})$. El algoritmo construido se denomina algoritmo de inclusión de la ciudad más próxima.

Ejercicio. Muéstrase que el algoritmo de inclusión de la ciudad más próxima puede ser realizado mediante un número de operaciones proporcional a n^2 .

Teorema I. Supongamos que la $(n \times n)$ -matriz de valores C en el problema del viajante de comercio es simétrica y satisface la desigualdad triangular: $c_{ik} + c_{kj} \geq c_{ij}$ ($i, j, k = 1, \dots, n$). Entonces

$$\frac{|I_n|}{|O_n|} < 2,$$

donde I_n es la ruta prefijada por el algoritmo de inclusión de la ciudad más próxima; O_n , la ruta óptima; $|I_n|$ y $|O_n|$ son sus valores sumarios.

Demostración. Supongamos que las ciudades a_1, a_2, \dots, a_n están enumeradas de un modo tal que en el i -ésimo paso del algoritmo se le agrega a la ruta la ciudad a_i . Demostraremos el teorema estableciendo una correspondencia biunívoca entre las ciudades a_2, \dots, a_n y todas las aristas en O_n (a excepción de la arista mayor) de un modo tal que el valor de inclusión de la ciudad a_k entre a_i y a_j , es decir, $c_{ik} + c_{kj} - c_{ij}$ (para que el primer paso sea correcto, tomamos $c_{ij} = 0$), no sobrepase el valor duplicado de la arista en O_n , que corresponde a a_k . Obtendremos en realidad una estimación más fuerte que la indicada en la formulación: $|I_n| \leq 2(|O_n| - |l_{\max}|)$, donde l_{\max} es la arista de mayor longitud en O_n .

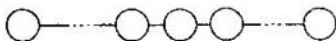


Fig.6.17.

Estudiemos en cada paso del algoritmo no sólo la ruta corriente $a_1, \dots, a_t, 1 \leq t \leq n$, sino también algunas aristas adicionales de $O_n \setminus \{I_{\max}\}$ que unen las ciudades $a_{t+1}, a_{t+2}, \dots, a_n$ con el itinerario corriente. En el primer paso el itinerario consta de la ciudad a_1 e incluye todas las aristas en O_n , a excepción de la I_{\max} (véase fig. 6.17).

Supóngase que en el $(k-1)$ -ésimo paso tenemos la ruta a_1, \dots, a_{k-1} (en la fig. 6.18 se indica con una línea gruesa) con ciudades agregadas a ella a_k, \dots, a_n con ayuda de las "piernas" (aristas dibujadas en línea no gruesa); una construcción de esta índole se llamará arácnea.

Admitamos que la ciudad a_k , que ha de ser incluida en la ruta en el k -ésimo paso, es la más próxima a la ciudad a_m perteneciente a la ruta (véase fig. 6.18). El valor de inclusión de a_k entre a_m y la siguiente ciudad a_t , es igual a $c_{tk} + c_{km} - c_{tm}$. Sea a_x un punto de la ruta al cual se une la pierna que incluye a_k , y sea (a_x, a_y) la primera arista de dicha pierna (puede ser que $a_y = a_k$). Por cuanto a_k es una ciudad próxima a la ruta, tenemos

$$c_{mk} \leq c_{xy}. \quad (1)$$

Según la desigualdad triangular, tenemos

$$c_{tk} \leq c_{tm} + c_{mk}. \quad (2)$$

De (1) y (2) obtenemos

$$c_{tk} \leq c_{tm} + c_{xy}. \quad (3)$$

Al adicionar (1) y (3) y al hacer uso de la simetría de la matriz C , obtenemos

$$c_{tk} + c_{km} \leq c_{tm} + 2c_{xy}.$$

lo que es equivalente a la desigualdad

$$c_{tk} + c_{km} - c_{tm} \leq 2c_{xy}.$$

es decir, la inclusión de a_k entre a_t y a_m cuesta $2c_{xy}$ como máximo. Incluida a_k , suprimimos la arista (a_x, a_y) de la configuración (la construcción sigue siendo arácnea). A la ciudad a_k se le hace corresponder la arista (a_x, a_y) . Después que el algoritmo pasa por todas las ciudades a_2, \dots, a_n , se establece la correspondencia requerida entre $\{a_2, \dots, a_n\}$ y $O_n \setminus \{I_{\max}\}$, lo que demuestra el teorema.

Mostremos ahora que se tiene un problema del viajante de comercio con n ciudades ($n \geq 6$), para el cual el algoritmo de inclusión de la ciudad más próxima proporciona una ruta que es dos veces mayor que la óptima. Veamos la siguiente matriz de valores C : $c_{ij} = c_{ji} = \min(j-i, n-j+i)$, es decir c_{ij} es la longitud de la ruta más corta de i a j , que va por las aristas del tipo $(k, (k+1) \pmod n)$. Para $n=8$, esta configuración viene expuesta en la fig. 6.19.

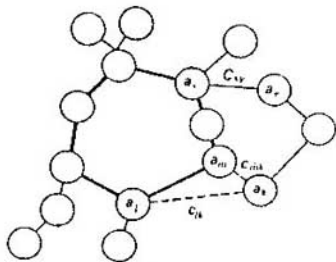


Fig. 6.18.

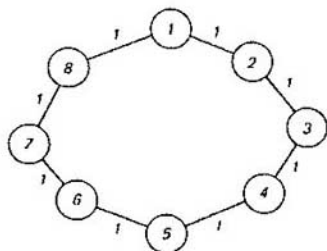


Fig. 6.19.

Demostremos que la ruta T_n compuesta por las aristas $(1, 2)$, $(n-1, n)$ y $(i, i+2)$ para $1 \leq i \leq n-2$ (en la fig. 6.20 se muestra la ruta T_8), puede obtenerse como resultado de aplicar el algoritmo de inclusión de la ciudad más próxima. Observemos que el algoritmo puede aplicarse a partir de la ciudad 1, luego se agrega a ésta la ciudad 2, la ciudad 3, etc. Sea T_k , $3 \leq k \leq n$ la ruta obtenida con ayuda de este algoritmo después de la inclusión de la ciudad k . Mostremos por inducción que T_k consta de las aristas $(1, 2)$, $(k-1, k)$ y $(i, i+2)$ para $1 \leq i \leq k-2$. Cuando $k=3$, esto es evidente. Suponiendo que esto es cierto para $3 \leq i \leq k$, observemos que T_{k+1} se obtiene de T_k por inclusión de la ciudad $k+1$ entre las ciudades $k-1$ y k .

La longitud de la ruta T_n es igual a $2n-2$, mientras que la ruta óptima O_n , compuesta por las aristas $(1, n)$ y $(i, i+1)$, $1 \leq i \leq n-1$, tiene una longitud igual a n . Por eso tenemos

$$\frac{|T_n|}{|O_n|} = 2 \left(1 - \frac{1}{n} \right).$$

Esta relación es la máxima que se admite por el teorema, puesto que $\lim_{n \rightarrow \infty} \frac{|T_n|}{|O_n|} = 1$.

La investigación teórica de la optimización local se realiza también con ayuda de los métodos teórico-probabilísticos. Con este fin en el conjunto de problemas que se resuelven por el algoritmo local dado se introduce una medida probabilística y se demuestra que con el crecimiento de la dimensión del problema, la medida de los problemas cuyas soluciones se desvían de la solución óptima en una magnitud superior a ϵ , tiende a 0, cuando la dimensión del problema crece. Tal comportamiento de las soluciones lleva el nombre de optimalidad asintótica. La optimalidad asintótica está demostrada para muchos algoritmos de solución del problema del viajante de comercio, en particular, para el más sencillo de ellos, cuando en cada paso se supone ir a la ciudad que está más próxima a aquella, en la que el viajante de comercio se encuentra en el momento dado. El lector puede encontrar una información más completa en [26].

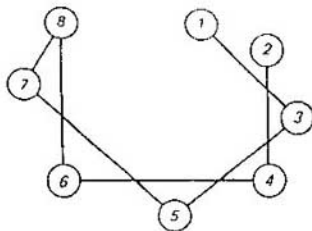


Fig. 6.20.

Un tipo más de algoritmos heurísticos de optimización está relacionado con la elección aleatoria de las combinaciones conforme a cierta distribución probabilística que se elige a partir de los razonamientos de obtención más probable de las combinaciones deseables. Para datos más detallados véase [81].

"Algoritmo ávido". Veamos un conjunto finito S , a cada elemento del cual s está asignado cierto número no negativo $w(s)$, llamado peso de s . El peso del subconjunto $A \subseteq S$ se determina como suma de pesos de todos los elementos de A . Sea \mathcal{F} una familia de subconjuntos del conjunto S .

Varios problemas de la optimización combinatoria se reducen al problema siguiente: hallar en la familia \mathcal{F} un subconjunto de peso máximo. Por ejemplo, a este tipo de problemas se reduce el problema de hallar el esqueleto (o el bosque) de peso máximo en un grafo ponderado G , si \mathcal{F} es un juego de todos los esqueletos (bosques) del grafo G .

Para resolver este problema resulta natural aplicar el siguiente algoritmo que se denomina "ávido" (greedy, en inglés). Este es como sigue:

Paso 1. Se elige un elemento s_1 tal que $\{s_1\} \in \mathcal{F}$, y $w(s_1) \geq w(s)$ para todos aquellos s que $\{s\} \in \mathcal{F}$. Si tal s_1 no existe, entonces tenemos una parada (indisponibilidad).

Paso 2. Se elige un elemento s_2 tal que $\{s_1, s_2\} \in \mathcal{F}$, y $w(s_2) \geq w(s)$ para todos aquellos $s \neq s_1$ que $\{s_1, s_2\} \in \mathcal{F}$. Si tal s_2 no existe, entonces tenemos una parada.

Paso k . Se elige tal elemento s_k , distinto de s_1, s_2, \dots, s_{k-1} , que $\{s_1, s_2, \dots, s_{k-1}, s_k\} \in \mathcal{F}$, y $w(s_k)$ sea máximo entre todos los s de este tipo. Si tal s_k no existe, tenemos una parada.

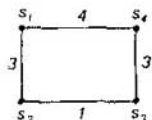
Es obvio que el algoritmo "ávido" finaliza su trabajo al obtener un subconjunto de la familia \mathcal{F} que sea máximo por inclusión. Sin embargo, dicho subconjunto puede tener en \mathcal{F} un peso no máximo. En efecto, si $S = \{a, b, c, d\}$, $w(a) = 4$, $w(b) = 3$, $w(c) = 2$, $w(d) = 2$, y $\mathcal{F} = \{\{a\}, \{a, c\}, \{b, c, d\}, \{b, d\}\}$, entonces, sirviéndonos del algoritmo "ávido", obtendremos el subconjunto $\{a, c\}$ cuyo peso es igual a 6, aunque el subconjunto de peso máximo es $\{b, c, d\}$ y el peso de éste es igual a 7. Al mismo tiempo, si cambiamos los pesos, poniendo $w(a) = 6$, $w(b) = 3$, $w(c) = 2$, $w(d) = 2$, el algoritmo "ávido" nos ofrece de nuevo el subconjunto $\{a, c\}$ de peso 9, que esta vez ya es un subconjunto de peso máximo.

El algoritmo "ávido" merece, obviamente, su nombre. Tiene una constante aspiración de incluir en un subconjunto de la familia \mathcal{F} el elemento de peso máximo posible. No lo hace sólo en el caso en que, al agregar tal elemento, el subconjunto se hace inadmisibles, es decir, no pertenecerá a \mathcal{F} .

Estudiamos la correlación existente entre la solución obtenida por el algoritmo "ávido" y la estructura de la familia \mathcal{F} .

Un par (S, \mathcal{F}) se llama *sistema de conjuntos*, si partiendo de que $A \in \mathcal{F}$, $A' \subseteq A$ se deduce que $A' \in \mathcal{F}$. Los elementos de la familia \mathcal{F} se denominan en este caso *conjuntos independientes*.

Se puede comprobar que con ayuda del algoritmo "ávido" se puede hallar en un grafo ponderado el bosque de peso máximo. Al mismo tiempo, el algoritmo "ávido" no trabaja para el problema sobre una combinación de pares máxima para el grafo. En efecto, el algoritmo "ávido", aplicado a un grafo ponderado G :



daría la combinación de pares $\{(s_1, s_4), (s_2, s_3)\}$, mientras que la combinación de pares máxima ponderada en G es $\{(s_1, s_2), (s_1, s_4)\}$.

El análisis de los sistemas de conjuntos que son distintos por su naturaleza y origen, pero que poseen la propiedad de que el algoritmo "ávido" resuelve los correspondientes problemas combinatorios de la optimización nos conduce a la siguiente definición.

Un sistema de conjuntos $M = (S, \mathcal{I})$ se denomina *matroide*, si el algoritmo "ávido" resuelve correctamente cualquier problema combinatorio concreto referente a la optimización para el sistema M .

Teorema 2. Sea $M = (S, \mathcal{I})$ un sistema de conjuntos. Entonces son equivalentes las siguientes afirmaciones:

- 1) M es un matroide;
- 2) si $F_1, F_2 \in \mathcal{I}$ donde $|F_1| = p$, y $|F_2| = p + 1$, existe un elemento $s \in F_2 - F_1$ tal que $F_1 \cup \{s\} \in \mathcal{I}$;
- 3) si A_1 y A_2 son subconjuntos independientes, máximos por inclusión, del conjunto A , entonces $|A_1| = |A_2|$.

Demostración. (1) \Rightarrow (2). Sea M un matroide, pero se supone que la afirmación 2 del teorema no se cumple, es decir, existen dos subconjuntos independientes $F_1, F_2 \in \mathcal{I}$ tales que $|F_1| = p$, $|F_2| = p + 1$, y el subconjunto $F_1 \cup \{s\} \notin \mathcal{I}$, cualquiera que sea $s \in F_2 - F_1$. Supongamos que la función ponderal sobre el conjunto S viene dada del modo siguiente:

$$w(s) = \begin{cases} p + 2, & \text{si } s \in F_1; \\ p + 1, & \text{si } s \in F_2 - F_1; \\ 0, & \text{si } s \notin F_1 \cup F_2. \end{cases}$$

Entonces, el conjunto independiente F_1 no es óptimo, es decir, no posee máximo. Efectivamente, $w(F_2) \geq (p+1)^2 > p(p+2) = w(F_1)$. El algoritmo "ávido" para este problema concreto empezará por la elección de todos los elementos del conjunto F_1 , puesto que estos elementos son de peso máximo. A continuación, el algoritmo "ávido" no podrá aumentar el peso total, ya que para todos los elementos restantes s o bien $F_1 \cup \{s\} \in \mathcal{F}$, si $s \in F_2$, o bien $w(s) = 0$. Por consiguiente, el algoritmo "ávido" da una solución no óptima de F_1 , por lo cual M no es un matroide, lo que contradice nuestra suposición. Por lo tanto, la afirmación 2 se cumple y la implicación en consideración queda demostrada.

(2) \Rightarrow (3). Supongamos que la afirmación 2 del teorema se cumple y sean A_1 y A_2 dos subconjuntos independientes, máximos por inclusión, del conjunto $A \subseteq S$. Admitamos que $|A_1| < |A_2|$. Rechacemos $|A_2| < |A_1| - 1$ elementos del conjunto A_2 . Entonces, por ser el sistema \mathcal{F} cerrado respecto de la inclusión, obtendremos un subconjunto independiente $A_3 \subseteq A$ tal que $|A_3| = |A_1| + 1$. En virtud de la afirmación 2 del teorema, existe un elemento $s \in A_3 - A_1$ tal que $A_1 \cup \{s\} \in \mathcal{F}$. Por consiguiente, el subconjunto A_1 no es un subconjunto independiente, máximo por inclusión, del conjunto A . La contradicción obtenida demuestra la implicación requerida.

(3) \Rightarrow (1) Supongamos que tiene lugar la afirmación 3 del teorema. Mostremos que en este caso el algoritmo "ávido" resuelve el problema de la optimización combinatoria sobre el sistema de conjuntos M . Admitamos que esto no es así, es decir, para cierto conjunto de pesos $w(s)$, donde $s \in S$, el algoritmo "ávido" lleva a un conjunto independiente $F = \{s_1, s_2, \dots, s_i\}$, mientras que existe un conjunto $F' = \{s'_1, s'_2, \dots, s'_j\}$ tal que $w(F') > w(F)$. Sin restringir la generalidad de los razonamientos podemos considerar que los elementos de los conjuntos F y F' están ordenados de tal modo que $w(s_1) \geq w(s_2) \geq \dots \geq w(s_i)$ y $w(s'_1) \geq w(s'_2) \geq \dots \geq w(s'_j)$. Podemos considerar evidentemente, que F' es un subconjunto independiente, máximo por inclusión, del conjunto S . Según la construcción, el subconjunto F' es también subconjunto del conjunto S , máximo por inclusión. Por eso, en virtud de la afirmación 3 del teorema, para $A = S$ obtenemos que $i = j$. Demostremos que para todos los $m = 1, 2, \dots, i$ se verifica la desigualdad $w(s_m) \geq w(s'_m)$ lo que contradirá la suposición de que $w(F') > w(F)$, con lo que quedará por terminada la demostración de nuestra implicación.

Realicemos la demostración por inducción respecto de m . Para $m = 1$ la afirmación es evidente. Supongamos ahora que $w(s_m) < w(s'_m)$ para cierto $m > 1$, y que $w(s_p) \geq w(s'_p)$ para $p = 1, 2, \dots, m-1$. Veamos un conjunto $A = \{s \in S: w(s) \geq w(s'_m)\}$. El conjunto $\{s_1, \dots, s_{m-1}\}$ es subconjunto independiente, máximo por inclusión, del conjunto A , puesto que si $\{s_1, s_2, \dots, s_{m-1}, s\} \in \mathcal{F}$, y $w(s) \geq w(s'_m) > w(s_m)$, entonces el algoritmo "ávido" debería elegir s , en lugar de s_m , a título del siguiente elemento del conjunto F_1 . Esto contradice la afirmación 3 del teorema, por cuanto

$\{s'_1, s'_2, \dots, s'_m\}$ es otro subconjunto independiente del conjunto A de mayor potencia. Con esto termina la inducción y toda la demostración del teorema.

En el teorema 2 se ha demostrado en realidad la equivalencia existente entre el concepto de matroide (introducido aquí) como sistema de conjuntos sobre los cuales el algoritmo "ávido" da soluciones óptimas y otras definiciones de los matroides que se consideran en el capítulo 8. En el párrafo 8.4 se aducirán diferentes ejemplos de los matroides, sobre los cuales el algoritmo "ávido" es siempre eficaz.

