

# Parte I

## Abordagem Exata

## 2 Formulações Matemáticas

### 2.1 Introdução

Neste capítulo são apresentadas e referenciadas as principais formulações já existentes na literatura para o TOP. Em seguida, são apresentadas três formulações de Programação Inteira aqui propostas para o TOP. A primeira delas é uma formulação baseada em arcos. A segunda, uma formulação estendida. E a terceira, uma formulação baseada em rotas não elementares e que é utilizada no algoritmo *Branch-Cut-and-Price* proposto neste trabalho.

### 2.2 Formulações da literatura

A literatura específica para o TOP ainda é escassa. Anteriormente a este trabalho, haviam sido propostas quatro formulações para o problema.

A primeira formulação para o problema, anteriormente conhecido como *Multiple Tour Maximum Collection Problem*, foi proposta por Butt e Cavalier (BC94). As variáveis de decisão nesta formulação representam os arcos que conectam os nós no grafo. Assim, a solução ótima consiste em selecionar o conjunto ótimo de arcos para cada rota.

Conforme os testes realizados por Butt e Cavalier (BC94), resolver este problema utilizando essa formulação é extremamente custoso em termos de tempo computacional, especialmente quando se utiliza um resolvedor comercial genérico, sem considerar a estrutura peculiar do problema.

Posteriormente, Butt e Ryan (BR99) propuseram uma nova formulação para o problema com uma leve diferença na definição do problema. Eles definem que, para cada veículo, há um tempo máximo de duração da sua respectiva rota. Nessa formulação é definido um conjunto  $V = \{i|1, \dots, n\}$ . Desse conjunto não participam os depósitos (nós 0 e  $n + 1$ ). Desde que haja um número finito de subconjuntos de  $V$ , define-se  $\lambda_j$  como sendo o  $j$ -ésimo subconjunto. Os  $m$  veículos são divididas em  $c$  subconjuntos, de modo que para cada um desses  $c$  subconjuntos há uma duração máxima de rota permitida

pré-definida  $L_k$ . O número de rotas dos veículos do conjunto  $k$  é denominado  $t_k$ .

Isso permite que se tenham as seguintes definições:

$L(\lambda_j)$  = tempo de viagem total do ciclo ótimo do *Problema do Caixeiro Viajante* para todos os nós pertencentes ao subconjunto  $\lambda_j$  além do depósito.

- $m$  = o número total de veículos da frota
- $n$  = o número total de vértices ( $|V|$ )
- $s_i$  = tempo de serviço no vértice  $i \in V$
- $p_i$  = prêmio associado à visita do nó  $i$
- $L_k$  = duração máxima permitida para a rota a ser realizada pelo veículo  $k$
- $Q^k = \{j | L(\lambda_j) + \sum_{i \in \lambda_j} s_i \leq L_k\}$
- $a_{ij} = \begin{cases} 1 & \text{se } i \in \lambda_j \text{ e } j \in Q^k \\ 0 & \text{caso contrário} \end{cases}$
- $x_j^k = \begin{cases} 1 & \text{se os nós em } \lambda_j \text{ devem ser visitados por uma das } t_k \text{ rotas} \\ & \text{dos veículos do conjunto } k \\ 0 & \text{caso contrário} \end{cases}$

Finalmente, é calculado o valor  $c_j^k$  que corresponde ao lucro total coletado dos nós em  $\lambda_j$ :

$$p_j^k = \sum_{i \in \lambda_j} p_i = \sum_{i=1}^n p_i a_{ij}^k; j \in Q^k, k = 1, \dots, m \quad (2-1)$$

