Fundamentos de Programación Entera
5. Ramificado y acotamiento

Carlos Testuri – Fernando Islas

Departamento de Investigación Operativa – Instituto de Computación
Facultad de Ingeniería – Universidad de la República

2012-2023

Contenido

Resolución mediante ramificado y acotamiento

El esquema de resolución mediante *ramificado y acotamiento* (branch and bound) utiliza la técnica de “dividir y conquistar” el espacio de soluciones factibles.

En el caso de que el problema es difícil de resolver, se divide el espacio de soluciones (ramificado) generando subproblemas más simples y se determinan cotas del valor objetivo para evitar resolver algunos de los subproblemas (acotamiento).

A su vez cada subproblema puede dividirse nuevamente, generando un esquema jerárquico de particiones.

El método enumera todas las soluciones factibles (enumerativo).

Esquema de resolución mediante ramificado y acotamiento

Sea el problema

$$\begin{aligned} z = \quad & \max \quad c^\tau x \\ & \text{s.a} \quad\ \ x \in X \subset \mathbb{Z}, \end{aligned}$$

se particiona el conjunto factible X en subconjuntos X_k, con $k = 1, ..., K$, y se resuelven separadamente los subproblemas

$$\begin{aligned} z^k = \quad & \max \quad c^\tau x \\ & \text{s.a} \quad\ \ x \in X_k. \end{aligned}$$

La resolución del problema original es equivalente a

$$z = \max_{k \in K} z^k.$$

La división recurrente de subproblemas se puede representar con un árbol de enumeración.

Ejemplos de ramificación

Dado el conjunto factible $X \subseteq \{0,1\}^n$, se lo puede particionar en $X_0 = \{x \in X : x_1 = 0\}$ y en $X_1 = \{x \in X : x_1 = 1\}$.
A su vez X_0 se podría particionar en $X_{00} = \{x \in X_0 : x_2 = 0\}$ y en $X_{01} = \{x \in X_0 : x_2 = 1\}$.

Dado el conjunto factible $X \subseteq \mathbb{Z}^n$, y cierto $r \in \mathbb{R}$ se lo puede particionar en $X_{<r} = \{x \in X : x_1 \leq \lfloor r \rfloor\}$ y en $X_{>r} = \{x \in X : x_1 \geq \lceil r \rceil\}$.

Dado el conjunto factible $X \subseteq \{a,b,c\}^n$, se lo puede particionar en $X_a = \{x \in X : x_1 = a\}$, en $X_b = \{x \in X : x_1 = b\}$, y en $X_c = \{x \in X : x_1 = c\}$.

La cantidad total de particiones crece exponencialmente con los niveles de particionado. ¿Cómo acotar el proceso de particionado?

Acotamiento

Mientras resolver un subproblema puede ser difícil, obtener cotas de su valor objetivo puede ser fácil. Un método para obtener cotas superiores es mediante una relajación a programación lineal o dualidad. Las cotas inferiores se pueden obtener mediante soluciones factibles.

Se asume que existe la posibilidad de determinar, en forma eficiente, cotas sobre los valores objetivos, z^k, de los subproblemas X_k.
Sean $\overline{z}^k$ y $\underline{z}^k$, las cotas superiores e inferiores de los subproblemas,

$$\underline{z}^k \leq z^k \leq \overline{z}^k.$$

Las cotas permiten evitar resolver suproblemas, estableciendo lo que se denomina *podas* en el árbol de enumeración.

Podas por optimalidad, acotamiento e infactibilidad

Poda por optimalidad
Si $\underline{z}^k = \bar{z}^k = z^k$ entonces la solución óptima del subproblema se ha alcanzado, por lo que no es necesario continuar revisando X_k y su rama correspondiente en el árbol de enumeración puede podarse.

Poda por acotamiento
Dada la división de X en dos subconjuntos X_1 y X_2.
Si $\bar{z}^1 < \underline{z}^2$ entonces la solución óptima del problema no puede tener valor menor a $\underline{z}^2$, y como el valor máximo de X_1 no lo puede superar la rama del árbol de enumeración correspondiente a X_1 puede podarse.

Poda por infactibilidad
Si X_k es vacío la rama del árbol de enumeración correspondiente puede podarse.

Seudocódigo

El proceso implica llevar una lista de subproblemas *activos*: que no han sido podados o que necesitan explorarse. Sea $\underline{z}$ el valor del objetivo de la mejor solución factible encontrada hasta el momento.