6.4. OPTIMIZACIÓN EN LOS GRAFOS

En los párrafos antecedentes de este capítulo ya se han examinado problemas combinatorios extremales, relacionados con la optimización en los grafos (véase el problema del viajante de comercio). En el presente párrafo estudiaremos una serie más de problemas del mismo tipo.

Al principio nos dedicaremos al estudio de los grafos orientados $G = (X, \Gamma)$ que no tienen contornos. Introduzcamos para ellos la noción de función ordinal. Así se denomina una función $O(x_i)$ definida sobre un conjunto de vértices $X = \{x_i\}$, que toma valores de números enteros. En este caso, si $x_j \in \Gamma(x_i)$, entonces $O(x_j) > O(x_i)$. Dicho de otro modo, si el vértice x_j sigue tras el vértice x_i , entonces el valor de la función ordinal en el vértice x_j es superior a su valor en el vértice x_i . La magnitud $O(x_i)$ se llamará nivel del vértice x_i .

Las funciones ordinales para un grafo concreto $G = (X, \Gamma)$ pueden construirse mediante métodos diferentes. He aquí uno de éstos. Veamos los subconjuntos de los vértices N_0, N_1, \dots, N_r definidos del modo siguiente:

$$\begin{aligned} N_0 &= \{x_i | x_i \in E; \Gamma^{-1}(x_i) = \emptyset\}; \\ N_1 &= \{x_i | x_i \in E \setminus N_0; \Gamma^{-1}(x_i) \in N_0\}; \\ &\dots\dots\dots \\ N_r &= \{x_i | x_i \in (E \setminus \bigcup_{k=0}^{r-1} N_k); \Gamma^{-1}(x_i) \in \bigcup_{k=0}^{r-1} N_k\}. \end{aligned}$$

Es fácil ver que los conjuntos N_0, N_1, \dots, N_r forman una partición X , y la función O , que toma el valor k sobre el conjunto N_k ($k = 0, 1, \dots, r$) es una función ordinal del grafo G . En la fig. 6.21 $\Gamma(N_r) = \emptyset$, está expuesto un grafo cuyo conjunto de vértices está partido en subconjuntos N_0, N_1, \dots, N_r .

A la par con la función ordinal sobre el conjunto de vértices del grafo se define también una función numérica φ . Por medio de esta función se determinan las características numéricas de los problemas, que han de ser optimizadas. Supongamos que (x_1, x_2, \dots, x_n) es un camino en el grafo G , es decir, $x_1, x_2, \dots, x_n \in X; x_{i+1} \in \Gamma(x_i)$ ($i = 1, 2, \dots, n-1$). Llamemos

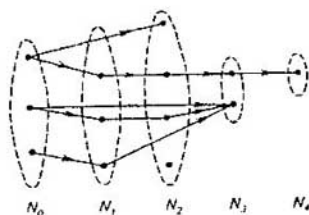


Fig. 6.21.

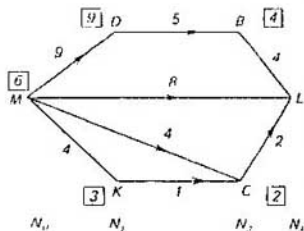


Fig. 6.22.

valor de este camino a un número $\lambda = \varphi(x_1) + \varphi(x_2) + \dots + \varphi(x_n)$.

A veces la función numérica φ se define no sobre los vértices, sino sobre los arcos. En tal caso se llama valor del camino (x_1, x_2, \dots, x_n) a un número $\lambda = \varphi(x_1, x_2) + \varphi(x_2, x_3) + \dots + \varphi(x_{n-1}, x_n)$.

Muchos problemas de optimización tienen la siguiente formulación: hallar entre cierto conjunto de caminos P el mínimo (o el máximo), es decir, un camino que posea el valor mínimo (o el máximo) de λ . Como conjunto P podemos elegir, por ejemplo, el conjunto de todos los caminos hamiltonianos. En esta sección analizaremos el problema de hallar un camino mínimo y máximo entre todas los caminos que unen dos vértices fijos.

Teorema 3 (de optimalidad). Supongamos que un cierto camino que une el vértice x de nivel m y el vértice x' de nivel s , es mínimo (máximo). Entonces, su subcamino entre el vértice y de nivel k y el vértice y' de nivel p ($m \leq k < p \leq s$) es también mínimo (máximo).

Demostración. Sea el camino $(x, \dots, y, \dots, y', \dots, x')$ mínimo. Si su subcamino (y, \dots, y') no es mínimo, existe, entonces, otro camino de y a y' de menor valor. Sustituycamos el subcamino anterior en el camino (x, \dots, x') por uno nuevo, y obtendremos, como resultado, el nuevo camino desde x hasta x' de menor valor, lo que lleva a una contradicción.

El teorema alegado yace en la base del método para hallar los caminos mínimos en un grafo sin contornos. Este método fue propuesto por Bellman y obtuvo el nombre de método de programación dinámica.

Expliquemos la esencia de este método con un ejemplo. Sea dado un grafo $G = (X, U)$ (fig. 6.22); la función numérica viene dada en los arcos. Buscaremos el camino de M a A con el valor mínimo.

La función ordinal $O(x)$ adquiere los siguientes valores:

$$O(M) = 0; O(D) = O(K) = 1; O(B) = O(C) = 2; O(A) = 3.$$

Examinamos sucesivamente, a partir de A , todos los vértices del grafo en el orden de decrecimiento de su función ordinal y asignamos a todo vértice un número igual al valor mínimo del camino que separa dicho vértice de A .

Para buscar los caminos mínimos en los grafos que tienen contornos

existen también métodos diferentes. Puede ocurrir que en tales grafos no haya caminos máximos.

Algoritmo de Ford. Supongamos que una función numérica l viene dada sobre los arcos del grafo G que no tiene contornos. Podemos considerar que l toma sólo valores positivos (de lo contrario, esto puede conseguirse con facilidad agregando un número suficientemente grande a todos los $l(x_i, x_j)$). Se requiere hallar el camino mínimo que lleva desde x_0 hasta x_n . A todo vértice x_i le asignaremos símbolos de acuerdo con el siguiente algoritmo:

1) pongamos $\lambda_i = \infty$ en cada vértice x_i , salvo $\lambda_0 = 0$;

2) buscamos un arco (x_i, x_j) tal que $\lambda_j - \lambda_i > l(x_i, x_j)$ y sustituimos λ_j por $\lambda_i + l(x_i, x_j) < \lambda_j$; procedemos de este modo hasta que se haga posible encontrar un arco que disminuya por lo menos un solo valor.

Demostremos que el número λ_n , hallado según la regla citada, será el valor del camino mínimo de x_0 a x_n (si $\lambda_n = \infty$, tal camino no existe).

Se puede indicar un vértice x_r , tal que con ayuda del arco (x_r, x_n) se haya disminuido el valor de λ_n por la última vez, y

$$\lambda_n = \lambda_r + l(x_r, x_n).$$

Análogamente, existe un vértice x_s , tal que

$$\lambda_r = \lambda_s + l(x_s, x_r),$$

etc. La sucesión $\lambda_n, \lambda_r, \lambda_s, \dots$ es estrictamente decreciente. Por consiguiente, para cierto k obtenemos $\lambda_{r_{k+1}} = 0$, es decir, $x_{r_{k+1}} = x_0$.

Demostremos que el itinerario $(x_0, x_{r_1}, x_{r_2}, \dots, x_n)$ es mínimo (su valor es igual a λ_n). Tomemos un camino arbitrario que lleva de x_0 a x_n : $(x_0, x_{v_1}, x_{v_2}, \dots, x_{v_m}, x_n)$. Son válidas las siguientes correlaciones

$$\lambda_{r_1} - 0 \leq l(x_0, x_{v_1});$$

$$\lambda_{r_2} - \lambda_{v_1} \leq l(x_{v_1}, x_{v_2});$$

$$\dots\dots\dots$$

$$\lambda_n - \lambda_{v_m} \leq l(x_{v_m}, x_n).$$

Al sumar término a término estas desigualdades, obtendremos

$$\lambda_n \leq l(x_0, x_{v_1}) + l(x_{v_1}, x_{v_2}) + \dots + l(x_{v_m}, x_n),$$

lo que se trataba de demostrar.

Observemos que el algoritmo de Ford puede emplearse para buscar el camino máximo en los grafos sin contornos. En este caso se debe poner $\lambda_i = 0$ para todos los vértices sin excepción alguna, y luego sustituir λ_j por $\lambda_j' = \lambda_i + l(x_i, x_j)$, si $\lambda_j' > \lambda_j$, hasta que se pueda aumentar λ_j .

Mencionemos, por fin, el algoritmo de Bellman—Kalaba. Este último emplea el llamado principio de optimalidad: "cualquier subcamino del camino mínimo es un camino mínimo entre los vértices correspondientes".

Sea un grafo orientado $G = (X, U)$; $|X| = n + 1$. Numeremos todos los vértices de 0 hasta n .

Sea l_{ij} un valor que se le asigna al arco (x_i, x_j) ($l_{ij} = \infty$, si $(x_i, x_j) \notin U$; $l_{ij} = 0$, si $i = j$). Buscaremos un camino $(x_0, x_{i_1}, \dots, x_{i_n}, x_n)$ tal que el valor del camino $l_{0,i_1} + l_{i_1,i_2} + \dots + l_{i_n,n}$ sea mínimo.

Examinemos el sistema de ecuaciones

$$v_i = \min_{j \neq i} (v_j + l_{ij}); \quad i = 0, 1, 2, \dots, n-1;$$

$$j = 0, 1, 2, \dots, n.$$

Entonces, las soluciones v_i de este sistema son valores de los caminos mínimos de x_i a x_n ; $i = 0, 1, \dots, n-1$. El hecho de que los valores de los caminos mínimos satisfacen el sistema mencionado se deduce del principio de optimalidad. La afirmación recíproca de que las soluciones del sistema son valores de los caminos óptimos se deducirá del algoritmo que se usa para hallar v_i .

Pongamos

$$v_i^{(0)} = l_{in}; \quad i = 0, 1, 2, \dots, n-1; \quad v_n^{(0)} = 0.$$

Calculamos sucesivamente:

$$v_i^{(1)} = \min_{j \neq i} (v_j^{(0)} + l_{ij}); \quad i, j = 0, 1, 2, \dots, n-1;$$

.....

$$v_i^{(k)} = \min_{j \neq i} (v_j^{(k-1)} + l_{ij});$$

$$v_n^{(k)} = 0.$$

Los cálculos se realizan hasta que se cumpla el sistema de igualdades

$$v_i^{(k)} = v_i^{(k-1)}; \quad i = 0, 1, 2, \dots, n.$$

Del procedimiento que se usa para el cálculo se deduce que $v_i^{(m)}$ es el camino mínimo del vértice x_i al vértice x_n entre todos los caminos que pasan no más que por m arcos. Por consiguiente, para hallar la solución del sistema son suficientes $n-1$ iteraciones.

Ha de notarse que el algoritmo de Bellman—Kalaba puede ser aplicado para la búsqueda del camino máximo en un grafo sin contornos. Para ello es suficiente sólo sustituir en todas las correlaciones mín por máx y poner l_{ij} igual a $-\infty$, si el arco (x_i, x_j) no se contiene en el grafo.

6.5. FLUJOS EN LAS REDES

Al resolver los problemas extremales se emplea a menudo su interpretación en términos de los flujos en las redes. Se denomina *red* a un grafo orientado conexo $G = (X, U)$ con un conjunto de vértices X y un conjunto de arcos U , donde a todo arco $(x_i, x_j) \in U$ se le asigna una característica numérica no negativa $c(x_i, x_j)$ que se llama *capacidad de paso*. En G se distinguen dos vértices fijos: s y t ; s se denomina *entrada* (también se conoce como fuente), t se denomina *salida* (también se llama, a veces, vertedero) y los restantes vértices son *intermedios*.

Se llama *flujo estacionario* de valor v desde s a t de la red $G = (X, U) = (X, \Gamma)$ a una función numérica f , definida en el conjunto de arcos U , que satisface las siguientes ecuaciones lineales y desigualdades:

$$\sum_{y \in \Gamma(x)} f(x, y) - \sum_{y \in \Gamma^{-1}(x)} f(y, x) = \begin{cases} v, & x = s; \\ 0, & x \neq s, t; \\ -v, & x = t; \end{cases} \quad (1)$$

$$0 \leq f(x, y) \leq c(x, y) \text{ para cualesquiera } (x, y) \in U. \quad (2)$$

Introduzcamos las siguientes designaciones con el fin de simplificar las notaciones de las correlaciones de flujo.

Sean X, Y ciertos conjuntos; entonces, por (X, Y) se entenderá el conjunto de todos los pares (x, y) ; $x \in X, y \in Y$.

Pongamos para cierta función $g(x, y)$ y $h(x)$:

$$g(X, Y) = \sum_{(x, y) \in (X, Y)} g(x, y);$$

$$h(X) = \sum_{x \in X} h(x).$$

El problema del flujo máximo estacionario es el problema de maximización de la variable v que obedece a las restricciones de flujo (1) y (2).

Se denomina *corte* L en la red (X, U) , que separa los vértices s y t , a un conjunto de arcos (Y, \bar{Y}) , donde $Y \subseteq X, \bar{Y} = X - Y, s \in Y, t \in \bar{Y}$.

El número $c(Y, \bar{Y})$ lleva el nombre de *capacidad de paso del corte* (Y, \bar{Y}) .

Señalemos que cualquier camino de s a t contiene al menos un arco del corte (Y, \bar{Y}) . Por eso, si de la red se excluyen los arcos de tal o cual corte, en la red nueva no quedará ningún camino de s a t , y el valor del flujo máximo por la citada red será igual a cero. Se pone claro intuitivamente que el valor v de un flujo arbitrario f no sobrepasa la capacidad de paso de cualquier corte. Demostremos que esta afirmación es cierta.

Lema 1. Sea f un flujo de s a t en la red (X, Γ) de valor v ; supongamos que (Y, \bar{Y}) es un corte que separa s de t . Entonces

$$v = f(Y, \bar{Y}) - f(\bar{Y}, Y) \leq c(Y, \bar{Y}).$$

Demostración. Reescribamos las relaciones (1) en una forma más compacta:

$$f(s, X) - f(X, s) = v;$$

$$f(x, X) - f(X, x) = 0; \quad x \neq s, t;$$

$$f(t, X) - f(X, t) = -v.$$

Sumemos término a término aquellas de estas relaciones que se refieren a $x \in Y$. Teniendo en cuenta que $s \in Y$, y $t \in \bar{Y}$, obtenemos

$$v = f(Y, X) - f(X, Y);$$

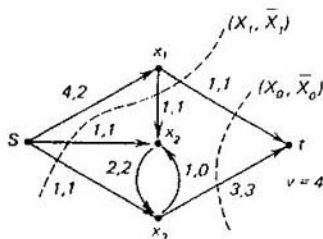


Fig 6.23.

haciendo uso de la relación $X = Y \cup \bar{Y}$, obtenemos

$$\begin{aligned} \nu &= f(Y, Y \cup \bar{Y}) - f(Y \cup \bar{Y}, Y) = \\ &= f(Y, Y) + f(Y, \bar{Y}) - f(Y, Y) - f(\bar{Y}, Y) = \\ &= f(Y, \bar{Y}) - f(\bar{Y}, Y). \end{aligned}$$

Resta por observar que $f(\bar{Y}, Y) \geq 0$; $f(Y, \bar{Y}) \leq c(Y, \bar{Y})$, con lo que queda demostrado el lema.

El corte de capacidad de paso mínima (el cual siempre existe, puesto que el conjunto de todos los cortes es finito) se denomina mínimo. De lo demostrado más arriba se deduce que la intensidad del flujo máximo no sobrepasa la capacidad de paso del corte mínimo.

Veamos la red expuesta en la fig. 6.23, donde el primer número en el arco es igual a su capacidad de paso mientras que el segundo número, al valor de la función ν . Entonces, (X_0, X_0) y (X_1, X_1) , donde $X_0 = \{s, x_1, x_2, x_3\}$ y $X_1 = \{s, x_1\}$, son cortes mínimos de la red con una capacidad de paso igual a 4. Observemos, además, que en dicha red fluye el flujo máximo cuya magnitud es también igual a 4.

El siguiente teorema sobre el flujo máximo y el corte mínimo es el resultado principal referente al flujo estacionario máximo en la red.

Teorema 1. En cualquier red el valor máximo de un flujo de s a t es igual a la capacidad de paso mínima del corte que separa s de t .

Demostración. Del lema 1 se deduce que resulta suficiente comprobar la existencia del flujo f y del corte (Y, \bar{Y}) , para los cuales el valor del flujo y la capacidad de paso del corte son iguales. En este caso el flujo será máximo y el corte, mínimo.

Tomemos el flujo máximo f (tal flujo existe, evidentemente, aunque pueden haber varios de ellos) y definamos de una manera recurrente algún corte:

a) $s \in Y$;

b) Si $x \in Y$, y $f(x, y) < c(x, y)$, entonces $y \in Y$; si $x \in Y$, y $f(y, x) > 0$, entonces $y \in Y$.

Demostremos que el conjunto obtenido (Y, \bar{Y}) es un corte en la red $G = (X, U)$. Para ello basta demostrar que $t \in \bar{Y}$.

Supongamos que esto no es así. Entonces, de la definición del conjunto

Y se desprende que existe un camino de s a t ($s, x_1, x_2, \dots, x_n, t$) que posee la propiedad de que para todos los arcos (x_i, x_{i+1}) del camino citado tiene lugar o bien

$$f(x_i, x_{i+1}) < c(x_i, x_{i+1}), \quad (3)$$

o bien

$$f(x_{i+1}, x_i) > 0. \quad (4)$$

Sea ε_1 el mínimo de la diferencia $c - f$, calculada en aquellos arcos del camino (s, x_1, \dots, x_n, t) , para los cuales se verifica la desigualdad (3), y sea ε_2 el mínimo de f en todos los arcos inversos, en los que se verifica la desigualdad (4). Pongamos $\varepsilon = \min\{\varepsilon_1, \varepsilon_2\} > 0$. Modifiquemos ahora el flujo f del modo siguiente: aumentemos f en la magnitud ε en los arcos del camino (s, x_1, \dots, x_n, t) , sobre los cuales se verifican las desigualdades (3) y disminuyamos f en la magnitud ε en los arcos inversos, sobre los cuales se verifica (4). Una función, definida de este modo, será, evidentemente, el flujo en la red de s a t , cuyo valor es igual a $v + \varepsilon$. Mas, esto contradice la suposición de que el flujo f es máximo. Por consiguiente, $t \in \bar{Y}$, y el conjunto (Y, \bar{Y}) es el corte que separa s de t .