Dadas essas definições, o problema pode ser formulado da seguinte forma:

$$\text{maximizar} \quad \sum_{k=1}^m \sum_{j \in Q^k} p_j^k \cdot x_j^k \quad (2-2)$$

$$\sum_{k=1}^m \sum_{j \in Q^k} a_{ij}^k \cdot x_j^k \leq 1, \quad i = 1 \dots n \quad (2-3)$$

$$\sum_{j \in Q^k} x_j^k \leq t_k, \quad k = 1 \dots m \quad (2-4)$$

$$x_j^k \in \{0, 1\}, \quad \forall j \in Q^k, k = 1 \dots m \quad (2-5)$$

Nessa formulação, a função objetivo (2-2) visa a maximizar o lucro total coletado. A restrição 2-3 garante que cada nó é visitado somente uma vez por rota e aparece em no máximo uma rota. A restrição 2-4 garante que no

máximo  $t_k$  rotas de tamanho  $L_k$  podem ser realizadas. Finalmente, as restrições 2-5 forçam as variáveis a assumirem um valor inteiro 0-1. A viabilidade das rotas é definida e satisfeita implicitamente. Essa é, portanto, uma formulação de *Generalized Set Partitioning* e possui um número exponencial de variáveis.

Mais tarde, Tang e Miller-Hooks (TMH05) propuseram uma nova formulação para o TOP. Eles elaboraram uma formulação inteira utilizando a seguinte notação:

- $L$  = a duração máxima permitida para as rotas
- $y_{ik} = \begin{cases} 1 & \text{se o vértice } i \text{ é visitado pelo veículo } k \\ 0 & \text{caso contrário} \end{cases}$
- $x_{ijk}$  = o número de vezes que a arestas  $(i, j)$  é percorrida pelo veículo  $k$   
 $x_{ijk} \in \{0, 1\}$  se  $i, j \in V \setminus \{0\}$ ,  $x_{ijk} \in \{0, 1, 2\}$  se  $i = 0$  e  $j \in V \setminus \{0\}$
- $l_{ij}$  = tempo de viagem gasto no arco  $a = (i, j) \in E$
- $s_i$  = tempo de serviço no vértice  $i \in V$
- $U$  um subconjunto de vértices de  $V$ ,  $U \subset V$

$$\text{maximizar } \sum_{i=1}^{n-1} \sum_{k=1}^m p_i \cdot y_{ik} \quad (2-6)$$

$$\sum_{j=1}^{n-1} \sum_{k=1}^m x_{0jk} = 2m \quad (2-7)$$

$$\sum_{i < j} x_{ijk} + \sum_{i > j} x_{jik} > 2y_{jk}, \quad j = 1, 2, \dots, n-1; \quad k = 1, 2, \dots, m \quad (2-8)$$

$$\sum_{j=0}^{n-2} \sum_{j > i} l_{ij} x_{ijk} + \sum_{i=1}^{n-1} s_i y_{ik} \leq L, \quad k = 1, 2, \dots, m \quad (2-9)$$

$$\sum_{k=1}^m y_{ik} \leq 1 \quad i = 1, 2, \dots, n-1 \quad (2-10)$$

$$\sum_{i, j \in U, i, j} x_{ijk} \leq |U| - 1, \quad (U \subset V \setminus \{0\}; \quad n-2 \geq |U| \geq 2; \quad k = 1, 2, \dots, m) \quad (2-11)$$

$$x_{0jk} \in \{0, 1, 2\}, \quad (j = 1, 2, \dots, n-1; \quad k = 1, 2, \dots, m), \quad (2-12)$$

$$x_{ijk} \in \{0, 1\}, \quad (1 \leq i < j \leq n-1; \quad k = 1, 2, \dots, m), \quad (2-13)$$

$$y_{ik} \in \{0, 1\}, \quad (i = 1, 2, \dots, n-1; \quad k = 1, 2, \dots, m) \quad (2-14)$$

A função objetivo maximiza o prêmio total coletado por todas as rotas.

A restrição 2-7 assegura que há  $m$  veículos partindo e retornando ao depósito. A restrição 2-8 garante a conectividade de cada rota, enquanto que o limite de duração das rotas é garantido pelas restrições 2-9. A restrição 2-10 assegura que todos os vértices devem ser visitados no máximo uma vez, exceto o depósito. Subciclos são proibidos pela restrição 2-11. Finalmente, as restrições 2-12, 2-13 e 2-14 garantem a integralidade das variáveis de decisão. Essa formulação tem  $(n^2 \cdot m + (n - 2) \cdot m) / 2$  variáveis inteiras. A primeira formulação elaborada para esse problema, que é descrita em Butt e Cavalier (BC94) tem  $2n^2m$  variáveis inteiras.