Ramificado y acotamiento

1. Agregar a la lista el problema inicial. Establecer $\underline{z} := -\infty$ y $x^* := \{\}$.
2. Mientras la lista no esté vacía, seleccionar un subproblema activo, X_k.
3. Resolver la relajación del subproblema y obtener su cota superior, $\bar{z}^k$.
4. Si el subproblema es infactible, se lo descarta. Ir al paso 2.
5. Si $\bar{z}^k \leq \underline{z}$, se lo descarta por acotamiento. Ir al paso 2.
6. Si $x^k(relaj.)$ es entero, actualizar la cota $\underline{z}$ y la solución corriente x^*. Si no lo es dividirlo en nuevos subproblemas. Agregar los nuevos subproblemas a la lista de problemas activos. Ir al paso 2.
7. Reportar solución óptima x^*.

Esquema de resolución con múltiples opciones

- Cotas inferiores (primales) provistas por soluciones factibles.
- Cotas superiores (duales) provistas por relajación o problemas duales; ¿cuáles usar?
- Compromiso entre cotas débiles, calculadas con menor esfuerzo, y cotas fuertes, calculadas con mayor esfuerzo.
- Partición de la región factible; ¿se establece una regla general o evoluciona a medida que se resuelve el problema?
- Orden de evaluación de subproblemas; ¿ordenamiento de cola o priorizar según mayor cota superior?

Ejemplo ilustrativo (1/5)

Dado el siguiente problema con región factible denominada X,

$$\begin{aligned} z = \quad \max \quad & x_1 + 2x_2 \\ \text{s.a} \quad & -x_1 + x_2 \leq 1 \\ & x_1 + x_2 \leq 4 \\ & 5x_1 + 3x_2 \leq 15 \\ & x_1, x_2 \geq 0, \text{enteras.} \end{aligned}$$

Para **establecer las cotas** se relaja el problema a programación lineal. La solución del problema relajado da la cota superior $\bar{z} = \frac{13}{2}$ con solución $\bar{x} = (\frac{3}{2}, \frac{5}{2})$, la cota inferior se asume $\underline{z} = 0$.
Dada la diferencia de las cotas se debe ramificar.

Ejemplo ilustrativo (2/5)

Para **ramificar** se elige una variable básica que no es entera en la solución, dividiendo la región factible en dos a partir de restricciones excluyentes de esta.
Dada $x_1 = \frac{3}{2}$, se particiona la región factible X en $X_1 = X \cap \{x : x_1 \leq \lfloor \frac{3}{2} \rfloor = 1\}$ y $X_2 = X \cap \{x : x_1 \geq \lceil \frac{3}{2} \rceil = 2\}$, generando dos nuevos vértices activos en el árbol.

Se selecciona el vértice X_1 para continuar.
Se resuelve la relajación lineal del vértice X_1 a partir del método simplex-dual, que es más eficiente dado que permite reusar la base obtenida en la resolución de la relajación de X.
La solución del problema relajado da la cota superior $\bar{z}^1 = 5$ con solución entera $\bar{x}^1 = (1, 2)$.

Ejemplo ilustrativo (3/5)

Se selecciona el vértice X_2 para continuar.
Se resuelve la relajación lineal del vértice X_2 a partir del método simplex-dual.
La solución del problema relajado da la cota superior $\bar{z}^2 = \frac{16}{3}$ con solución entera $\bar{x}^2 = (2, \frac{5}{3})$. Por lo que X_2 no puede ser podado.

Dada $\bar{x}_2^2 = \frac{5}{3}$, se particiona la región factible X_2 en
$X_{21} = X_2 \cap \{x : x_2 \leq \lfloor \frac{5}{3} \rfloor = 1\}$ y $X_{22} = X_2 \cap \{x : x_2 \geq \lceil \frac{5}{3} \rceil = 2\}$, generando dos nuevos vértices activos en el árbol.

Ejemplo ilustrativo ejemplo (4/5)

Se selecciona el vértice X_{21} para continuar.
Se resuelve la relajación lineal del vértice X_{21} a partir del método simplex-dual.
La solución del problema relajado da la cota superior $\bar{z}^{21} = \frac{22}{5} = 4.4$ con solución $\bar{x}^{21} = (\frac{12}{5} = 2.4, 1)$. Dado que su cota superior es inferior a la de X_1 el vértice X_{21} es *podado por acotamiento*.