De la definición de Y se deduce:

$$\begin{aligned} f(x, \bar{x}) &= c(x, \bar{x}) \text{ para } (x, \bar{x}) \in (Y, \bar{Y}); \\ f(\bar{x}, x) &= 0 \text{ para } (\bar{x}, x) \in (\bar{Y}, Y). \end{aligned}$$

Por consiguiente,

$$v = f(Y, Y) - f(\bar{Y}, Y) = c(Y, Y).$$

El teorema está demostrado.

Del lema demostrado 1 y del teorema 4 pueden deducirse algunos corolarios.

Diremos que cierto itinerario de s a t aumenta el flujo f , si para cada arco U o bien $f(U) < c(U)$, o bien $f > 0$ en el arco inverso de U .

Corolario 1. El flujo f es máximo, cuando y sólo cuando no existe ningún itinerario que aumente el flujo f .

El arco (x, y) se llamará *saturado de flujo f* , si $f(x, y) = c(x, y)$, y *libre del flujo*, si $f(x, y) = 0$.

Corolario 2. Un corte (Y, \bar{Y}) es mínimo, si y sólo si cada flujo máximo f satura todos los arcos del corte (Y, \bar{Y}) y deja libres todos los arcos pertenecientes a (\bar{Y}, Y) .

El siguiente teorema demuestra que el corte mínimo (Y, \bar{Y}) construido en el teorema 4 no depende, en realidad, de cómo se elige el flujo máximo f .

Teorema 5. Supongamos que (Z, \bar{Z}) es un corte mínimo arbitrario; f , cierto flujo máximo y (Y, \bar{Y}) , un corte construido por el flujo f en el teorema 4. Entonces

$$Y \subseteq Z.$$

Demostración. Sea $Y \not\subseteq Z$; entonces $Y \in Z \subset Y$. Demostremos la siguiente afirmación: $(Y \cap Z, \overline{Y \cap Z})$ es también un corte mínimo. Para la demostración hagamos uso del corolario 2. Sea $x \in Y \cap Z$, y $\ell \in \overline{Y \cap Z}$. Por consiguiente, $x \in Y$; $x \in Z$; $\bar{x} \in \overline{Z \cup Y}$, y, por tanto, (x, ℓ) pertenece por lo menos a uno de los cortes: o bien a (Y, \overline{Y}) , o bien a (Z, \overline{Z}) y entonces, de acuerdo con el corolario 2, el flujo máximo satura el arco (x, ℓ) . Análogamente se puede demostrar que el flujo máximo arbitrario f deja libre el arco (ℓ, x) . Por consiguiente, $(Y \cap Z, \overline{Y \cap Z})$ es un corte mínimo.

Supongamos que el vértice $x \in Y$, y $x \notin Y \cap Z$. Entonces, $x \neq s$, y existe un camino de s a x , digamos, $(s = x_1, x_2, \dots, x_n = x)$ tal que cada uno de sus arcos o bien está saturado de flujo f , o bien el arco inverso no es libre del flujo. Por cuanto $s \in Y \cap Z$, y $x \notin Y \cap Z$, se encontrará un par (x_i, x_{i+1}) , en el que $x_i \in Y \cap Z$ y $x_{i+1} \in Y \cap Z$. Demostrado el hecho de que $(Y \cap Z, \overline{Y \cap Z})$ es el corte mínimo, podemos concluir, de acuerdo con el corolario 2, que todos los arcos rectos de él están saturados de flujo f , y todos los arcos inversos de él son libres del flujo f . Hemos llegado, pues, a una contradicción y, por lo tanto, $Y \subseteq Z$. El teorema está demostrado.

Varias entradas y salidas. Supongamos que el conjunto de vértices x de la red (X, U) está partido en tres conjuntos:

S es un conjunto de entradas,

T , un conjunto de salidas,

R , un conjunto de vértices intermedios.

Por flujo f de S a T se entiende una función numérica f que está definida en U y satisface las condiciones:

$$f(x, X) - f(X, x) = 0; \quad x \in R;$$

$$f(S, X) - f(X, S) = v;$$

$$f(T, X) - f(X, T) = -v;$$

$$0 \leq f(x, y) \leq c(x, y).$$

Ensanchemos la red (X, U) , agregando dos vértices s^* y t^* y todos los arcos (s^*, S) y (T, t^*) . Obtendremos una nueva red (X^*, U^*) . Definamos adicionalmente la función de la capacidad de paso c en U^* , poniendo

$$c^*(S^*, x) = \infty; \quad x \in S;$$

$$c^*(x, t^*) = \infty; \quad x \in T;$$

$$c^*(x, y) = c(x, y); \quad (x, y) \in U.$$

Es fácil ver que el estrechamiento f de cualquier flujo f^* de s^* a t^* en la red (X^*, U^*) representa el flujo de S a T en la red (X, Y) . Y viceversa, cualquier flujo f de S a T en la red (X, U) continúa como el flujo f^* de s^* a t^* en la red (X^*, U^*) :

$$f^*(s^*, x) = f(x, X) - f(X, x); \quad x \in S;$$

$$f^*(x, t^*) = f(X, x) - f(x, X); \quad x \in T;$$

$f^*(x, y) = f(x, y)$ en todos los demás casos.

Por consiguiente, el problema del flujo máximo de S a T en la red (X, U) es equivalente al del flujo máximo en una red ensanchada con una entrada y una salida.

Método de marcación para resolver problemas del flujo máximo. Existen varios algoritmos para hallar el flujo máximo. No obstante, en la base de la mayoría de ellos yace el llamado "método de marcación". Apoyándonos en la demostración del teorema 1 y de sus corolarios, damos a conocer este método eficaz de construcción del flujo máximo y del corte mínimo.

Para asegurar la terminación del proceso, convengámonos en considerar que la función de la capacidad de paso c toma sólo valores de números enteros. En la práctica esto no es una restricción sustancial, puesto que las capacidades de paso racionales pueden reducirse a los números enteros a cuenta del cambio de la escala.

Sea dado un flujo f de números enteros en la red (X, U) . (En el instante inicial el flujo f puede ser nulo). Asignemos a los vértices de la red dada unas marcas que pueden tener la forma (x^+, ε) ó (x^-, ε) , donde $x \in X$, y ε es un número natural, o bien ∞ . Al realizarse la operación A , pueden ser posibles tres estados de un vértice: a) el vértice no está marcado; b) el vértice está marcado, pero no observado; c) el vértice está marcado y observado.

Operación A (marcación). La entrada S recibe la marca $(-\infty)$, es decir, $\varepsilon(S) = \infty$. Ahora la entrada está marcada, pero no observada, mientras que todos los demás vértices no están marcados. En general, elijamos cualquier vértice x que está marcado, pero no observado. Supongamos que tiene la marca $(z^+, \varepsilon(x))$. Entonces, a todos los vértices y , que no están marcados y para los cuales $f(x, y) < c(x, y)$, les asignamos la marca $(x^+, \varepsilon(y))$, donde

$$\varepsilon(y) = \min[\varepsilon(x), c(x, y) - f(x, y)].$$

Ahora los vértices y están marcados, pero no observados. A los vértices que después de esto quedan no marcados, pero para los cuales $f(x, y) > 0$, les asignamos la marca $(x^-, \varepsilon(y))$, donde

$$\varepsilon(y) = \min[\varepsilon(x), f(x, y)].$$

Tales vértices y están ahora marcados y no observados, y el vértice x está marcado y observado. Esta operación se repite hasta que resulte marcada la salida t , o bien hasta que no sea posible marcar ningún vértice, mientras que la salida quede no marcada. En el último caso el proceso de construcción de f se considera terminado. En el primer caso pasamos a la operación B.

Operación B (cambio del flujo). Supongamos que la salida t tiene la marca $(y^+, \varepsilon(t))$; entonces $f(y, t)$ se sustituye por $f(y, t) + \varepsilon(t)$; en cambio, si t está marcado con $(y^-, \varepsilon(t))$, entonces $f(t, y)$ se sustituye por $f(t, y) - \varepsilon(t)$. Luego, en cualquiera de estos casos pasamos al vértice y . En

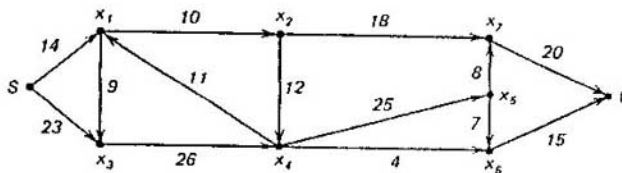


Fig. 6.24.

general, si y tiene la marca $(x^+, c(y))$, entonces $f(x, y)$ se sustituye por $f(x, y) + c(t)$, y si el vértice y está marcado con $(x^-, c(y))$, entonces $f(y, x)$ se sustituye por $f(x, y) - c(t)$, después de lo cual se pasa al vértice x . Conseguida la entrada s , el cambio del flujo se da por terminado. Se borran todas las marcas y de nuevo se pasa a la operación Λ para el nuevo aumento del flujo.

En el proceso de marcación, si pasamos de la operación A a la B , esto quiere decir que hemos encontrado el camino de s a t , que aumenta el flujo f . En cambio, si la operación A está terminada, mientras que la salida, no marcada, entonces el flujo fue máximo y el conjunto de arcos que unen los vértices marcados con los no marcados, forman el corte mínimo.

Veamos ahora un ejemplo que ilustra la aplicación del método de marcación para resolver los problemas de flujo máximo. Sea una red expuesta en la fig. 6.24 con la entrada s y la salida t . Las capacidades de paso de los arcos vienen indicadas también en dicha figura. Se requiere hallar el flujo máximo de s a t .

Operación A . Asignemos al vértice s una marca $(-\infty)$.

El conjunto de vértices $\{x_j | x_j \in \Gamma(s), f(s, x_j) < c(s, x_j), x_j \text{ no está marcado}\}$ es $\{x_1, x_3\}$, al vértice x_1 se le asigna la marca $(s^+; \min\{\infty, 14-0\})$, es decir, $(s^+; 14)$; al vértice x_3 se le asigna la marca $(s^+; \min\{\infty, 23-0\})$, es decir, $(s^+; 23)$.

El conjunto $\{x_j | x_j \in \Gamma^{-1}(s), f(x_j, s) > 0, x_j \text{ no están marcados}\}$ es vacío. Así pues, el vértice s está marcado y observado; los vértices x_1 y x_3 están marcados y no observados, y todos los demás vértices, no marcados.

Observemos ahora el vértice x_1 $\{x_j | x_j \in \Upsilon(x_1), f(x_1, x_j) < c(x_1, x_j), x_j \text{ no está marcado}\} = \{x_2\}$, para x_2 de marca sirve $(x_1^+; \min\{14, 10-0\}) = (x_1^+; 10)$. $\{x_j | x_j \in \Gamma^{-1}(x_1), f(x_j, x_1) > 0, x_j \text{ no está marcada}\} = \emptyset$. Ahora los vértices s y x_1 están marcados y observados, mientras que los x_2 y x_3 , marcados, pero no observados.

Al tomar para la observación el vértice x_2 , obtenemos las siguientes marcas: $(x_2^+; 10)$ para x_4 y $((x_2^+; 10)$, para x_7 .

Al tomar para la observación x_3 , llegamos a que no se puede poner ninguna marca. Al continuar la observación desde el vértice x_4 , obtendremos las siguientes marcas: $((x_4^+; 10)$ para x_5 , $((x_4^+; 4)$ para x_6 y $(x_6^+; 4)$, para t .

Operación B . Aquí obtenemos: $x^* = t$; $f(x_6, t) = 0 + 4 = 4$; $x^* = x_6$;

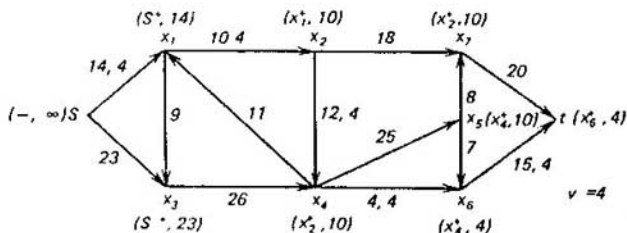


Fig. 6.25.

$f(x_4, x_6) = 0 + 4 = 4$; $x = x_4$; $f(x_2, x_4) = 0 + 4 = 4$; $x = x_2$; $f(x_1, x_2) = 0 + 4 = 4$, $x = x_1$;

La forma del flujo al final de la operación B y las marcas de los vértices antes de que ellas quedan borradas se muestran en la fig. 6.25, donde el primer número en el arco indica la capacidad de paso, y el segundo, el valor de la función f .

Borrando las marcas de los vértices y volviendo a la operación A para la segunda pasada, obtendremos nuevas marcas de los vértices (los vértices marcados, pero no observados, los observaremos en el orden de crecimiento de sus números ordinales).

Operación A. La marca para s será $(-, \infty)$; para x_1 , $(s^+, 10)$, y para x_3 , $(s^+, 23)$; ahora el vértice s está marcado y observado. La marca para x_2 será $(x_1^+, 6)$ y el vértice x_1 queda marcado y observado. La marca para x_4 será $(x_2^+, 6)$ y para x_7 , $(x_2^+, 6)$. Ahora el vértice x_2 está marcado y observado, igual que el vértice x_3 . La marca para x_5 será $(x_4^+, 6)$ y el vértice x_4 se hace marcado y observado. La marca para x_6 será $(x_5^+, 6)$, después de lo cual el vértice x_5 será marcado y observado. Por último, la marca para t será $(x_6^+, 6)$.

Operación B. El nuevo flujo ha aumentado del modo siguiente:

$$f(x_6, t) = 4 + 6 = 10; \quad f(x_5, x_6) = 0 + 6 = 6; \quad f(x_4, x_5) = 0 + 6 = 6; \\ f(x_2, x_4) = 4 + 6 = 10; \quad f(x_1, x_2) = 4 + 6 = 10; \quad f(s, x_1) = 4 + 6 = 10.$$

Todos los demás valores de la función f quedan los mismos. La forma nueva del flujo y las marcas de los vértices se muestran en la fig. 6.26.

Continuando el proceso, obtenemos tras cada paso del algoritmo, los flujos y las marcas que se muestran sucesivamente en las figs. 6.27—6.29.

El algoritmo finaliza su trabajo cuando ningún vértice puede ser marcado, como tampoco la salida t ; las marcas "finales" se muestran en la fig. 6.30. El flujo expuesto en la fig. 6.30 es, por eso, el máximo del valor 29, mientras que el correspondiente corte mínimo en este dibujo está representado con una línea punteada.

Ejercicios.

1. Hállese, empleando el método de marcación, el flujo máximo de la entrada s a la salida t en las redes expuestas en las figs. 6.31 a y 6.31b. Junto a los arcos en la fig. 6.31 se indican las capacidades de paso y los valores corrientes de la función del flujo f .

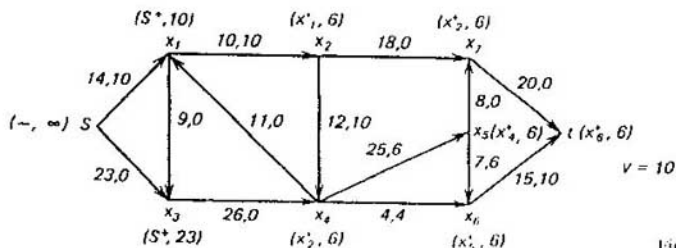


Fig. 6.26.

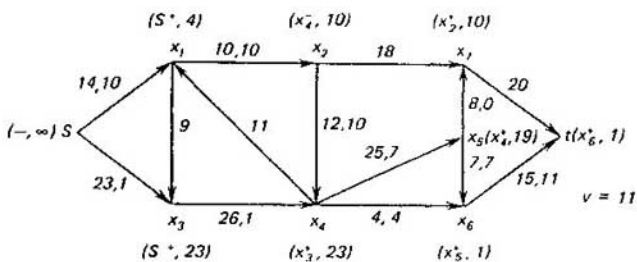


Fig. 6.27.

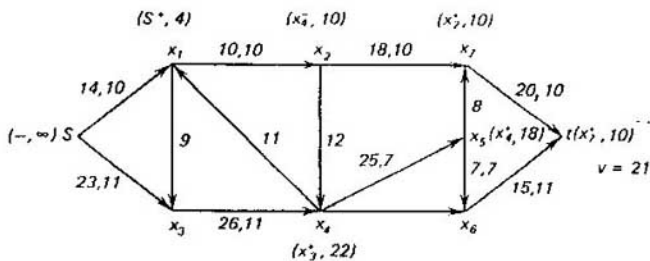


Fig. 6.28.

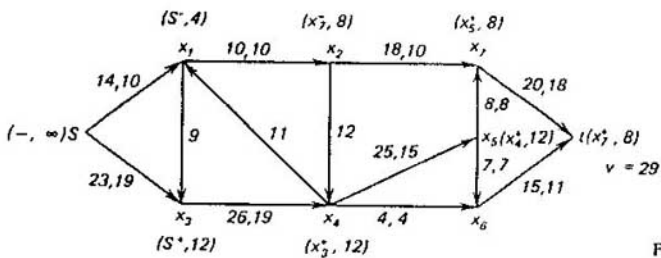


Fig. 6.29.

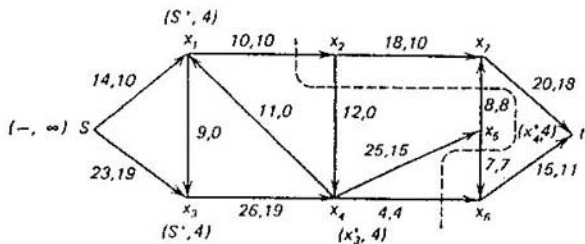


Fig. 6.30.

2. Hállese el corte mínimo que separa la entrada s de la salida t para las redes del ejercicio 1. ¿Saturará el flujo máximo todos los arcos de este corte?

3. Enúnciense las condiciones, bajo las cuales en una red arbitraria existe un camino, que aumenta el flujo, de la entrada s a la salida t , que sea compuesto íntegramente de arcos inversos (arcos rectos).

Del algoritmo de marcación se deduce que si el flujo original era de números enteros (es decir, $f(u)$ representaba números enteros para cualquier $u \in U$), entonces todos los flujos que siguen después de los aumentos resultarán ser de números enteros. Por consiguiente, es lícito el

Teorema 6. Si la función de la capacidad de paso c es de números enteros, existe un flujo máximo que también es de números enteros.

Restricciones inferiores de los flujos de arco. En la definición de flujo, en lugar de la desigualdad (2) podemos hacer

$$l(x, y) \leq f(x, y) \leq c(x, y), \quad (5)$$

donde $l(x, y)$ es una función real definida sobre los arcos del conjunto U , y $0 \leq l(x, y) \leq c(x, y)$.