Mais recentemente, Vansteenwegen et al. (VSBO09) propuseram uma formulação compacta para o TOP. Para essa formulação, foi utilizada a seguinte notação:

- $T_i$  é o tempo necessário para visitar o vértice  $i$
- $u_{ik}$  é a posição do vértice  $i$  na rota  $k$

$$x_{ijk} = \begin{cases} 1 & \text{se, na rota } k \text{ uma visita ao vértice } i \\ & \text{é seguida de uma visita ao vértice } j \\ 0 & \text{caso contrário} \end{cases}$$

$$y_{ik} = \begin{cases} 1 & \text{se a rota } k \text{ visita o vértice } i \\ 0 & \text{caso contrário} \end{cases}$$

$$\max \sum_{k=1}^m \sum_{i=2}^{n-1} S_i \cdot y_{ik} \quad (2-15)$$

$$\sum_{k=1}^m \sum_{i=2}^{n-1} x_{0jk} = \sum_{k=1}^m \sum_{i=2}^{n-1} x_{i(n+1)k} = m \quad (2-16)$$

$$\sum_{k=1}^m y_{jk} \leq 1, \forall j = 2, \dots, n \quad (2-17)$$

$$\sum_{i=1}^{n-1} x_{ijk} = \sum_{v=2}^n x_{jvk} = y_{jk}, \forall j = 1, \dots, n, \forall k = 1, \dots, m \quad (2-18)$$

$$\sum_{i=1}^{n-1} \left( T_i y_{ik} + \sum_{j=2}^n t_{ij} x_{ijk} \right) \leq L, \forall k = 1, \dots, m \quad (2-19)$$

$$2 \leq u_{ik} \leq n, \forall i = 1, \dots, n + 1; \forall k = 1, \dots, m \quad (2-20)$$

$$u_{ik} - u_{jk} + 1 \leq (n - 1)(1 - x_{ijk}), \forall i, j = 1, \dots, n + 1; \forall k = 1, \dots, m \quad (2-21)$$

$$x_{ijk}, y_{ik} \in \{0, 1\}, \forall i, j = 0, \dots, n + 1; k = 1, \dots, m \quad (2-22)$$

A função objetivo maximiza o lucro total recolhido nas visitas. A restrição 2-16 garante que cada rota tem início no vértice 1 e termina no vértice  $n$ . A restrição 2-17 assegura que cada vértice é visitado no máximo uma vez. A restrição 2-18 impõe que, se um vértice é visitado em uma rota, essa visita deve ser antecedida por exatamente uma visita e, da mesma forma, deve ser sucedida de exatamente uma visita, ambas na mesma rota. A restrição 2-19 força as rotas a terminarem antes do tempo máximo pré-estabelecido. As restrições 2-20, 2-21 e 2-22 são necessárias para evitar subciclos. Essas restrições de eliminação de subciclos foram formuladas de acordo com a formulação para o *Problema do Caixeiro Viajante* proposta em Miller et al. (MTZ60).

Além dessas formulações, há ainda a formulação proposta por Boushier et al. (BFG07), que é uma formulação baseada em rotas elementares, cuja relaxação é resolvida pela técnica de geração de colunas.

Nesta formulação que segue,  $Q$  é o conjunto de todas as possíveis rotas viáveis e  $p_j$  é o lucro total coletado pela rota  $j$ .

$$\max \sum_{\lambda_j \in Q} p_j \cdot \lambda_j \tag{2-23}$$

$$\sum_{a \in \delta^-(v)} \sum_{j \in Q} g_a^j \cdot \lambda_j \leq 1, v \in V \setminus v_0 \cup v_{n+1} \tag{2-24}$$

$$\sum_{a \in \delta^+(0)} \sum_{j \in Q} g_a^j \cdot \lambda_j \leq m, \tag{2-25}$$

$$y_v \in \{0, 1\}, \quad \forall v \in V^- \tag{2-26}$$

$$\lambda_j \in \{0, 1\}, \quad \forall j \in Q \tag{2-27}$$

Esta formulação é muito semelhante à formulação do Problema Mestre Restrito (subseção ) que foi utilizada neste trabalho. A diferença entre as duas formulações está no fato de que a formulação desta tese é baseada no grafo estendido, com um índice extra indicando o tempo de início do percurso dos arcos. Utilizar o grafo estendido permite que se identifique cortes que não teriam como ser separados utilizando-se apenas o grafo original.