Ejemplo ilustrativo (5/5)

Se selecciona el vértice X_{22} para continuar.
Se resuelve la relajación lineal del vértice X_{22} a partir del método simplex-dual.
El problema relajado correspondiente no es factible, X_{22} es *podado por infactibilidad*.

Habiendo agotado la lista de subproblemas activos, se determina que la solución óptima del problema es la establecida en el vértice X_1: $z^* = 5$ con solución $x^* = (1, 2)$.

Ramificado y acotamiento en TSP

Dado el problema del vendedor viajero en grafo dirigido $G(V,A)$ con $N = \{1, ..., n\}$ y costos c_{ij} por cada arco $(i,j) \in A$.
Donde la variable $x_{ij} = 1$ si el vendedor va de la ciudad i a la j, y si no $x_{ij} = 0$.

$$
\begin{array}{lll}
\min & \sum_{(i,j)\in A} c_{ij}x_{ij} & \\
\text{s.a} & \sum_{j:(i,j)\in A} x_{ij} = 1, \quad i \in N, & (1) \\
& \sum_{i:(i,j)\in A} x_{ij} = 1, \quad j \in N, & (2) \\
& \sum_{i\in S} \sum_{j\in S} x_{ij} \leq |S| - 1, \quad \forall S \subseteq N,\ S \neq \emptyset, N, & (3) \\
& x_{ij} \in \{0,1\}, \text{enteras}, \quad (i,j) \in A. &
\end{array}
$$

Supongamos que se relaja el problema eliminando la restricción (3).
¿Qué características tiene la solución relajada?

Si la solución relajada contiene subtours. ¿Cómo se podría aplicar ramificado y acotamiento?

Resolución mediante ramificado y acotamiento: generalidades (1/2)

Se lleva registro de (por lo menos)

- los subproblemas activos,
- la mejor cota superior (dual),
- las cotas inferiores y superiores de los subproblemas,
- una base óptima o próximo a óptima de los subproblemas, a ser reusada.

Dependiendo de la disposición de memoria se puede llevar más registro y evitar la regeneración de información; un caso típico es el tamaño de la lista de subproblemas activos.

Resolución mediante ramificado y acotamiento: generalidades (2/2)

Para la **elección de la variable** a ramificar se puede elegir la de mayor valor fraccional (más próximo a 0,5).

Para **podar significativamente** el árbol se necesita una cota inferior ajustada. Dicha cota se puede obtener generando una solución factible a partir de una búsqueda en profundidad en el árbol.

Para **reducir la cantidad de subproblemas** a explorar se puede tener el criterio de siempre elegir el subproblema activo con mayor valor de cota superior (más promisorio).

Sistemas informáticos de resolución mediante ramificado y acotamiento

Los sistemas consisten en componentes de

- preprocesamiento: mecanismos de inferencia que mejoran la formulación inicial.
- heurísticas para determinación de soluciones factibles iniciales.
- algoritmos de resolución: simplex primal, simplex dual, punto interior, etc.
- mecanismos de selección de subproblemas.
- estrategias de selección de variables.
- mecanismos de ramificado GUB o SOS, clique, cobertura, flujos, etc.

Preprocesamiento: ejemplo (1/2)

Se trata de eliminar variables y restricciones redundantes, y ajustar cotas de variables, para que el problema de partida sea más fácil de resolver. Dado el ejemplo

$$
\begin{array}{rlll}
z = & \max & x_1 + 2x_2 & \\
 & \text{s.a} & -x_1 + x_2 \leq 1 & (1) \\
 & & x_1 + x_2 \leq 6 & (2) \\
 & & 5x_1 + 3x_2 \leq 15 & (3) \\
 & & 0 \leq x_1 & \\
 & & 0 \leq x_2 \leq 2. &
\end{array}
$$

Se puede *ajustar la cota* de la variable x_1: a partir de la restricción (3) y que $x_2 \geq 0$ se tiene que $x_1 \leq (15 - 3x_2)/5 \leq (15 - 0)/5 = 3$.

Preprocesamiento: ejemplo (2/2)

Se puede *eliminar la restricción* redundante (2): valorando su expresión con las cotas de las variables x_1 y x_2 se tiene que $x_1 + x_2 \leq 3 + 2 < 6$ por lo que la restricción es redundante.