El proceso de marcación se extiende inmediatamente a este caso con algunas variaciones insignificantes. Siendo vigentes las restricciones (5), se demuestra con facilidad el análogo del teorema básico 4.

Teorema 7. Si existe una función f que satisface, para cierto v , las desigualdades (5) y ecuaciones (1), entonces el valor máximo del flujo de s a t es igual al mínimo de la diferencia $c(\overline{Y}, \overline{Y}) - l(\overline{Y}, \overline{Y})$ tomado en todos los cortes (Y, \overline{Y}) de la red.

Más abajo se demostrarán los teoremas que proporcionan las condi-

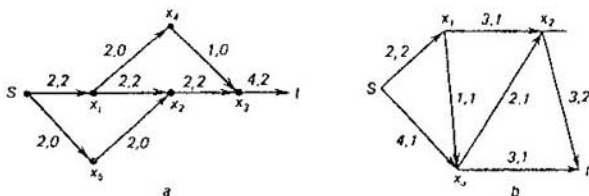


Fig. 6.31.

ciones necesarias y suficientes para la existencia de los flujos en las redes, las cuales satisfacen diferentes desigualdades lineales.

Teorema de demanda y oferta. Sea $G = (X, U)$ una red arbitraria con la capacidad de paso $c = c(x, y)$, donde $(x, y) \in U$, y X está partido en tres conjuntos: $X = S \cup R \cup T$; S son las entradas, R son los vértices intermedios, T son las salidas. A todo vértice $x \in S$ se le asigna un número $a(x)$, y a cada vértice $y \in T$ se le asigna el número $b(y)$; $a(x) \geq 0$, $x \in S$; $b(y) \geq 0$, $y \in T$. El número $a(x)$ puede entenderse como oferta de la mercancía en la entrada x ; $b(y)$, como la demanda de la mercancía en la salida y . Surge el siguiente problema: ¿cuáles son las condiciones necesarias y suficientes, bajo las cuales la demanda en las salidas puede ser satisfecha por la oferta en las entradas? En el lenguaje algebraico esto puede expresarse así:

$$\begin{aligned} f(x, X) - f(X, x) &\leq a(x); & x \in S; \\ f(x, X) - f(X, x) &= 0; & x \in R; \\ f(X, x) - f(x, X) &\geq b(x); & x \in T; \\ 0 &\leq f(x, y) \leq c(x, y). \end{aligned} \quad (6)$$

Teorema 8. Las restricciones (6), donde $a(x) \geq 0$ y $b(x) \geq 0$ son admisibles cuando y sólo cuando para cada subconjunto $Y \subseteq X$:

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq c(Y, \bar{Y}). \quad (7)$$

Demostración. Demostremos al principio la necesidad de la desigualdad (7). Supongamos que existe un flujo f que satisface las condiciones (6). Al sumar estas desigualdades respecto de todos los vértices $x \in \bar{Y}$, obtenemos una desigualdad

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq f(X, \bar{Y}) - f(\bar{Y}, X). \quad (8)$$

$X = Y \cup \bar{Y}$, por lo cual (8) puede ser transformada:

$$b(T \cap \bar{Y}) - a(S \cap \bar{Y}) \leq f(Y, \bar{Y}) - f(\bar{Y}, Y).$$

Haciendo uso del resultado del lema 1, demostramos la necesidad de la condición 7.

Demostremos ahora la suficiencia de la condición (7). Hagamos ensanchar la red $G = (X, U)$, uniendo la entrada ficticia s , la salida t y los arcos (s, S) y (T, t) . Obtendremos una red nueva X^*, U^* . La función que caracteriza la capacidad de paso sobre U^* la definiremos según las fórmulas:

$$\begin{aligned} c^*(s, x) &= a(x); & x \in S; \\ c^*(x, t) &= b(x); & x \in T; \\ c^*(x, y) &= c(x, y); & (x, y) \in U. \end{aligned}$$

La validez de la desigualdad (7) significa que el corte (T, t) en la red (X^*, U^*) es mínimo. Demostremoslo. Sea (Y^*, \bar{Y}^*) un corte arbitrario que separa

s de t . Pongamos $Y = Y^* \setminus s$, $\bar{X} = X^* \setminus t$, entonces

$$\begin{aligned} c^*(Y^*, \bar{Y}^*) - c^*(T, t) &= c^*(Y, t) + c^*(s, \bar{Y}) + \\ + c^*(Y, \bar{Y}) - c^*(T, t) &= b(T \cap Y) + a(S \cap \bar{Y}) + c(Y, \bar{Y}) - b(T) = \\ &= -b(T \cap \bar{Y}) + a(S \cap \bar{Y}) + c(Y, \bar{Y}). \end{aligned}$$

Aplicando el teorema de flujo máximo y corte mínimo, llegamos a que el cumplimiento de la condición (7) tiene por resultado la existencia del flujo f^* de s a t en la red (X^*, U^*) que satura todos los arcos del corte (T, t) . En tal caso el flujo f que representa un estrechamiento del flujo f^* en la red (X, U) satisface las condiciones (6), puesto que

$$\begin{aligned} a(x) \geq f^*(s, x) &= f^*(x, X) - f^*(X, x) = f(x, X) - f(X, x); \quad x \in S; \\ b(x) = f^*(x, t) &= f^*(X, x) \text{ y } f^*(x, X) = f(X, x) - f(x, X); \quad x \in T. \end{aligned}$$

El teorema está demostrado.

Teorema de circulación. Demostremos un teorema más, en el cual se trata de la existencia de las circulaciones, es decir, de los flujos que están privados de entradas y salidas y que están limitados en los arcos por las fronteras superiores e inferiores prefijadas.

Supongamos que para la red (X, U) están prefijadas l y c , es decir, las funciones de frontera inferior y superior, respectivamente, en U . La circulación admisible en la red (X, Y) es f , que está definida sobre U y que satisface las condiciones:

$$f(x, X) - f(X, x) = 0; \quad x \in X; \quad (9)$$

$$0 \leq l(x, y) \leq f(x, y) \leq c(x, y); \quad (x, y) \in U. \quad (10)$$

Teorema 9. Para que las restricciones (9) y (10) sean admisibles, es necesario y suficiente que para cualquier $Y \subseteq X$ se cumpla la desigualdad

$$c(Y, \bar{Y}) \geq l(\bar{Y}, Y).$$

Demostración. Hagamos ensanchar la red (X, U) añadiendo dos vértices s y t , y un conjunto de arcos (s, X) y (X, t) . En la nueva red (X^*, U^*) definamos una función que tiene una capacidad de paso según las fórmulas:

$$c^*(x, y) = c(x, y) - l(x, y); \quad (x, y) \in U;$$

$$c^*(s, x) = l(X, x); \quad x \in X;$$

$$c^*(x, t) = l(x, X); \quad x \in X.$$

Si en la red (X, U) existe una circulación admisible f , entonces ésta engendra el flujo f^* de s a t en la red (X^*, U^*) :

$$f^*(x, y) = f(x, y) - l(x, y); \quad (x, y) \in U;$$

$$f^*(s, x) = l(X, x); \quad x \in X;$$

$$f^*(x, t) = l(x, X); \quad x \in X.$$

Al contrario, si en la red (X^*, U^*) existe un flujo $f^*(x, y)$, entonces en la red (X, U) puede determinarse la circulación $f(x, y)$ al poner:

$$f(x, y) = f^*(x, y) + l(x, y) \text{ para } (x, y) \in U.$$

Se comprueba con facilidad el cumplimiento de las relaciones (9) y (10).

Así pues, el problema se ha reducido a la cuestión sobre la existencia del flujo de s a t en la red (X^*, U^*) de valor $l(X, X)$. La condición necesaria y suficiente para la existencia del flujo citado de s a t de valor $l(X, X)$ consiste en que las capacidades de paso de todos los cortes sean no menos de $l(X, X)$.

Sea (Y, \bar{Y}^*) un corte que separa s de t en la red (X^*, U^*) . Definamos el conjunto $Y \subseteq X$ del modo siguiente:

$$Y = Y^* \setminus s; \quad \bar{Y} = \bar{Y}^* \setminus t.$$

Entonces:

$$\begin{aligned} c^*(Y^*; \bar{Y}^*) &= c^*(Y \cup s, \bar{Y} \cup t) = c^*(Y, \bar{Y}) + c^*(s, \bar{Y}) + c^*(Y, t) = \\ &= c(Y, \bar{Y}) - l(Y, \bar{Y}) + l(X, \bar{Y}) + l(Y, X) = c(Y, \bar{Y}) + \\ &+ l(\bar{Y}, \bar{Y}) + l(Y, X) = c(Y, \bar{Y}) + l(X, X) - l(\bar{Y}, Y). \end{aligned}$$

Por consiguiente, $c^*(Y^*; \bar{Y}^*) \geq l(X, X)$ cuando y sólo cuando

$$c(Y, \bar{Y}) \geq l(\bar{Y}, Y).$$

Con ello queda demostrado el teorema.

Por último demostremos el análogo del teorema de König (véase § 4.1), ilustrando de este modo su conexión con el teorema sobre el flujo máximo y el corte mínimo.

Sea $G = G(S \cup T, U)$ un grafo bipartido arbitrario sobre un conjunto de vértices $S \cup T$ con un conjunto de arcos U , cuyos arcos están todos orientados de S en T . Se denomina combinación de pares en el grafo G un subconjunto E de aristas, de las cuales ningún par tiene vértices comunes. Un conjunto de vértices se denomina (S, T) -separador, si la eliminación en el grafo de los citados vértices junto con los arcos incidentes con ellos rompe todos los caminos que van desde los vértices del conjunto S a los vértices del conjunto T .

Teorema 10. Sea $G = G(S \cup T, U)$ un grafo bipartido. Entonces, el número máximo de arcos de la combinación de pares en el grafo G es igual al número mínimo de vértices en el conjunto (S, T) -separador.

Demostración. Juntemos al grafo en consideración G dos vértices s y t , y también los arcos (s, S) y (T, t) , y para todos los $x \in S$ e $y \in T$ definamos las capacidades de paso de los arcos del grafo:

$$\begin{aligned} c(s, x) &= 1; \quad x \in S; \\ c(x, t) &= 1; \quad x \in T; \\ c(x, y) &= \infty; \quad (x, y) \in U. \end{aligned} \tag{11}$$

Sea f un flujo máximo arbitrario de s a t , que se mide en números enteros, y sea (Y, \bar{Y}) cierto corte mínimo que separa s de t . Teniendo presente (11), concluimos que los arcos del conjunto $I = \{(x, y) \in U; f(x, y) = 1\}$ no tienen vértices comunes de dos en dos. El conjunto $D = (S \cap \bar{Y}) \cup (T \cap Y)$ es (S, T) -separador, con la particularidad de que D se encuentra en una correspondencia biunívoca con los arcos del corte mínimo (Y, \bar{Y}) . Al hacer uso del teorema sobre el flujo máximo y el corte mínimo, llegamos a que si v es el valor del flujo máximo f , entonces $|I| = |D| = v$. Por otra parte, el número máximo de arcos del grafo G que no tienen vértices comunes de dos en dos no es, obviamente, superior al número mínimo de vértices en cualquier conjunto (S, T) -separador. El teorema está demostrado.

A las relaciones existentes entre el teorema sobre el flujo máximo y el corte mínimo y otros problemas combinatorios extremales volveremos una vez más en el § 8.1.

El dominio de los problemas extremales de carácter discreto es enorme. Los esfuerzos de muchos matemáticos están dirigidos a resolverlos. Han sido elaborados algoritmos para la resolución de algunas clases de problemas. No obstante, no existe todavía una teoría íntegra y única en dicho dominio.

CAPÍTULO 7 MÉTODOS PROBABILÍSTICOS EN EL ANÁLISIS COMBINATORIO

7.1. EJEMPLOS DE APLICACIÓN DE LOS MÉTODOS PROBABILÍSTICOS

En el § 2.4 se han estudiado las aplicaciones del aparato combinatorio de funciones generatrices a los problemas de la teoría de las probabilidades (para un conocimiento más detallado de esta última se recomienda el libro [82]). En el presente capítulo el planteamiento del problema será inverso. Daremos a conocer un método (o, incluso, un grupo de métodos) de resolución de los problemas combinatorios por medios de carácter probabilístico. La esencia del método consiste en que los objetos combinatorios que se investigan se interpretan como sucesos en cierto espacio probabilístico discreto. Se demuestra que en este espacio la probabilidad de que un objeto, poseedor de las propiedades dadas, se realice es positiva. De este modo se demuestra la existencia de tal objeto, aunque dicho objeto no se construye explícitamente. El método no es perfecto por su carácter no constructivo, sin embargo, en muchos problemas combinatorios resulta suficiente establecer el propio hecho de existencia de un objeto. El método que se describe emplea el aparato desarrollado de la teoría de las probabilidades, lo que permite simplificar razonamientos engorrosos.

Explicaremos la esencia del método con ejemplos concretos.

Estimaciones de los números de Ramsey. Enunciemos el teorema de Ramsey del párrafo 3.3 para un caso particular en que $r = t = 2$, $\varphi_1 = \varphi_2 = k$.

Teorema 1. Existe un número natural mínimo $n = R(k)$ tal que para cualquier coloración en dos colores de las aristas de un grafo no orientado completo de n vértices se encontrará un subgrafo completo de k vértices, cuyas aristas están coloreadas de un mismo color.

Demostremos por el método probabilístico la estimación inferior para los números de Ramsey $R(k)$.

Teorema 2.

$$R(k) \geq k \cdot 2^{k/2} \cdot \left(\frac{1}{\sqrt{2e}} + o(1) \right) \text{ para } k \rightarrow \infty.$$

Demostración. Analicemos un grafo completo con un conjunto de vértices X , donde $|X| = n$. A título de espacio probabilístico elijamos el conjunto de todas las coloraciones de sus aristas en 2 colores: rojo y azul. Todas las coloraciones son equiprobables, es decir, cada arista se colorea, independientemente de las otras, de color rojo con una probabilidad igual a $1/2$, y de color azul, con una probabilidad de $1/2$. Examinemos un k -subconjunto $Y \subseteq X$. La probabilidad de que todas las aristas del subgrafo

generado por Y estén coloreadas de rojo será igual a $2^{-\binom{k}{2}}$. Para el color azul esta probabilidad es la misma. Por eso, la probabilidad de que el subconjunto dado de los vértices Y genere un subgrafo monocromático es igual a $2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$. El conjunto Y puede ser elegido por $\binom{n}{k}$ métodos. Por eso, la probabilidad de la existencia del k -subconjunto Y , que genera un subgrafo monocromático no es superior a $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. Por consiguiente, si los números n y k son de tal género que $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$, entonces el suceso se realiza con la probabilidad positiva $p = 1 - \binom{n}{k} \cdot 2^{1-\binom{k}{2}}$, cuando ningún subgrafo generado de k vértices es monocromático. En virtud de la finitud del conjunto de coloraciones esto quiere decir que existe una coloración, para la cual no existen subgrafos completos monocromáticos de k vértices. Así pues, de $\binom{n}{k} < 2^{\binom{k}{2}-1}$ se deduce: $n < R(k)$. Por consiguiente,

$$\binom{R(k)}{k} \geq 2^{\binom{k}{2}-1}.$$

Haciendo uso del desarrollo de Stirling $k! = k^k e^{-k} \cdot \sqrt{2\pi k} \cdot (1 + o(1))$ para $k \rightarrow \infty$, obtendremos

$$R(k)^k \geq k! \binom{R(k)}{k} \geq k! \cdot 2^{\binom{k}{2}-1}.$$

$$R(k) \geq \left(k^k e^{-k} \cdot \sqrt{2\pi k} \cdot 2^{\frac{k(k-1)}{2}-1} \right)^{\frac{1}{k}} \cdot (1 + o(1)) = \frac{k \cdot 2^{k/2}}{\sqrt{2e}} \cdot (1 + o(1)),$$

lo que se requería demostrar

Propiedad \mathcal{B} (véase la definición en el § 1.5). Designemos con $m(k)$ el número mínimo de miembros en el sistema de k -subconjuntos que no poseen la propiedad \mathcal{B} . Es fácil ver que el sistema de todos los k -subconjuntos del $(2k-1)$ -conjunto no posee la propiedad \mathcal{A} , por lo cual

$$m(k) \leq \binom{2k-1}{k}.$$

La estimación de debajo para $m(k)$ la obtendremos por el método de los primeros momentos.

Teorema 3.

$$m(k) \geq 2^{k-1}.$$

Demostración. Veamos un conjunto V y un sistema $A = \{S_1, S_2, \dots, S_m\}$ de sus k -subconjuntos que no posee la propiedad \mathcal{A} , donde $m = m(k)$.

Sea $C = (V_1, V_2)$ una partición del conjunto V , es decir, $V = V_1 \cup V_2$, $V_1 \cap V_2 = \emptyset$. Pongamos

$$f_{ij}(C) = \begin{cases} 1, & \text{si } S_j \subseteq V_i, \\ 0, & \text{si } S_j \not\subseteq V_i, \end{cases}$$

donde $i = 1, 2$; $j = 1, 2, \dots, m$. Entonces

$$\beta(C) = \sum_{j=1}^m \sum_{i=1}^2 f_{ij}(C)$$

es el número de miembros del sistema A que están íntegramente contenidos en una de las partes de la partición C .

Sea ahora ζ una partición accidental, donde cada elemento del conjunto V pertenece con la probabilidad $1/2$, independientemente de los otros elementos, a la primera parte de la partición, y con la misma probabilidad $1/2$, a la segunda parte. Entonces, cuando $i = 1, 2$ para cualquier subconjunto S_j de k elementos, tenemos:

$$Mf_{ij}(C) = P\{f_{ij}(C) = 1\} = 2^{-k}.$$

Calculemos la esperanza matemática $\beta(C)$:

$$M\beta(C) = \sum_{j=1}^m \sum_{i=1}^2 Mf_{ij}(C) = \sum_{j=1}^m \sum_{i=1}^2 2^{-k} = m \cdot 2^{1-k}.$$

Por cuanto A no posee la propiedad \mathcal{B} , entonces $\beta(C) \geq 1$ para cualquier partición C y, por lo tanto,

$$M\beta(C) \geq 1.$$

Así pues, $m \cdot 2^{1-k} \geq 1$, lo que se trataba de demostrar.