### 2.3 Novas formulações

Esta seção apresenta as três formulações propostas neste trabalho para o TOP.

A primeira formulação é equivalente à que foi apresentada em Vansteenwegen et al. (VSBO09), que utiliza variáveis indicando se uma rota percorre ou não um determinado arco. Com a indexação por veículos, é necessário consid-

erar na formulação a restrição de duração máxima das rotas. Essa formulação foi denominada de *Formulação Baseada em Arcos*.

Na segunda formulação, essas variáveis indexadas são também indexadas pelo instante no qual elas iniciam na rota. Isso permite que seja evitada a indexação por veículos, pois não será considerada nenhuma variável indicando um arco que termine após o tempo máximo de duração das rotas. A primeira formulação é dita ser baseada em arcos. O número de variáveis é polinomial e, embora o número de restrições seja exponencial, elas podem ser separadas em tempo polinomial (restrições de eliminação de subciclos).

A segunda formulação tem um número de variáveis pseudo-polinomial e, portanto, é menos compacta do que a primeira e foi denominada de *Formulação Estendida*.

Por fim, é apresentada uma formulação com um número exponencial de colunas. Cada coluna representa uma possível rota e essa formulação pode ser vista como uma decomposição da primeira.

### 2.3.1

#### Notação

A notação usada nessas formulações pode ser descrita da seguinte forma:

- $G(V, A)$  é um grafo completo orientado, onde  $V$  é o conjunto de vértices e  $A$  é o conjunto de arcos;
- $V^- = \{1, \dots, n\}$  é conjunto dos vértices clientes que possuem prêmio associado;
- $v_0$  é o depósito inicial e  $v_{n+1}$  o depósito final;
- $p_v$  denota os prêmios não negativos, sendo  $v \in V^-$ ;
- $l_a$  representa o tempo de viagem nos arcos, sendo  $a = (i, j) \in A$ ;
- $m$  é o número de veículos da frota;
- $L$  é a tempo máximo de duração das rotas.

## 2.3.2

## Formulação de arcos

Essa formulação para o TOP usa variáveis binárias  $x_a^k$  e  $y_v$  de modo que:

$$x_a^k = \begin{cases} 1 & \text{indica que o arco } a \text{ é percorrido pelo veículo } k \\ 0 & \text{caso contrário} \end{cases}$$

$$y_v = \begin{cases} 1 & \text{indica que o vértice } v \text{ é visitado} \\ 0 & \text{caso contrário} \end{cases}$$

$$\max \sum_{v \in V^-} p_v \cdot y_v \quad (2-28)$$

$$\sum_{k=1}^m \sum_{a \in \delta^-(v)} x_a^k - y_v = 0, \forall v \in V^- \quad (2-29)$$

$$\sum_{k=1}^m \sum_{a \in \delta^-(S)} x_a^k - y_v \geq 0, S \subset V \forall v \in S \quad (2-30)$$

$$\sum_{k=1}^m \sum_{a \in \sigma(S)} x_a^k \leq |S| - 1, S \subset V \quad (2-31)$$

$$\sum_{k=1}^m x_a^k \leq 1, \forall a \in A \quad (2-32)$$

$$\sum_{a \in A} l_a x_a^k \leq L, k = 1, \dots, m \quad (2-33)$$

$$\sum_{a \in \delta^+(v_0)} x_a^k = 1, k = 1, \dots, m \quad (2-34)$$

$$\sum_{a \in \delta^-(v_{n+1})} x_a^k = 1, k = 1, \dots, m \quad (2-35)$$

$$y_v \in \{0, 1\}, \forall v \in V^- \quad (2-36)$$

$$x_a^k \in \{0, 1\}, \forall a \in A \forall k = 1, \dots, m \quad (2-37)$$

A função objetivo maximiza (2-28) a soma dos prêmios associados aos vértices visitados. Restrições (2-29) asseguram quem um cliente é visitado por no máximo um veículo.