Con lo que el problema queda reducido a

$$\begin{array}{rll} z = & \max & x_1 + 2x_2 \\ & \text{s.a} & -x_1 + x_2 \leq 1 \\ & & 5x_1 + 3x_2 \leq 15 \\ & & 0 \leq x_1 \leq 3 \\ & & 0 \leq x_2 \leq 2. \end{array}$$

Preprocesamiento: formalización (1/3)

Dado el problema según restricciones: $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$, $i = 1, ..., m$, y cotas de las variables $l_j \leq x_j \leq u_j$, $j = 1, ..., n$.

1. *Cota de variable*. Dada la variable k-ésima, x_k,
   Si $a_{ik} > 0$ entonces

$$x_k \leq (b_i - \sum_{j \neq k : a_{ij} > 0} a_{ij} l_j - \sum_{j \neq k : a_{ij} < 0} a_{ij} u_j)/a_{ik},$$

   si $a_{ik} < 0$ entonces

$$x_k \geq (b_i - \sum_{j \neq k : a_{ij} > 0} a_{ij} l_j - \sum_{j \neq k : a_{ij} < 0} a_{ij} u_j)/a_{ik}.$$

Preprocesamiento: formalización (2/3)

2. *Restricción redundante*. La restricción i-ésima es redundante si

$$\sum_{j:a_{ij}>0} a_{ij}u_j + \sum_{j:a_{ij}<0} a_{ij}l_j < b_i.$$

3. *Infactibilidad*. El problema es infactible si alguna restricción, $i = 1, ..., m$, cumple

$$\sum_{j:a_{ij}>0} a_{ij}l_j + \sum_{j:a_{ij}<0} a_{ij}u_j > b_i.$$

Preprocesamiento: formalización (3/3)

4. *Fijación de variable*. Para el caso de un problema max $\sum_{j=1}^{n} c_j x_j$
   Si $a_{ij} \geq 0$ para $i = 1, \ldots, m$, y $c_j < 0$, entonces $x_j = l_j$;
   si $a_{ij} \leq 0$ para $i = 1, \ldots, m$, y $c_j > 0$, entonces $x_j = u_j$.

5. *Cota de variable entera*. Si la variable x_j es entera y sus cotas, l_j o u_j, no son enteras, entonces
   $$\lceil l_j \rceil \leq x_j \leq \lfloor u_j \rfloor.$$

Preprocesamiento: ejemplo BIP (1/2)

En este caso se utiliza inferencia lógica. Dadas las restricciones

$$\begin{array}{cl} 6x_1 - 4x_2 + 3x_3 - x_4 \leq 1 & (1) \\ -2x_1 + x_2 - 4x_3 + 7x_4 \leq 2 & (2) \\ x_1 + 2x_2 + 3x_3 \geq 4 & (3) \\ x \in \{0,1\}^4. & \end{array}$$

1. *Inecuaciones lógicas*
   Si $x_1 = 1$ la restricción (1) implica que $x_2 = 1$ y $x_4 = 1$. Por lo que se infieren la inecuaciones $x_1 \leq x_2$ y $x_1 \leq x_4$.
   La restricción (1) es infactible si $x_1 = 1$ y $x_3 = 1$, por lo que se infiere la inecuación $x_1 + x_3 \leq 1$.
   La restricción (2) implica: $x_4 \leq x_3$ y $x_4 \leq x_1$.
   La restricción (3) implica: $x_1 \leq x_3$, $x_2 + x_3 \geq 1$ y $x_1 + x_3 \geq 1$.

Preprocesamiento: ejemplo BIP (2/2)

Dadas las inecuaciones lógicas de las restricciones:
- La restricción (1) implica: $x_1 \leq x_2, x_1 \leq x_4$ y $x_1 + x_3 \leq 1$.
- La restricción (2) implica: $x_4 \leq x_3$ y $x_4 \leq x_1$.
- La restricción (3) implica: $x_1 \leq x_3, x_2 + x_3 \geq 1$ y $x_1 + x_3 \geq 1$.

2. *Combinación de inecuaciones lógicas*
   De las restricciones (1) y (2) se tiene que $x_1 \leq x_4$ y $x_4 \leq x_1$, las que implican que $x_1 = x_4$.
   De las restricciones (1) y (3) se tiene que $x_1 + x_3 \leq 1$ y $x_1 \leq x_3$, las que implican que $x_1 = 0$. De (3) se tiene que $x_1 + x_3 \geq 1$ y dado que $x_1 = 0$ se tiene que $x_3 = 1$.
   De (2) se tiene que $x_4 \leq x_1$ y dado que $x_1 = 0$ se cumple $x_4 = 0$.