Problema sobre la cantidad máxima de números cuyas sumas son todas diferentes. Designemos con $g(n)$ la cantidad máxima k de números naturales a_1, \dots, a_k , no superiores a n , cuyas sumas $\sum_{i \in S} a_i$ son todas diferentes para diferentes $S \subseteq \{1, \dots, k\}$. Está claro que los números $1, 2, 4, \dots, 2^{k-1}$ satisfacen esta condición cuando $k \leq 1 + \log_2 n$, por lo cual

$$g(n) \geq 1 + \lceil \log_2 n \rceil.$$

Por otra parte, cualquiera de las sumas no sobrepasa kn . Si $k = g(n)$, se tienen en total $2^{k(n)}$ sumas diferentes y todas ellas están encerradas entre 0 y $ng(n)$, de donde $2^{k(n)} \leq ng(n)$. Por consiguiente

$$\begin{aligned} g(n) &\leq \log_2 n + \log_2 g(n) \leq \log_2 n + \log_2(\log_2 n + \log_2 g(n)) = \\ &= \log_2 n + \log_2(\log_2 n) + o(1). \end{aligned}$$

Esta estimación superior puede ser mejorada, si hacemos uso del método teórico-probabilístico de los segundos momentos.

Teorema 4.

$$g(n) \leq \log_2 n + \frac{1}{2} \log_2(\log_2 n) + o(1) \text{ para } n \rightarrow \infty.$$

Demostración. Supongamos que los números a_1, \dots, a_k satisfacen las condiciones del problema. Examinemos una magnitud aleatoria $\xi = \sum_{i \in S} a_i$, donde S es un subconjunto elegido al azar del conjunto $\{1, \dots, k\}$. Cada uno de los elementos de dicho conjunto integra S independientemente de los otros elementos con una probabilidad igual a $1/2$. Por eso,

$$M\xi = \sum_{i=1}^k P\{i \in S\} \cdot a_i = \frac{1}{2} \sum_{i=1}^k a_i.$$

Cada par de elementos $\{i, j\}$, donde $1 \leq i < j \leq k$, integra el subconjunto S con la probabilidad igual a $1/4$. Por eso,

$$\begin{aligned} M\xi^2 &= M\left(\sum_{i \in S} a_i\right)^2 = \sum_{i=1}^k P\{i \in S\} \cdot a_i^2 + \sum_{1 \leq i < j \leq k} P\{i, j \in S\} \cdot 2a_i a_j = \\ &= \frac{1}{2} \sum_{i=1}^k a_i^2 + \frac{1}{4} \sum_{1 \leq i < j \leq k} 2a_i a_j = \frac{1}{4} \sum_{i=1}^k a_i^2 + \frac{1}{4} \left(\sum_{i=1}^k a_i\right)^2. \\ D\xi &= M\xi^2 - (M\xi)^2 = \frac{1}{4} \sum_{i=1}^k a_i^2. \end{aligned}$$

Emplearemos la desigualdad de Chébishev ($\sigma > 0$):

$$P\{|\xi - M\xi| < \sigma\} \geq 1 - \frac{D\xi}{\sigma^2} = 1 - \frac{1}{4\sigma^2} \sum_{i=1}^k a_i^2 \geq 1 - \frac{kn^2}{4\sigma^2}.$$

Por cuanto todas las sumas $\sum_{i \in S} a_i$ son distintas, el número de subconjuntos $S \subseteq \{1, \dots, k\}$, que satisfacen la condición $\left|\sum_{i \in S} a_i - M\xi\right| < \sigma$, no sobrepasa 2σ . Por consiguiente,

$$P\{|\xi - M\xi| < \sigma\} \leq 2\sigma \cdot \left(\frac{1}{2}\right)^k$$

y

$$1 - \frac{kn^2}{4\sigma^2} \leq \sigma \cdot 2^{1-k}$$

Al poner $\sigma = n\sqrt{k}$, obtenemos que $\frac{3}{8} \cdot 2^k \leq n\sqrt{k}$, de donde encontramos:

$$k \leq \log_2 n + \frac{1}{2} \log_2 (\log_2 n) + o(1) \text{ para } n \rightarrow \infty,$$

lo que se requería demostrar.

7.2. PROBLEMAS DE PLANIFICACIÓN DEL EXPERIMENTO

Examinemos el siguiente modelo matemático del problema de experimentos de rechazo. En un conjunto $X = \{x_1, \dots, x_n\}$, compuesto de n elementos, hay que hallar s elementos fijos, desconocidos de antemano (llamémoslos *defectuosos*). Concretamente estos pueden ser los problemas siguientes: búsqueda de las monedas falsas, determinación de los defectos en un televisor, comprobación de un cable telefónico por medio de un tester, localización de un hombre enfermo en un numeroso grupo de gente con ayuda del análisis de la sangre, etc.

Los experimentos se llevan a cabo del modo siguiente. Se toma cierto subconjunto $Y \subseteq X$. Como resultado de las pruebas (test) obtenemos una información que se interpreta según el planteamiento del problema. Por ejemplo, podemos determinar si pertenece al conjunto Y por lo menos uno de los s elementos defectuosos o bien cuántos, precisamente, elementos defectuosos se contienen en Y .

El juego de conjuntos Y_i ($i = 1, 2, \dots, N$) se llama *plan de experimentos*. Partiendo de los resultados de N experimentos se necesita hallar unívocamente todos los s elementos defectuosos, con la particularidad de que en la elección de $Y_i \subseteq X$ ($i = 1, 2, \dots, N$) pueden imponerse algunas restricciones.

Resulta natural dividir el problema de planificación de los experimentos en dos clases. A la primera clase se refiere la planificación *estática*, donde la elección del i -ésimo conjunto Y_i no depende de los resultados de los $i - 1$ experimentos anteriores. La segunda clase es la planificación *sucesiva*, donde los resultados de las comprobaciones anteriores influyen en el planteamiento del experimento de turno. Veamos primeramente el problema de búsqueda de un elemento defectuoso.

Sea $s = 1$, es decir, entre los elementos del conjunto $X = \{x_1, \dots, x_n\}$ hay exactamente un elemento defectuoso. Supongamos, además, que está prefijado el plan estático de los experimentos, es decir, el juego de conjuntos $Y_i \subseteq X$, $i = 1, \dots, N$. Basándose en dicho juego se construye la matriz $A = |a_{ij}|$ de dimensión $N \times n$:

$$a_{ij} = \begin{cases} 1, & \text{si } x_j \in Y_i, \\ 0, & \text{si } x_j \notin Y_i \quad (1 \leq j \leq n; 1 \leq i \leq N). \end{cases}$$

A la inversa, a base de esta matriz se restablecen unívocamente los conjuntos Y_i ($i = 1, 2, \dots, N$).

Lema 1. Sea $s = 1$. Los experimentos contruidos con arreglo a la matriz $A = |a_{ij}|$ restablecen unívocamente el elemento defectuoso cuando y sólo cuando todas las columnas de la matriz A son diferentes.

Demostración. Supongamos que en la matriz A la k -ésima columna es igual a la columna l -ésima. En este caso no podremos distinguir cuál de los elementos es defectuoso: x_k ó x_l , puesto que dichos elementos o bien

pertenecen simultáneamente al conjunto Y_m ($1 \leq m \leq N$), o bien no lo pertenecen, razón por la cual todas las columnas de la matriz A han de ser diferentes.

Al contrario, sea e_i el resultado de la i -ésima comprobación, es decir, $e_i = 1$, si el elemento defectuoso pertenece al conjunto Y_i del plan de experimentos construido a base de la matriz $A = [a_{ij}]$, y $e_i = 0$, en el caso contrario. Así pues, realizadas N comprobaciones, obtendremos un vector de resultados $\vec{e} = (e_1, e_2, \dots, e_N)^T$ (T es una operación de transposición). Si el elemento x_k es defectuoso, entonces \vec{e} coincide con la k -ésima columna de la matriz $A = [a_{ij}]$. Por consiguiente, si todas las columnas de la matriz A son distintas, entonces partiendo del vector de los resultados restableceremos unívocamente el elemento defectuoso. La demostración está terminada.

Así pues, el problema de planificación estática consiste en hallar la matriz A , compuesta de ceros y unidades, tal que todas las n columnas cuyas sean distintas dos a dos y el número de filas, mínimo. Este número se denotará con $N_{\text{exp}}^{\text{est}}$.

Teorema 5. En el caso de un elemento defectuoso

$$N_{\text{exp}}^{\text{est}} = \lceil \log_2 n \rceil.$$

Demostración. Si todas las n columnas de una $(0, 1)$ -matriz A , compuesta de N filas, son distintas dos a dos, entonces su cantidad no sobrepasa la cantidad de todas las palabras binarias de longitud N , es decir, $n \leq 2^N$. Teniendo presente que N es un número entero y suponiendo $N = N_{\text{exp}}^{\text{est}}$, obtendremos: $N_{\text{exp}}^{\text{est}} \geq \lceil \log_2 n \rceil$.

La validez de la desigualdad recíproca se deduce de que la $(0, 1)$ -matriz, cuyas columnas representan la notación de los números naturales $1, 2, \dots, n$ en el sistema binario de numeración, contiene $\lceil \log_2 n \rceil$ filas. El teorema queda demostrado.

Analícemos ahora un caso en que se trata de dos elementos defectuosos. Si para $s = 1$ tenemos la fórmula exacta de $N_{\text{exp}}^{\text{est}}$, entonces cuando $s = 2$, las estimaciones superior e inferior obtenidas no coinciden. Ya en este caso se manifiestan dificultades de los problemas de búsqueda en la planificación del experimento para s grandes.

Al igual que para $s = 1$, al plan estático corresponderá unívocamente una matriz $A = [a_{ij}]$.

En el modelo lineal de la planificación estática obtendremos, como resultado del i -ésimo experimento, una información referente al número de elementos defectuosos que están contenidos en el conjunto Y_i . De este modo, el vector de los resultados será una palabra de longitud N sobre el alfabeto $\{0, 1, 2\}$.

Lema 2. Una matriz $A = [a_{ij}]$ es el plan estático para $s = 2$, es decir, restablece unívocamente el par de elementos defectuosos cuando y sólo cuando son diferentes todas las sumas de columnas cuyas tomadas de dos en dos.

Las sumas de las columnas se entienden como sumas tomadas por componentes, mientras que la sumación será habitual:

$$0 + 0 = 0, 0 + 1 = 1 + 0 = 1, 1 + 1 = 2.$$

La demostración del lema 2 es igual a la del lema 1. La única diferencia consiste en que esta vez el vector de los resultados es una suma de las columnas i -ésima y j -ésima, si (x_i, x_j) es el par de elementos defectuosos.

Teorema 6. En el caso de dos elementos defectuosos

$$N_{\text{exp}}^{\text{est}} \geq \log_3 \binom{n}{2}. \quad (1)$$

Demostración. Supongamos que la $(0, 1)$ -matriz A es de tal género que todas las sumas de las columnas suyas tomadas de dos en dos son diferentes. Por cuanto se pueden comparar en total $\binom{n}{2}$ sumas de las columnas tomadas de dos en dos y cada una de dicha suma es una palabra de longitud N sobre el alfabeto $\{0, 1, 2\}$, donde N es el número de filas en A , entonces resulta válida la desigualdad

$$\binom{n}{2} \leq 3^N$$

Al poner $N = N_{\text{exp}}^{\text{est}}$, obtendremos la afirmación del teorema.

En adelante demostraremos una estimación más fuerte para $N_{\text{exp}}^{\text{est}}$, y por ahora pasaremos a las estimaciones superiores para $N_{\text{exp}}^{\text{est}}$.

Con el fin de obtener estimaciones superiores hagamos uso del método de construcción aleatoria de la matriz del plan estático. Supongamos que la matriz $A = [a_{ij}]$ está construida del modo siguiente:

$$P\{a_{ij} = 1\} = p \geq 0, P\{a_{ij} = 0\} = q \geq 0, p + q = 1;$$

$$(i = 1, \dots, N; j = 1, \dots, n),$$

y que todas las magnitudes aleatorias a_{ij} son independientes en su totalidad.

Veamos un vector columna $\vec{z} = (z_1, \dots, z_n)^T$ de longitud n , compuesto de 2 unidades y $n - 2$ ceros. Este vector se denominará *vector de desarreglo*. En términos de los vectores de desarreglo el lema 2 puede enunciarse del modo siguiente.

Lema 2'. La matriz $A = [a_{ij}]$ es el plan estático, cuando y sólo cuando para cualquier par de diferentes vectores de desarreglo \vec{z} y \vec{v} se verifica la desigualdad: $A\vec{z} \neq A\vec{v}$.

La matriz A se multiplica por el vector de un modo habitual.

Veamos dos vectores de desarreglo \vec{z} y \vec{v} tales que ambas unidades del vector \vec{z} y las unidades del vector \vec{v} ocupan cuatro posiciones diferentes. Calculemos la probabilidad de que $A\vec{z} = A\vec{v}$. Calculemos con anterioridad cuál es la probabilidad de que las i -ésimas componentes de las columnas $A\vec{z}$ y $A\vec{v}$ sean iguales ($i = 1, \dots, N$).

Revisando todos los casos posibles, obtenemos

$$P\{(A\bar{z})_i = (A\bar{v})_i\} = p^4 + 4p^2q^2 + q^4.$$

Por cuanto los elementos de la matriz son magnitudes aleatorias independientes, entonces

$$P\{A\bar{z} = A\bar{v}\} = (p^4 + 4p^2q^2 + q^4)^N.$$

Veamos ahora un par de vectores de desarreglo \bar{u} y \bar{w} , en los cuales una componente unitaria coincide, y las segundas son diferentes. Por analogía con el primer caso obtendremos

$$p\{A\bar{u} = A\bar{w}\} = (p^2 + q^2)^N.$$

Estimemos ahora la probabilidad de existencia de los vectores \bar{z}_1 y \bar{z}_2 tales que $A\bar{z}_1 = A\bar{z}_2$:

$$P\{\exists \bar{z}_1, \bar{z}_2: A\bar{z}_1 = A\bar{z}_2\} \leq M_1(p^4 + 4p^2q^2 + q^4)^N + M_2(p^2 + q^2)^N,$$

donde M_1 es el número de pares de vectores de primer tipo, y M_2 , el número de pares de vectores de segundo tipo.

Los razonamientos combinatorios sencillos conducen a las siguientes igualdades:

$$M_1 = 3 \binom{n}{4}, \quad M_2 = 3 \binom{n}{3}.$$

De este modo,

$$P\{\exists \bar{z}_1, \bar{z}_2: A\bar{z}_1 = A\bar{z}_2\} \leq 3 \binom{n}{4} (p^4 + 4p^2q^2 + q^4)^N + 3 \binom{n}{3} (p^2 + q^2)^N.$$

Supongamos que para ciertos p y q ($p, q \geq 0, p + q = 1$) el número $\bar{N} = \bar{N}(n)$ es tal que

$$3 \binom{n}{4} (p^4 + 4p^2q^2 + q^4)^{\bar{N}} + 3 \binom{n}{3} (p^2 + q^2)^{\bar{N}} < 1. \quad (2)$$

De esto se deducirá que con la probabilidad superior a cero existe una matriz A de dimensión $\bar{N} \times n$, para la cual la desigualdad

$$A\bar{z} \neq A\bar{v}$$

se verifica para cualesquiera dos diferentes vectores de desarreglo \bar{z} y \bar{v} . Así pues, la matriz A corresponderá al plan estático.

Por consiguiente, $N_{\text{est}}^{\text{est}} \leq \bar{N}$.

Como probabilidades p y q podemos elegir cualesquiera números no negativos tales que $p + q = 1$. Sin embargo, resulta más ventajoso tomar tales p y q con los que las expresiones $p^4 + 4p^2q^2 + q^4$ y $p^2 + q^2$ sean las más pequeñas posibles. Las búsquedas no complejas de los valores extremos conducen a los siguientes resultados:

$$\min_{\substack{p, q \geq 0 \\ p + q = 1}} (p^4 + 4p^2q^2 + q^4) = \frac{3}{8} \quad \text{y se consigue para } p = q = \frac{1}{2}.$$

$\min_{\substack{p, q \geq 0 \\ p+q=1}} (p^2 + q^2) = \frac{1}{2}$ y se consigue para $p = q = \frac{1}{2}$.

Pongamos, por eso, $p = q = \frac{1}{2}$ y reescribamos la relación (2):

$$3 \binom{n}{4} \binom{3}{8}^N + 3 \binom{n}{3} \binom{1}{2}^N < 1. \quad (3)$$

Si se verifican simultáneamente dos desigualdades:

$$\begin{aligned} 3 \binom{n}{4} \binom{3}{8}^N &< \frac{1}{2}, \\ 3 \binom{n}{3} \binom{1}{2}^N &< \frac{1}{2}, \end{aligned} \quad (4)$$

la desigualdad (3) también se cumplirá. No es difícil comprobar que las desigualdades (4) son válidas para $N = \lceil 3 \log_2 n \rceil$. De este modo, hemos demostrado el teorema siguiente.

Teorema 7. En el caso de dos elementos defectuosos

$$N_{\text{exp}}^{\text{est}} \leq \lceil 3 \log_2 n \rceil.$$

Observación. El resultado aducido se ha obtenido de una manera que no puede considerarse constructiva, pues no hemos construido una matriz concreta del plan estático de dimensión $\lceil 3 \log_2 n \rceil \times n$, sino que sólo demostramos la existencia de la misma.

Ha de notarse, sin embargo, que en dicho problema la victoria queda por el lado de los "constructivistas". El plan estático con el número de filas $2 \cdot \lceil \log_2 n \rceil$ (véase [83]), construido con ayuda de los códigos de Bose nos da la siguiente estimación:

$$N_{\text{exp}}^{\text{est}} \leq 2 \lceil \log_2 n \rceil.$$

No obstante, para s grandes de construcción aleatoria de una matriz del plan estadístico proporciona resultados considerablemente mejores que el método de los "constructivistas" (véase [84]).

Empleando métodos análogos podemos obtener también estimaciones superiores de la longitud del plan estadístico para el llamado modelo disyuntivo, en el cual tenemos, como resultado de la i -ésima comprobación, sólo información referente a la presencia de por lo menos un elemento defectuoso en el conjunto Y_i .

7.3. MÉTODO DE ENTROPÍA

Este método permite mejorar la estimación (1) obtenida en el teorema 6. Recordemos previamente la definición y propiedades principales de la entropía.

Sea una distribución finita de las probabilidades $\mathcal{P} = \{p_1, \dots, p_m\}$;

$\sum_{i=1}^m p_i = 1$, $p_i \geq 0$, $i = 1, \dots, m$. Se denominará entropía de la distribución \mathcal{P} a la magnitud $H(\mathcal{P}) = - \sum_{i=1}^m p_i \log_2 p_i$. Además, si alguna p_i es nula, entonces, en lugar de un sumando indeterminado $-p_i \log_2 p_i$ se pone en la suma el cero (lo que es natural, puesto que $\lim_{x \rightarrow 0} x \log_2 x = 0$).

Examinemos las propiedades principales del concepto introducido.

Propiedad 1. Se verifica la desigualdad

$$H(\mathcal{P}) \geq 0,$$

donde la igualdad se consigue cuando y sólo cuando una de las probabilidades p_i es igual a la unidad, y las demás, a cero.

Propiedad 2. Se verifica la desigualdad:

$$H(\mathcal{P}) \leq \log_2 m, \quad (5)$$

en la cual la igualdad se consigue cuando y sólo cuando $p_1 = \dots = p_m = 1/m$.

En efecto,

$$H(\mathcal{P}) - \log_2 m = \sum_{i=1}^m p_i \log_2 (1/p_i m) \leq \log_2 e \sum_{i=1}^m p_i \left(\frac{1}{p_i m} - 1 \right) = 0.$$

Aquí se ha aprovechado la desigualdad $\log_2 x \leq (x-1)\log_2 e$, en la cual el signo de igualdad se consigue sólo para $x = 1$. Por consiguiente, la igualdad en (5) tiene lugar cuando y sólo cuando $p_1 = \dots = p_m = 1/m$.

Veamos la cuestión sobre la entropía de la distribución compleja de las probabilidades $\mathcal{P} = \{p_{ij}\}$, $i = 1, \dots, m_1$; $j = 1, \dots, m_2$; $p_{ij} \geq 0$, $\sum_{i,j} p_{ij} = 1$. Formemos dos distribuciones simples de las probabilidades \mathcal{P}_1 y \mathcal{P}_2 del modo siguiente:

$$\mathcal{P}_1 = \{p_i^{(1)}\}, p_i^{(1)} = \sum_{j=1}^{m_2} p_{ij}, i = 1, \dots, m_1;$$

$$\mathcal{P}_2 = \{p_j^{(2)}\}, p_j^{(2)} = \sum_{i=1}^{m_1} p_{ij}, j = 1, \dots, m_2;$$

Está claro que \mathcal{P}_1 y \mathcal{P}_2 son realmente distribuciones de las probabilidades, es decir, $p_i^{(1)} \geq 0$, $p_j^{(2)} \geq 0$, $i = 1, \dots, m_1$; $j = 1, \dots, m_2$; $\sum_{i=1}^{m_1} p_i^{(1)} = 1$, $\sum_{j=1}^{m_2} p_j^{(2)} = 1$. La propiedad siguiente se refiere a la relación existente entre

las magnitudes $H(\mathcal{P})$, $H(\mathcal{P}_1)$ y $H(\mathcal{P}_2)$.

Propiedad 3 (subaditividad de la entropía). Se verifica la desigualdad:

$$H(\mathcal{P}) \leq H(\mathcal{P}_1) + H(\mathcal{P}_2),$$

donde el signo de igualdad tiene lugar cuando y sólo cuando las distribuciones \mathcal{A}_1 y \mathcal{A}_2 son independientes, es decir, para cualesquiera i, j se cumple la relación: $p_{ij} = p_i \cdot p_j$.

En efecto, $H(\mathcal{A}) - H(\mathcal{A}_1) - H(\mathcal{A}_2) = \sum_{i,j} p_{ij} \log_2 \frac{p_i^{(1)} p_j^{(2)}}{p_{ij}}$. Aplican-

do la desigualdad $\log_2 x \leq (x - 1) \log_2 e$, obtenemos para cada término de la suma en el segundo miembro la siguiente estimación:

$$p_{ij} \log_2 \frac{p_i^{(1)} p_j^{(2)}}{p_{ij}} \leq p_{ij} \left(\frac{p_i^{(1)} p_j^{(2)}}{p_{ij}} - 1 \right) \log_2 e,$$

con la particularidad de que la igualdad tiene lugar cuando y sólo cuando $p_{ij} = p_i^{(1)} \cdot p_j^{(2)}$. Por consiguiente,

$$H(\mathcal{A}) - H(\mathcal{A}_1) - H(\mathcal{A}_2) \leq \log_2 e \left(- \sum_{i,j} p_{ij} + \sum_{i,j} p_i^{(1)} p_j^{(2)} \right) = 0,$$

lo que se trataba de demostrar.

Hemos de notar que la propiedad de subaditividad de la entropía puede extenderse con facilidad a un número arbitrario de distribuciones $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$.

Ejemplo. Sea $1 \leq k \leq n/2$, $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$.

Demostremos que $\sum_{i=0}^k \binom{n}{i} \leq 2^{nh(k/n)}$.

Hagamos $p = k/n$. Tenemos $\binom{n-1}{i} \binom{n}{i+1}^{-1} = \frac{i+1}{n} \leq p$

para $i = 0, \dots, k-1$. Por consiguiente, $\bar{p} = \left(\sum_{i=0}^{k-1} \binom{n-1}{i} \right) \left(\sum_{i=0}^k \binom{n}{i} \right)^{-1} < p$. Sea $i_j \in [0, 1], j = 1, \dots, n$,

$$p_{i_1, \dots, i_n} = \begin{cases} 0, & \text{si } \sum_{j=1}^n i_j > k, \\ \left(\sum_{i=0}^n \binom{n}{i} \right)^{-1}, & \text{si } \sum_{j=1}^n i_j \leq k, \end{cases}$$

$$p_{i_1, \dots, i_n} = \begin{cases} \bar{p}, & \text{si } i_j = 1 \\ 1 - \bar{p}, & \text{si } i_j = 0, \text{ así que} \end{cases}$$

$$p_{i_1, \dots, i_n} = \sum_{i_1, \dots, i_n} p_{i_1, \dots, i_n}$$

Aprovechando la propiedad de subaditividad de la entropía, obtenemos

$$\begin{aligned} \log_2 \sum_{i=0}^k \binom{n}{i} &= - \sum_{i_1, \dots, i_n} p_{i_1, \dots, i_n} \log_2 p_{i_1, \dots, i_n} \\ &= H(\{p_{i_1, \dots, i_n}\}) \leq \sum_{j=1}^n H(\{p_{i_1, \dots, i_{j-1}, i_{j+1}, \dots, i_n}\}) = \\ &= nh(\bar{p}) < nh(p) \quad \left(\text{puesto que } \bar{p} < p \leq \frac{1}{2} \right), \end{aligned}$$

lo que se requería demostrar.

Volvamos al problema de experimentos de rechazo. Examinemos el caso de un modelo lineal del plan estático en que se tienen dos elementos defectuosos.

Teorema 8.

$$N_{\text{exp}}^{\text{est}} \geq (1 + o(1)) \cdot \frac{3}{4} \log_2 n.$$

Demostración. Sea \bar{z} un vector columna de desarreglo que contiene dos unidades y $(n-2)$ ceros, y sea \bar{e} una columna de resultados de los experimentos, mientras que A representa una matriz del plan estático. Entonces

$$\bar{e} = A\bar{z}.$$

Introduzcamos en el conjunto Ω de vectores de desarreglo una distribución uniforme \mathcal{P} : $P(\bar{z}) = \binom{n}{2}^{-1}$. Siendo la aplicación $A: \Omega \rightarrow \mathcal{A}(\Omega)$ biunívoca, en el conjunto $\mathcal{A}(\Omega)$ se induce la distribución uniforme de probabilidades Q :

$$P(\bar{e}) = \binom{n}{2}^{-1} \quad \text{y} \quad H(Q) = H(\mathcal{P}) = \log_2 \binom{n}{2}.$$

Supongamos que en la i -ésima fila de la matriz A se tienen $n_0^{(i)}$ ceros y $n_1^{(i)}$ unidades. La distribución \mathcal{P} induce en el conjunto $\{0, 1, 2\}^N$ distribuciones Q_i que se determinan del modo siguiente:

$$P_0^{(i)} = P\{e_i = 0\} = \binom{n_0^{(i)}}{2} \binom{n}{2}^{-1} = \frac{n_0^{(i)} \cdot (n_0^{(i)} - 1)}{n \cdot (n-1)},$$

$$P_1^{(i)} = P\{e_i = 1\} = \binom{n_0^{(i)}}{1} \binom{n_1^{(i)}}{1} \cdot \binom{n}{2}^{-1} = \frac{2n_0^{(i)} \cdot n_1^{(i)}}{n \cdot (n-1)},$$

$$P_2^{(i)} = P\{e_i = 2\} = \binom{n_1^{(i)}}{2} \cdot \binom{n}{2}^{-1} = \frac{n_1^{(i)} \cdot (n_1^{(i)} - 1)}{n \cdot (n-1)},$$

Está claro que $P_i^{(j)} = \sum_{\bar{z} \in A^{-1}(e_i)} P(\bar{z})$. Recurriendo a la propiedad de subaditividad de la entropía, obtenemos: $\log_2 \binom{n}{2} = H(Q) \leq \sum_{i=1}^N H(Q_i)$. Nos

convencemos inmediatamente de que $H(Q_i) = \varphi(x_i) \cdot (1 + o(1))$, donde $x_i = \frac{n_0^{(i)}}{n}$, $n \rightarrow \infty$, y $\varphi(x) = -x^2 \log_2 x^2 - 2x(1-x) \log_2 x(1-x) - (1-x)^2 \log_2 (1-x)^2$.

Por cuanto

$$\max_{x \in [0,1]} \varphi(x) = \varphi\left(\frac{1}{2}\right) = \frac{3}{2} \quad \text{y} \quad \sum_{i=1}^N H(Q_i) \leq (1 + o(1))N \max_{x \in [0,1]} \varphi(x),$$

entonces

$$N \geq \frac{\log_2 \binom{n}{2}}{\max_{x \in [0,1]} \varphi(x)} \cdot (1 + o(1)) = (1 + o(1)) \cdot \frac{4}{3} \log_2 n.$$

El teorema está demostrado.

Haciendo uso de las ideas análogas (pero en otro lenguaje matemático) y aplicando adicionalmente razonamientos combinatorios más finos, Lindström [85] obtuvo una estimación más fuerte:

$$N_{\text{exp}}^{\text{est}} \geq (1 + o(1)) \cdot \frac{5}{3} \log_2 n.$$

7.4. MÉTODO DE BALANCE ALEATORIO

En los párrafos antecedentes se analizaba el modelo matemático del problema de experimentos de rechazo, en el que se necesitaba hallar s elementos fijos, pero desconocidos de antemano, en un conjunto X compuesto de n elementos. Ahora suponemos que a todo elemento $x \in X$ se le ha hecho corresponder cierto número real $\omega(x)$, llamado *peso* del elemento citado, y entre los elementos del conjunto X sólo algunos, que no se conocen de antemano y cuya cantidad es igual a $s \geq 2$, poseen peso no nulo. Se requiere separar estos s elementos del conjunto X y determinar los pesos de ellos. Se permite en cada experimento elegir cierto subconjunto $Y \subseteq X$ y determinar la suma de pesos de los elementos pertenecientes a Y (es decir, "ponderar" los elementos de Y). Hace falta planificar los experimentos de una manera tal que a base de sus resultados se puedan determinar los pesos $\omega_j = \omega(x_j)$ de los elementos, pertenecientes a X ($j = 1, 2, \dots, n$).

En lo que sigue nos adheriremos a la siguiente terminología: los pesos $\omega_1, \dots, \omega_n$ de los elementos x_1, \dots, x_n se llamarán *factores*: el factor ω_j se denominará *significativo*, si $\omega_j \neq 0$. El número s de factores significativos se supone pequeño en comparación con el número total de factores n .

Partiendo de la reunión de los subconjuntos $Y_i \subseteq X$ ($i = 1, \dots, N$) construyamos una $(N \times n)$ -matriz de los experimentos $A = [a_{ij}]$, donde $a_{ij} = 1$, siempre que $x_j \in Y_i$, y $a_{ij} = 0$, en el caso contrario. Así pues, en cada i -ésimo experimento ($i = 1, \dots, N$) se halla la magnitud $e_i = \sum_{j=1}^n a_{ij} \omega_j$.

La matriz de los experimentos $A = [a_{ij}]$ se denomina, además, plan de los experimentos. El plan ha de satisfacer los siguientes requisitos principales: sencillez en la realización y economía. El cumplimiento de las condiciones mencionadas conduce a la necesidad de minimizar el número N de experimentos en el plan.

En el párrafo anterior se ha considerado el caso cuando $s = 2$ y todos los factores significativos son iguales a uno. En cambio, si el número s de factores significativos es arbitrario, no será difícil mostrar que para su separación unívoca del número total de factores n harán falta no menos de $\log_{s+1} \binom{n}{s} \sim \frac{s}{\log_2 s} \log_2 n$ experimentos ($s \leq n$). L. D. Meshalkin fue el primero en acentuar que los factores significativos son diferentes en la práctica e, incluso, poseen cierto carácter de inconmensurabilidad (véase [86]). Esta circunstancia hace posible construir los planes que permiten determinar con eficacia dichos factores en el transcurso de $\log_2 n$ experimentos, lo que es considerablemente inferior a la estimación inferior aducida.

Definición. Los factores $\omega_1, \dots, \omega_n$ se llaman *inconmensurables* sobre el conjunto $\Lambda \subseteq \mathbb{Z}$ (donde $0 \in \Lambda$), si de la condición $\sum_{j=1}^n \lambda_j \omega_j = 0$, donde $\lambda_j \in \Lambda$, se deduce que $\lambda_j \omega_j = 0$ para cualesquiera $j = 1, \dots, n$, es decir, $\lambda_j = 0$ para todos los factores significativos ω_j .

El problema planteado se resolverá para $\Lambda = \mathbb{Z}$, por el método de balance aleatorio. He aquí el esquema que debe seguirse. Se elabora un algoritmo, para el cual de los datos de partida sirve una $(0, 1)$ -matriz $A = [a_{ij}]$ que contiene N filas y n columnas, y, además, un vector columna $\vec{e} = (e_1, \dots, e_N)^T \in \mathbb{R}^N$ (T es la operación de transposición). Si la matriz y el vector \vec{e} poseen ciertas propiedades, las cuales trataremos más abajo, el algoritmo elabora el vector $\vec{\omega}' \in \mathbb{R}^n$; de lo contrario el algoritmo no nos da nada. Este algoritmo resulta ser de tal índole que de $\vec{e} = A\vec{\omega}$ se deduce: $\vec{\omega}' = \vec{\omega}$. Se demuestra que para todo $\beta \in]0, 1[$ se puede indicar tal número $N = N(n, s, \beta)$ que la matriz A , construida al azar, y el vector $\vec{e} = A\vec{\omega}$ posean las propiedades indispensables con la probabilidad no inferior a $1 - \beta$.

Ha de indicarse que el número s de factores significativos no es obligatoriamente conocido de antemano y puede determinarse en el transcurso de los experimentos.

Teorema 9. Supongamos que los factores $\omega_1, \dots, \omega_n$ son inconmensurables sobre el conjunto \mathbb{Z} , y el plan de los experimentos $A = [a_{ij}]$ se construye de un modo aleatorio:

$$P\{a_{ij} = 0\} = P\{a_{ij} = 1\} = \frac{1}{2}; \quad i = 1, \dots, N; \quad j = 1, \dots, n.$$

Entonces, si s es el número desconocido de factores significativos, y β , un número real, con la particularidad de que $0 < \beta < 1$, entonces para

$$N > s + \log_2(n - s + 1) - \log_2 \beta$$

podemos determinar, con la probabilidad no inferior a $1 - \beta$, todos los factores significativos y su número s . En este caso el error en la determinación de los factores está excluido.

Observemos que si el número s de factores significativos no sobrepasa cierto número conocido de antemano k : $s \leq k \ll n$, el teorema citado fija la frontera para el número de experimentos que han de ser realizados, para que el plan aleatorio construido determine los factores significativos con una probabilidad no inferior a la prefijada.

Previamente a la demostración del teorema aduzcamos el siguiente lema de L. D. Meshalkin.

Lema 3. Cualquier plano l -dimensional en un espacio real n -dimensional ($2 \leq n$; $l \leq n$) contiene no más de 2^l vértices de un cubo n -dimensional.

La demostración se realizará por inducción respecto de n . Cuando $n = 2$ puede igualarse a 1 ó a 2, y la afirmación del lema resulta evidente. Supongamos que dicha afirmación ya está demostrada para todos los números $n' < n$, donde $n \geq 3$. Demostremosla también para $n' = n$.

Elijamos un sistema de coordenadas en el espacio R^n de un modo tal que el cubo n -dimensional dado se haga unitario, es decir, que las coordenadas de sus vértices sean iguales a 0 ó 1. Está claro que si $l = n$, cada plano l -dimensional puede contener no más de 2^l vértices del cubo (número total de todos los vértices suyos). Si $l = 1$, entonces la recta $\bar{x}(t) = \bar{a} + (\bar{b} - \bar{a})t$, que pasa por diferentes vértices $\bar{a} = (a_1, \dots, a_n)^T$ y $\bar{b} = (b_1, \dots, b_n)^T$ ($a_i, b_i \in [0, 1]$, $i = 1, \dots, n$), no puede contener ningún otro vértice más. En efecto, sea i tal que $a_i \neq b_i$. Entonces, $x_i(0) = x_i(1)$, y $x_i(0), x_i(1) \in \{0, 1\}$. Por consiguiente, para todos los $t \neq 0, 1$:

$$x_i(t) = a_i + b_i - a_i)t \notin \{0, 1\}.$$

Sea ahora l tal que $1 < l < n$. Denotemos $R_i^n = \{\bar{x} \in R^n : x_n = i\}$, $i = 0, 1$. Si un plano l -dimensional L es paralelo a R_0^n o a R_1^n y no yace en estos últimos, entonces el plano no contiene ningún vértice del cubo. En cambio, si L está situado en uno de los hiperplanos R_0^n ó R_1^n , la afirmación del lema se cumplirá por la hipótesis de la inducción. Resta analizar un caso en que L interseca ambos hiperplanos. Designemos $L_i = L \cap R_i^n$, $i = 0, 1$. Está claro que las dimensiones de los planos L_0 y L_1 son iguales a $l - 1$. Los vértices del cubo, situados en cada uno de los hiperplanos R_i^n ($i = 0, 1$), forman un cubo $(n - 1)$ -dimensional y, por consiguiente, por hipótesis de la inducción, el plano L_i contiene 2^{l-1} tales vértices a lo sumo. Por eso, el plano L contiene no más de $2^{l-1} + 2^{l-1} = 2^l$ vértices del cubo n -dimensional dado. El lema queda demostrado.

Demostración del teorema 9. Introduzcamos las siguientes designaciones: \bar{a}_j es la j -ésima columna de la matriz $A = |a_{ij}|$; $L(\bar{a}_1, \dots, \bar{a}_m)$ es la cápsula lineal de los vectores $\bar{a}_1, \dots, \bar{a}_m$. Diremos que la matriz A y el vector \bar{e} poseen la propiedad S , si en A existen columnas linealmente

independientes $\vec{a}_{j_1}, \dots, \vec{a}_{j_m}$, cuya cápsula lineal no contiene ninguna otra columna y el sistema de ecuaciones

$$\vec{e} = \sum_{r=1}^m \vec{a}_{j_r} x_r \quad (6)$$

tiene una única solución. Si la matriz A y el vector \vec{e} poseen la propiedad S , entonces, al resolver el sistema (6), formamos un vector $\vec{\omega}' = (\omega'_1, \dots, \omega'_m)^T$, haciendo

$$\omega'_j = \begin{cases} \bar{x}_r & \text{para } j = j_r \ (r = 1, \dots, m), \\ 0 & \text{para } j \notin \{j_1, \dots, j_m\}, \end{cases}$$

donde $(\bar{x}_1, \dots, \bar{x}_m)$ es la solución del sistema (6). Observemos que si el juego de números de las columnas j_1, \dots, j_m en la propiedad S se define de una manera no unívoca, entonces puede elegirse arbitrariamente. Por consiguiente, el vector construido $\vec{\omega}'$ es precisamente vector buscado $\vec{\omega}$, si $\vec{e} = A\vec{\omega}$. Para demostrar esto, basta comprobar que los números de todos los factores significativos ω_j pertenecen al conjunto $\{j_1, \dots, j_m\}$. En efecto, si esto no es así, se encontraría un juego $\{\alpha_1, \dots, \alpha_l\}$ de números de las columnas que no está íntegramente contenido en el conjunto citado y que satisface el sistema de ecuaciones:

$$\vec{e} = \sum_{r=1}^l a_{\alpha_r} \omega_{\alpha_r},$$

donde $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$ son todos factores significativos.

Veamos un sistema de ecuaciones

$$\sum_{r=1}^m \vec{a}_{j_r} x_r - \sum_{r=1}^l \vec{a}_{\alpha_r} y_r = 0 \quad (7)$$

con las incógnitas $x_1, \dots, x_m, y_1, \dots, y_l$. De la definición de la propiedad S y de la elección del juego $\{\alpha_1, \dots, \alpha_l\} \not\subseteq \{j_1, \dots, j_m\}$ se deduce que el rango del sistema no es inferior a $m+1$. Por consiguiente, haciendo uso del procedimiento de Gauss, podemos excluir de dicho sistema de ecuaciones las variables x_1, \dots, x_m y obtener por lo menos una ecuación que contenga sólo las variables y_r ($r = 1, \dots, l$):

$$0 = \sum_{r=1}^l b_r y_r, \quad \sum_{r=1}^l b_r^2 \neq 0. \quad (8)$$

Los coeficientes de las incógnitas en el sistema (7) eran enteros, razón por la cual pueden considerarse enteros también los coeficientes b_r en (8). El sistema de ecuaciones (7) tiene la solución: $x_r = \bar{x}_r$, $r = 1, \dots, m$; $y_r = \omega_{\alpha_r}$, $r = 1, \dots, l$. Por eso, el juego $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$ forma la solución de la ecuación (8), lo que contradice la condición de inconmensurabilidad de los factores significativos $\omega_{\alpha_1}, \dots, \omega_{\alpha_l}$.

Se ha mostrado, pues, que si la matriz A y el vector \vec{e} poseen la propiedad S , entonces el vector desconocido $\vec{\omega}$ queda unívocamente definido. Estimemos la probabilidad P de que construyendo al azar un plan con el

número de experimentos igual a $N > s + \log_2(n - s + 1) - \log_2 \beta$, el par (A, \vec{e}) poseerá la propiedad S , donde $\vec{e} = A\vec{\omega}$.

Designemos mediante j_1, \dots, j_s los números desconocidos de los factores significativos. Estudiemos los sucesos:

$$\begin{aligned} X &= \{ \dim L(\vec{a}_{j_1}, \dots, \vec{a}_{j_s}) = s \}, \\ Y_j &= \{ \vec{a}_{j_i} \notin L(\vec{a}_{j_1}, \dots, \vec{a}_{j_s}) \}, j \in \{ j_1, \dots, j_s \}, \\ Y &= \bigcap_{j \in \{j_1, \dots, j_s\}} Y_j, Z = X \cap Y. \end{aligned}$$

Está claro que la aparición del suceso Z lleva consigo el que el par (A, \vec{e}) , donde $\vec{e} = A\vec{\omega}$, posee la propiedad S . Por consiguiente, $P \geq P(Z)$. Estimemos la probabilidad del suceso Z . Tenemos

$$P(Z) = P(X \cap Y) = P(X) \cdot P(Y|X).$$

Ciertos razonamientos combinatorios no complejos con la aplicación del lema 3 nos conducen a una desigualdad

$$P(Y|X) \geq (1 - 2^{-N+s})^{n-s}.$$

Para poder estimar la probabilidad del suceso X analicemos los siguientes sucesos:

$$\begin{aligned} X_1 &= \{ \dim L(\vec{a}_{j_1}) = 1 \}, X_2 = \{ \dim L(\vec{a}_{j_1}, \vec{a}_{j_2}) = 2 \}, \dots, \\ X_{s-1} &= \{ \dim L(\vec{a}_{j_1}, \dots, \vec{a}_{j_{s-1}}) = s-1 \}, \\ \bar{X}_1 &= \{ \vec{a}_{j_1} \notin L(\vec{a}_{j_1}) \}, \bar{X}_2 = \{ \vec{a}_{j_2} \notin L(\vec{a}_{j_1}, \vec{a}_{j_2}) \}, \dots, \\ \bar{X}_{s-1} &= \{ \vec{a}_{j_{s-1}} \notin L(\vec{a}_{j_1}, \dots, \vec{a}_{j_{s-1}}) \}. \end{aligned}$$

Al hacer uso del lema 3, no es difícil notar que

$$P(X_1) = 1 - 2^{-N}, P(\bar{X}_r | X_r) \geq 1 - 2^{-N+r}, r = 1, \dots, s-1.$$

Por eso, $P(X) = P(X_1) \prod_{r=1}^{s-1} P(\bar{X}_r | X_r) \geq \prod_{r=0}^{s-1} (1 - 2^{-N+r}) \geq 1 - 2^{-N+s}$

para $s \geq 2$. La demostración de la última desigualdad en esta cadena puede fácilmente obtenerse por inducción respecto de s . Así pues, $P(Z) \geq (1 - 2^{-N+s})^{n-s+1} > 1 - (n-s+1)2^{-N+s}$ para $N > s$, puesto que para $0 < \alpha < 1$ es válida la desigualdad: $(1 - \alpha)^m > 1 - m\alpha$. Por consiguiente, para $N > s + \log_2(n - s + 1) - \log_2 \beta$ se verifica la desigualdad: $P(Z) > 1 - (n - s + 1)2^{-N+s} > 1 - \beta$, lo que finaliza la demostración del teorema.

M. B. Maljútov y M. S. Pínsker [87] han estudiado el problema aducido para las condiciones más generales, cuando los factores $\omega_1, \dots, \omega_n$ se suponen incommensurables sobre el conjunto $\Lambda = [-1, 0, 1]$ (los experimentos se realizan en este caso no de un modo aleatorio, sino sucesivamente, o sea, tomando en consideración los resultados anteriores. Con ellos se ha demostrado el teorema siguiente.

Teorema 10. Supongamos que los factores $\omega_1, \dots, \omega_n$ son inconmen-

surables sobre el conjunto $\Lambda = \{-1, 0, 1\}$; y sea s un número desconocido de factores significativos. Existe, entonces, un método no aleatorio de realizar los experimentos que permite determinar ciertamente todos los factores en el transcurso de no más de $s \log_2 s + \log_2 n$ experimentos. El número s se halla en este caso en el proceso de los experimentos.

Demostración. Con ayuda del primer experimento determinamos $e_1 = \sum_{j=1}^n \omega_j$. Los experimentos ulteriores se realizan de un modo inductivo.

Supongamos que tras cierta cantidad de experimentos el conjunto de todos los factores haya resultado partido en $r_m + 1$ subconjuntos G_i^m , $i = 0, 1, \dots, r_m$ ($1 \leq r_m \leq s$), con la particularidad de que G_0^m no contiene factores significativos, mientras que todos los G_i^m restantes contienen cada uno por lo menos un factor significativo, y se conocen las sumas $\sigma_i^m = \sum_{\omega \in G_i^m} \omega_j$,

$i = 1, \dots, r_m$. En el experimento siguiente partiremos cada uno de los subconjuntos G_i^m ($i = 1, \dots, r_m$) en partes aproximadamente iguales: $G_i^m = G_{i0}^{m+1} \cup G_{i1}^{m+1}$, $G_{i0}^{m+1} \cap G_{i1}^{m+1} = \emptyset$, $|G_{i0}^{m+1}| - |G_{i1}^{m+1}| \leq 1$, y hallaremos la suma $c = \sum_{i=1}^{r_m} \sum_{\omega \in G_{i0}^{m+1}} \omega_j$. Si $e \in \left\{ \sum_{i=1}^{r_m} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$, entonces, haciendo uso de

la condición de inconmensurabilidad de los factores, llegamos a la conclusión de que cada uno de los subconjuntos $G_{i\varepsilon}^{m+1}$ contiene sólo factores insignificativos. Entonces, al elegir $G_i^{m+1} = G_{i\varepsilon}^{m+1}$, $\sigma_i^{m+1} = \sigma_i^m$ ($i = 1, \dots, r_m$), $r_{m+1} = r_m$, $G_0^{m+1} = \left(\bigcup_{i=1}^{r_m} G_{i\varepsilon}^{m+1} \right) \cup G_0^m$, donde $\bar{\varepsilon}_i = (\varepsilon_i + 1) \bmod 2$,

continuaremos los experimentos según el esquema citado hasta que queden determinados todos los factores significativos, es decir, hasta que r_m no se haga igual a s , y cada uno de los subconjuntos G_i^m no se componga de un único factor significativo. En cambio, si $e \in \left\{ \sum_{i=1}^{r_m} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$,

tendremos que realizar un ciclo de experimentos para determinar las sumas $\sigma_{i\varepsilon}^m = \sum_{j \in G_{i\varepsilon}^m} \omega_j$ de los factores significativos en todas las mitades $G_{i\varepsilon}^m$,

$i = 1, \dots, r_m$; $\varepsilon = 0, 1$. Con este fin realicemos un experimento y determinemos la suma $\sum_{i=1}^{\lfloor r_m/2 \rfloor} \sigma_{i0}^m$. Aclaremos si dicha suma pertenece al conjunto

$\left\{ \sum_{i=1}^{\lfloor r_m/2 \rfloor} \varepsilon_i \sigma_i^m \mid \varepsilon_i \in \{0, 1\} \right\}$. Si la respuesta resulta positiva, llegamos a una conclusión de que cada uno de los subconjuntos $G_{i\varepsilon}$, $i = 1, \dots, \left\lfloor \frac{r_m}{2} \right\rfloor$

contiene sólo factores insignificativos. De lo contrario, realizamos un experimento más y determinamos la suma $\sum_{i=1}^{\lfloor r_m/4 \rfloor} \sigma_{i0}^m$, etc. Indiquemos que en este

caso las sumas complementarias, tales como $\sum_{i \in [r_m/2]+1}^{r_m} \sigma_{i0}^m$, se determinan automáticamente y en adelante han de tratarse análogamente, dividiendo, si es necesario, el conjunto en el que viene definido el índice de sumación, aproximadamente por la mitad. Así pues, como resultado de tal ciclo de experimentos, determinamos todas las sumas $\sigma_{ic}^m = \sum_{j \in Q_j} w_j$, $i = 1, \dots, r_m$; $\varepsilon = 0, 1$, y aclaramos cuáles de los subconjuntos G_{ic}^m se componen íntegramente sólo de los factores significativos. Al relacionar todos estos subconjuntos, junto con los subconjuntos G_0^m , al subconjunto G_0^{m+1} y al reenumerar todos los subconjuntos nuevamente mediante los símbolos $G_1^{m+1}, \dots, G_{r_{m+1}}^{m+1}$, donde $r_{m+1} > r_m$, llegamos a la misma situación que teníamos al principio de nuestros razonamientos.

Estimemos el número de experimentos del plan construido sucesivamente. No es difícil notar que en cada ciclo el número de experimentos no sobrepasa de $(r_{m+1} - r_m)(1 + \log_2 r_m)$. Cada uno de los demás experimentos que no integran los ciclos disminuye el número de factores, que se supone son significativos, en dos veces (con una corrección evidente por el hecho de que este número puede no dividirse por 2). Por consiguiente, el número total N de experimentos del plan no sobrepasa $s \log_2 s + \log_2 n$:

$$N \leq 1 + \log_2 n + \sum_m (r_{m+1} - r_m)(1 + \log_2 r_m) \leq 1 + \log_2 n + \sum_m \sum_{i: r_m \leq i < r_{m+1}} (1 + \log_2 i) = 1 + \log_2 n + \sum_{i=1}^{s-1} (1 + \log_2 i)$$

(puesto que $1 = r_1 < r_2 < \dots < r_m < r_{m+1} < \dots < r_p = s$).

Por consiguiente,

$$N \leq s + \log_2 n + \sum_{i=2}^{s+1} \log_2 i \leq s + \log_2 n + \int_2^s \log_2 x dx \leq s \log_2 s + \log_2 n.$$

El teorema queda demostrado.

7.5. SISTEMAS SEPARADORES DE SUBCONJUNTOS

Recordemos que el sistema de subconjuntos $\{Y_1, Y_2, \dots, Y_N\}$ de un conjunto finito X lleva el nombre de sistema *separador*, si en él para cualesquiera dos elementos distintos de X existe Y_i ($1 \leq i \leq N$) que contiene un solo elemento de los dos mencionados (véase § 1.5). Pongamos $n = |X|$. En este párrafo se analiza el problema de buscar un sistema de separación mínimo, es decir, un sistema compuesto de un número mínimo posible $N = N(n, k)$ de subconjuntos, a condición de que cada subconjunto contiene exactamente k elementos. Nos será de interés también un sistema separador mínimo, cada subconjunto del cual se compone de no más k

elementos. La cantidad de subconjuntos de tal sistema la designamos con $N(n, \leq k)$.

Al concepto de sistema separador llevan algunos problemas de planificación de los experimentos de rechazo (véase § 7.2). Supongamos que en nuestra disposición se tiene una balanza analítica, con ayuda de la cual se pueden determinar los pesos de los subconjuntos del conjunto X , que contiene n monedas. Se conoce que en el proceso de un solo experimento pueden pesarse simultáneamente no más de k monedas. Todas las monedas de X son iguales, a excepción de una moneda que es falsa y difiere de las demás por el peso. Se requiere ofrecer un algoritmo estático de la búsqueda, que en un número mínimo posible de experimentos revele la moneda falsa.

Convengamos en considerar, para mayor precisión, que los pesos de las monedas auténticas son iguales a 0, y el de la moneda falsa, a la unidad. Supongamos que en el i -ésimo experimento ($i = 1, \dots, N$) se determina el peso sumario de los subconjuntos de monedas $Y_i \subseteq X$, $|Y_i| \leq k$. Está claro que el sistema $\{Y_1, \dots, Y_N\}$ de los subconjuntos de X representa la estrategia estática de la búsqueda, si y sólo si este sistema es separador. El problema se ha reducido, pues, a la construcción de un sistema separador mínimo, cada subconjunto del cual contiene no más de k elementos.

Designaremos con un mismo símbolo x_j la j -ésima moneda y el peso de la misma; $j = 1, \dots, n$. Igual que antes, construyamos, a base del juego de subconjuntos $\{Y_1, \dots, Y_N\}$, una $(N \times n)$ -matriz de los experimentos $A = [a_{ij}]$, al hacer $a_{ij} = 1$, si $x_j \in Y_i$, y $a_{ij} = 0$, en el caso contrario. Observemos que cada fila de la matriz A contiene no más de k unidades, con la particularidad de que el sistema $\{Y_1, \dots, Y_N\}$ es separador, si y sólo si todas las columnas de la matriz A son diferentes. Tal $(0, 1)$ -matriz A lleva el nombre de $F(n, \leq k)$ -plan. En este caso, si todas las filas de A contienen exactamente k unidades, la matriz A se denomina $F(n, k)$ -plan. No es difícil ver que las cantidades mínimas de filas en $F(n, k)$ y, correspondientemente, en $F(n, \leq k)$ -plan son iguales a $N(n, k)$ y $N(n, \leq k)$.

Procedamos a estudiar estas magnitudes y construir los planes correspondientes.

Teorema 11. Son válidas las siguientes afirmaciones:

- 1) $N(n, k) = N(n, n - k)$,
- 2) $N(n, k) = N(n, \leq k)$ para $n \geq 2k$.

Demostración. La validez de la primera afirmación se deduce de que todo $F(n, k)$ -plan se transforma, por inversión de 0 y 1, en $F(n, n - k)$ -plan. Demostremos 2). Por cuanto todo $F(n, k)$ -plan es, a la vez, el $F(n, \leq k)$ -plan, entonces $N(n, \leq k) \leq N(n, k)$. La desigualdad inversa para $n \geq 2k$ se deduce del algoritmo que convierte el $F(n, \leq k)$ -plan en $F(n, k)$ -plan. En efecto, veamos una fila arbitraria del $F(n, \leq k)$ -plan y denotemos con m el número de unidades en ella. Consideremos las columnas del plan que

en la fila elegida contienen cero. Está claro que entre estas $n - m$ columnas pueden elegirse no menos de $(n - m) - m = n - 2m$ tales columnas que la sustitución en ellas del elemento nulo destacado por la unidad conserve intacta la propiedad de distinguibilidad de las columnas del plan. Aprovechando esta circunstancia, completemos en cada fila el número de unidades del $F(n, \leq k)$ -plan que faltan para obtener k unidades y, de este modo, transformémoslo en $F(n, k)$ -plan.

Los resultados de los siguientes dos teoremas los obtuvo el matemático húngaro Katona [37].

Teorema 12.

$$\left\lceil \frac{2(n-1)}{k+1} \right\rceil \leq N(n, \leq k) \leq \max \left\{ \left\lceil \frac{2(n-1)}{k+1} \right\rceil, k \right\}.$$

Demostración. Por definición del $F(n, \leq k)$ -plan, la cantidad total de unidades en el plan no sobrepasa de Nk . Sea n_i el número de columnas del plan que contienen i unidades. Por cuanto $n_0 \leq 1$, $n_i \leq N$, tenemos

$$Nk \geq \sum_{i=0}^N in_i \geq n_1 + 2 \sum_{i=2}^N n_i = n_1 + 2(n - n_1 - n_0) \geq 2(n-1) - N,$$

de donde se desprende precisamente la estimación inferior. Para demostrar

la estimación superior, pongamos $N = \max \left\{ \left\lceil \frac{2(n-1)}{k+1} \right\rceil, k \right\}$ y

mostremos que existe un $F(n, \leq k)$ -plan de dimensión $N \times n$.

Examinaremos diferentes columnas binarias de altura N . Dos columnas de este género, $\vec{a} = (a_0, \dots, a_{N-1})^T$ y $\vec{b} = (b_0, \dots, b_{N-1})^T$ se llamarán semejantes, si $a_i = b_{(i+j) \bmod N}$ para cierto j natural y para todo $i = 0, \dots, N-1$. El conjunto de todas las columnas se divide en clases de equivalencia con relación a la semejanza. Está claro que la potencia de cada clase de este género será un divisor del número N . Cualesquiera dos columnas semejantes contienen un número igual de unidades. Una matriz compuesta por todas las columnas de una misma clase de equivalencia se denominará bloque.

Veamos un bloque de dimensión $N \times N$, generado por una columna cuyos primeros $m < N$ lugares están ocupados por unidades, y los demás lugares, por ceros. No es difícil notar que para cualquier número natural r , $0 < r < N$, pueden elegirse r columnas de este bloque de un modo tal que el número de unidades en cualesquiera dos filas de la matriz, formada por dichas columnas, se diferenciará no más que en 1. La matriz formada por las citadas r columnas se denominará bloque incompleto.

El $F(n, \leq k)$ -plan se construirá ahora del modo siguiente. Contiene una columna compuesta sólo por ceros y N columnas, cada una de las cuales contiene un solo elemento unidad. Las demás columnas contienen cada una dos elementos unidad. Estos se disponen en el plan en bloques. El

último bloque puede ser incompleto. La cantidad total de columnas que contienen no más de dos unidades no es inferior a n :

$$1 + N + \frac{N(N-1)}{2} = 1 + \frac{N(N+1)}{2} \geq \frac{2(n-1)}{k+1} \frac{k+1}{2} = n.$$

Por consiguiente, al construir el plan, podemos en realidad servirnos sólo de tales columnas. En el plan construido la cantidad de unidades en cada fila no es superior a k . Efectivamente, la cantidad total de unidades en el plan es igual a $2n - N - 2$. Por cuanto la cantidad de unidades en cualesquiera dos filas se diferencia no más que en 1, cada fila, entonces, contiene no más de $\left\lfloor \frac{2(n-1) - N}{N} \right\rfloor \leq k$ unidades. El teorema está demostrado.

Corolario. $N(n, k) = \left\lfloor \frac{2(n-1)}{k+1} \right\rfloor$ para $n \geq \frac{k(k+1)}{2} + 1$.

Teorema 13. Se verifica la siguiente desigualdad:

$$N(n, k) \geq \frac{\log_2 n}{h(k/n)}$$

donde $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$.

Demostración. Hagamos uso de la propiedad de subaditividad de la entropía para obtener la estimación inferior de la magnitud $N(n, k)$. Supongamos que la $(N \times n)$ -matriz $A = [a_{ij}]$, donde $N = N(n, k)$, $a_{ij} \in \{0, 1\}$, es el $F(n, k)$ -plan. En este caso A aplica el conjunto B_n^1 de todas las sucesiones de longitud n , que se componen de una sola unidad y ceros, biunivocamente sobre cierto subconjunto del conjunto B_N de todas las $(0, 1)$ -sucesiones de longitud N :

$$\vec{e} = A\vec{x}, \vec{x} = (x_1, \dots, x_n)^T \in B_n^1, \vec{e} = (e_1, \dots, e_N)^T \in B_N.$$

Introduzcamos sobre el conjunto B_n^1 una distribución uniforme \mathcal{P} : $P(\vec{x}) = 1/n$, $\vec{x} \in B_n^1$. Debido a ello, sobre el conjunto $A(B_n^1) \subseteq B_N$ se induce una distribución uniforme que se designará con la misma letra \mathcal{P} : $P(\vec{e}) = 1/n$, donde $\vec{e} \in A(B_n^1)$. La entropía $H(\mathcal{P})$ de distribución de las probabilidades \mathcal{P} es igual, como es fácil de notar, a $\log_2 n$. Por otra parte, la distribución de las probabilidades \mathcal{P} en $A(B_n^1)$ forma N distribuciones simples de probabilidades \mathcal{P}_i , definidas sobre las coordenadas de los vectores de B_N : $P\{e_i = 1\} = k/n$, $P\{e_i = 0\} = 1 - k/n$, $i = 1, \dots, N$. Al calcular las últimas probabilidades se ha utilizado el hecho de que cada fila de la matriz A contiene exactamente k unidades. Al aplicar la propiedad de subaditividad de la entropía, obtenemos: $\log_2 n = H(\mathcal{P}) \leq \sum_{i=1}^N H(\mathcal{P}_i) = Nh(p)$, donde $p = k/n$, $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ es la entropía de Shannon. Por consiguiente, $\frac{\log_2 n}{h(p)} \leq N = N(n, k)$.

Corolario. Sea $n \rightarrow \infty$, $k = k(n) = \lceil pn \rceil$, donde el número p no depende de n , y $0 < p < 1/2$. Entonces resulta válida la siguiente fórmula asintótica:

$$N(n, k) = \frac{\log_2 n}{h(p)} (1 + o(1)).$$

Demostración. Construyamos un $F(n, k)$ -plan concreto de longitud $N = \frac{\log_2 n}{h(p)} (1 + o(1))$. Por cuanto cada $F(N, \leq k)$ -plan puede ser transformado, para $n \geq 2k$, en un $F(n, k)$ -plan de la misma longitud será suficiente construir el correspondiente $F(n, \leq k)$ -plan. Elijamos N mínimo natural, para el cual se verifica la desigualdad $\binom{N}{\lfloor pN \rfloor} \geq n$. Hallando por logaritmos esta desigualdad y empleando la fórmula de Stirling para los factoriales, no es difícil de concluir que para tal N es lícita la fórmula asintótica: $N = \frac{\log_2 n}{h(p)} (1 + o(1))$. El plan se construye a partir de las columnas de altura N , cada una de las cuales contiene exactamente $\lfloor pN \rfloor$ unidades. Las columnas en el plan se disponen por bloques, empleando en caso de necesidad al final el bloque incompleto. Merced a tal disposición se asegura la diferencia entre el número de unidades en las filas del plan en no más de 1. La desigualdad $\binom{N}{\lfloor pN \rfloor} \geq n$ garantiza que al construir el plan nos será suficiente la reserva de columnas, con la cantidad de unidades indicada en cada una de ellas. Por cuanto la parte de las unidades en cada columna del plan no es superior a p , mientras que la cantidad de unidades en cualesquiera dos filas se diferencia en no más de 1, cada fila contiene no más de $k = \lceil pn \rceil$ unidades, es decir, dicho plan es realmente $F(n, \geq k)$ -plan. El teorema queda así demostrado.

Aduzcamos una estimación superior para la magnitud $N(n, \leq k)$, la cual es próxima a la estimación inferior obtenida en el teorema anterior por el método de entropía (véase [88 y 89]). Observemos que es un tanto mejor que la estimación superior de Katona de [37].

Teorema 14. Se verifica la desigualdad

$$N(n, \leq k) \leq \left\lceil \frac{\log_2 n}{\log_2 \frac{n}{k}} \right\rceil \left\lceil \frac{n}{k} - 1 \right\rceil$$

La demostración está basada en la construcción del $N(n, \leq k)$ -plan, partiendo del código binario de Rademajer. Así llamaremos una $(0, 1)$ -matriz de dimensión $\lceil \log_2 n \rceil \times n$, cuyas columnas son todas diferentes. Para obtener algún código de Rademajer es suficiente tomar una potencia del número 2, la más próxima a n superiormente, y, al elegir cualesquiera

n números inferiores a la citada potencia, escribirlos en columnas conforme al sistema de numeración con base 2.

Sea $n = 3$. Construyamos para este caso todos los códigos de Rademajer equivalentes con una exactitud de hasta una permutación de las columnas:

$$\begin{array}{cccccc} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}$$

Nos hará falta la siguiente identidad sencilla, válida para cualquier número real x y natural n :

$$\left\lceil \left\lfloor \frac{x}{n} \right\rfloor \right\rceil = \left\lfloor \frac{x}{n} \right\rfloor. \quad (9)$$

Demostración del teorema. Sean $R = R(k)$ un código arbitrario de Rademajer de longitud k , t , un número natural y s_0, \dots, s_{t-1} , números arbitrarios del conjunto $\{0, 1\}$. Partiendo de R , construyamos otro código de Rademajer $R_{s_0}, \dots, R_{s_{t-1}}$ del modo siguiente. Tomemos cada fila con número $i \in \{1, \dots, \lceil \log_2 k \rceil\}$ y sustituyamos en ella cada número $l \in \{0, 1\}$ por el número $(l + s_{(i-1) \bmod t}) \bmod 2$.

Escribamos arbitrariamente uno tras otro todos los 2^t códigos $R_{s_0}, \dots, R_{s_{t-1}}$ y obtendremos una matriz M de dimensión $\lceil \log_2 k \rceil \times 2^t k$. Esta matriz se divide en $\lambda_0 = \lceil t^{-1} \rceil \lceil \log_2 k \rceil = \lceil t^{-1} \log_2 k \rceil$ submatrices M_λ , $\lambda = 1, \dots, \lambda_0$ (véase (9)). Las submatrices M_λ están formadas, para $\lambda \neq \lambda_0$, por t filas de la matriz M con los números $i = t(\lambda - 1) + 1, t(\lambda - 1) + 2, \dots, t\lambda$. La submatriz M_{λ_0} está formada por las filas con los números $i = t(\lambda_0 - 1) + 1, t(\lambda_0 - 1) + 2, \dots, \lceil \log_2 k \rceil$ y puede tener menos de t filas.

No es difícil notar que cualquier submatriz M_λ contiene, para $\lambda \neq \lambda_0$, cada una de las 2^t columnas binarias de altura t exactamente k veces. Designemos el número de filas de la submatriz M_λ con τ . Análogamente, M_{λ_0} contiene cada una de las 2^t columnas binarias de altura τ exactamente $2^{t-\tau} \cdot k$ veces.

Denotemos con L_μ una matriz compuesta de k columnas iguales de altura t que representa una notación en columna del número $\mu \in \{0, 1, \dots, 2^t - 1\}$ con arreglo al sistema binario de numeración. Escribamos sucesivamente, una tras otra, todas las 2^t matrices L_μ ; designemos con M_0 la matriz obtenida. Agreguemos superiormente a la matriz M la matriz M_0 y designemos con \tilde{M} el resultado.

La matriz \tilde{M} contiene $t + \lceil \log_2 k \rceil$ filas y $2^t \cdot k$ columnas que son distintas dos a dos. Por consiguiente \tilde{M} es el código de Rademajer.

La demostración del teorema se basa en la construcción de un $F(n, \leq k)$ -plan, partiendo del código \tilde{M} , donde el número natural t se determina unívocamente de la condición $2^{t-1} < n/k \leq 2^t$. Para cada $\lambda = 0, \dots, \lambda_0 - 1$ construyamos la matriz M'_λ compuesta de n columnas de altura

$\left\lceil \frac{n}{k} \right\rceil - 1$. En cada columna todos los elementos son nulos, a excepción, quizás, de un elemento igual a 1, y en M'_λ cada columna no nula figura no más de k veces. En este caso dos columnas en M'_λ son iguales cuando y sólo cuando lo son las columnas con los números correspondientes en M_λ . Observemos que mediante dichas propiedades la matriz M' se define no unívocamente. Para nosotros tiene importancia el hecho de que tal matriz siempre existe. Está claro que cada fila de M'_λ contiene no más de k unidades.

De un modo análogo construyamos la matriz M'_{λ_0} . Esta consta de n columnas de altura $\left\lceil \frac{n}{k} \right\rceil - 2^{t-1}$. Cada una de estas columnas es o bien íntegramente nula, o bien contiene una sola unidad. Cada columna no nula figura en M'_{λ_0} no más de k veces, mientras que la columna nula, no más de $2^{t-1} \cdot k$ veces. Las columnas en M'_{λ_0} se disponen de un modo tal que de la condición de igualdad de dos columnas en M'_{λ_0} se deduce la igualdad de las columnas correspondientes en M'_{λ_0} . Cada fila de M'_{λ_0} contiene también no más de k unidades.

Escribamos sucesivamente, una debajo de la otra, las matrices $M'_0, M'_1, \dots, M'_{\lambda_0}$ y designemos el resultado con \tilde{M}' . Todas las columnas en \tilde{M}' , y, por tanto, en M' son distintas de dos en dos. Cada fila de \tilde{M}' contiene k unidades a lo sumo. Por consiguiente, \tilde{M}' es un $F(n, \leq k)$ -plan. Estimemos su longitud N :

$$\begin{aligned} N &= \lambda_0 \left(\left\lceil \frac{n}{k} \right\rceil - 1 \right) + \left(\left\lceil \frac{n}{k} \right\rceil - 2^{t-1} \right) \leq (\lambda_0 + 1) \times \\ &\times \left(\left\lceil \frac{n}{k} \right\rceil - 1 \right) = \left\lceil \frac{n}{k} \right\rceil - 1 + \frac{\log_2 k}{t} - \left\lceil \frac{n}{k} \right\rceil - 1 \leq \\ &\leq \left\lceil \frac{n}{k} \right\rceil + 1 \frac{\log_2 k}{\log_2 \frac{n}{k}} \left\lceil \frac{n-k}{k} \right\rceil = \left\lceil \frac{\log_2 n}{\log_2 \frac{n}{k}} \right\rceil \left\lceil \frac{n-k}{k} \right\rceil. \end{aligned}$$

El teorema está demostrado.

Ilustremos la construcción del $F(n, \leq k)$ -plan con un ejemplo en que $n = 20$ y $k = 5$. Tenemos: $t = \left\lceil \log_2 \frac{n}{k} \right\rceil = 2$. Veamos el código de Rademajer R de longitud 5:

$$R: \begin{array}{cccccc} & 1 & 1 & 1 & 0 & 0 \\ & 1 & 1 & 0 & 1 & 0 \\ & 1 & 0 & 1 & 1 & 0 \end{array}$$

Determinemos $R_{00}, R_{01}, R_{10}, R_{11}, L_0, L_1, L_2, L_3$:

$$\begin{array}{l}
 R_{00}: \begin{array}{l} 1\ 1\ 1\ 0\ 0 \\ 1\ 1\ 0\ 1\ 0; \\ 1\ 0\ 1\ 1\ 0 \\ 0\ 0\ 0\ 1\ 1 \end{array} \quad
 R_{01}: \begin{array}{l} 1\ 1\ 1\ 0\ 0 \\ 0\ 0\ 1\ 0\ 1; \\ 1\ 0\ 1\ 1\ 0 \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \end{array} \quad
 R_{10}: \begin{array}{l} 0\ 0\ 0\ 1\ 1 \\ 1\ 1\ 0\ 1\ 0; \\ 0\ 1\ 0\ 0\ 1 \\ 0\ 0\ 0\ 0\ 0; \\ 1\ 1\ 1\ 1\ 1; \end{array} \\
 R_{11}: \begin{array}{l} 0\ 0\ 0\ 1\ 1 \\ 0\ 0\ 1\ 0\ 1; \\ 0\ 1\ 0\ 0\ 1 \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \end{array} \quad
 L_0: \begin{array}{l} 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \end{array} \quad
 L_1: \begin{array}{l} 0\ 0\ 0\ 0\ 0; \\ 1\ 1\ 1\ 1\ 1; \\ 0\ 0\ 0\ 0\ 0; \\ 0\ 0\ 0\ 0\ 0; \\ 1\ 1\ 1\ 1\ 1; \end{array} \\
 L_2: \begin{array}{l} 1\ 1\ 1\ 1\ 1 \\ 0\ 0\ 0\ 0\ 0; \end{array} \quad
 L_3: \begin{array}{l} 1\ 1\ 1\ 1\ 1 \\ 1\ 1\ 1\ 1\ 1; \end{array}
 \end{array}$$

Las construcciones posteriores del código \tilde{M} y del $F(20, 5)$ -plan se ven del diagrama siguiente.

Código \tilde{M} :

$$\begin{array}{l}
 M_0: \begin{array}{l} 0\ 0\ 0\ 0\ 0 \quad 0\ 0\ 0\ 0\ 0 \quad 1\ 1\ 1\ 1\ 1 \quad 1\ 1\ 1\ 1\ 1 \\ 0\ 0\ 0\ 0\ 0 \quad 1\ 1\ 1\ 1\ 1 \quad 0\ 0\ 0\ 0\ 0 \quad 1\ 1\ 1\ 1\ 1 \end{array} \\
 M: \left\{ \begin{array}{l} M_1: \begin{bmatrix} 11100 & 11100 & 00011 & 00011 \\ 11010 & 00101 & 11010 & 00101 \end{bmatrix} \\ M_2: \begin{bmatrix} 10110 & 10110 & 01001 & 00101 \end{bmatrix} \end{array} \right.
 \end{array}$$

$\Gamma(20, 5)$ -plan

$$\begin{array}{l}
 M_0^1: \begin{bmatrix} 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 0 & 1\ 1\ 1\ 1\ 1 & 0\ 0\ 0\ 0\ 0 \end{bmatrix} \\
 M_1^1: \begin{bmatrix} 1\ 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0\ 0 & 0\ 0\ 0\ 1\ 0 & 0\ 0\ 0\ 0\ 1 \\ 0\ 0\ 1\ 0\ 0 & 1\ 1\ 0\ 0\ 0 & 0\ 0\ 0\ 0\ 1 & 0\ 0\ 0\ 1\ 0 \\ 0\ 0\ 0\ 1\ 0 & 0\ 0\ 0\ 0\ 1 & 1\ 1\ 0\ 0\ 0 & 0\ 0\ 1\ 0\ 0 \end{bmatrix} \\
 M_2^1: \left\{ \begin{array}{l} 1\ 0\ 1\ 1\ 0 \quad 1\ 0\ 1\ 0\ 0 \quad 0\ 0\ 0\ 0\ 0 \quad 0\ 0\ 0\ 0\ 0 \\ 0\ 0\ 0\ 0\ 0 \quad 0\ 0\ 0\ 1\ 0 \quad 0\ 1\ 0\ 0\ 1 \quad 0\ 1\ 0\ 0\ 1 \end{array} \right.
 \end{array}$$

De los teoremas 13 y 14 se deduce un

Corolario. Sea $n \rightarrow \infty$, $k = k(n) = o(n)$. En este caso es válida la fórmula asintótica:

$$N(n, k) = \frac{n \log_2 n}{k \log_2 k} (1 + o(1)).$$