A conectividade das rotas é garantida pelas restrições (2-30), enquanto que a eliminação de subciclos é assegurada pelas restrições (2-31), onde  $\sigma(S)$  denota todos os arcos com ponto inicial e final em  $S$ . O conjunto de restrições (2-32) proíbe o uso de um arco por duas ou mais rotas. A duração máxima das rotas é imposta pela restrição (2-33). Os conjuntos de restrições (2-34) e (2-35) forçam  $m$  veículos a saírem do depósito inicial e retornarem ao depósito

final.

### 2.3.3

#### Formulação Estendida

Nessa formulação, cada arco tem um índice extra  $l$ . Esse índice representa o tempo de partida do veículo ao utilizar esse arco.

$$x_a^{lk} = \begin{cases} = 1 & \text{indica que o veículo } k \text{ percorre o arco } a \\ & \text{com } l \text{ unidades de tempo consumidas.} \\ = 0 & \text{caso contrário} \end{cases}$$

$$y_v = \begin{cases} = 1 & \text{indica que o vértice } v \text{ é visitado} \\ = 0 & \text{caso contrário} \end{cases}$$

Como cada arco pode ser utilizado somente uma vez, então ele só pode iniciar o percurso em uma única ocasião com um  $l$  unidades de tempo já gastas:

$$x_a^k = \sum_{l=0}^L x_a^{lk} \tag{2-38}$$

Com o intuito de considerar a duração associada aos arcos, o grafo original foi modificado. Foi criado um vértice intermediário  $w_a$  para cada arco  $a \in A$ . Esses vértices artificiais têm uma demanda igual ao tempo gasto no percurso do arco original  $a$ . Por exemplo, seja  $w_{ij}$  o ponto intermediário do arco  $a = (i, j)$ . O arco original dá origem a dois novos arcos  $a_1 = (i, w_{ij})$  e  $a_2 = (w_{ij}, j)$ . O grafo modificado resultante possui portanto o conjunto de arcos  $A_1 \cup A_2$ , onde  $A_1 = \{(i, w_{ij}), (i, j) \in A\}$  e  $A_2 = \{(w_{ij}, j), (i, j) \in A\}$ . O conjunto de vértices é dado por  $V \cup A$ . A Figura 2.1 mostra a transformação do grafo. O vértice intermediário (ponto cinza) consome uma demanda (tempo) que corresponde ao tempo de viagem do arco original  $(i, j)$ . Nesse caso, o tempo de viagem do arco  $(i, j)$  é 2.

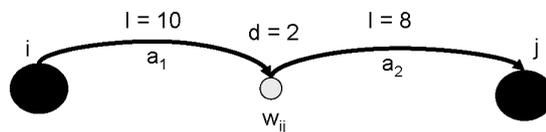


Figura 2.1: Transformação sobre o grafo

A formulação sobre as variáveis  $x_a^{lk}$  é obtida substituindo variáveis  $x_a^k$  de acordo com a equação (2-38) na restrição (2-29) na formulação baseada em arcos 2.3.2. Como mencionado acima, restrições (2-33) não são necessárias, desde que a duração limite das rotas pode ser respeitada considerando-se

somente os índices  $l$  apropriados. Além disso, impõe-se que  $m$  arcos deixem o vértice 0 e retornem ao vértice  $n + 1$ . Dessa forma, cada rota é unicamente identificada e, conseqüentemente, isso nos permite a remoção do índice  $k$  das variáveis  $x_a^{lk}$ . Essa segunda formulação é então escrita sobre as variáveis  $x_a^l$ ,  $a \in A_1 \cup A_2$ ,  $l = 0, 1, \dots, L$ . De fato, seria mais preciso que usar com sendo o maior valor de  $l$  a diferença entre  $L$  e o tempo de viagem do arco conectando o último vértice visitado pela rota ao depósito final  $n + 1$ . Todavia, usa-se simplesmente  $L$  no intuito de simplificar a notação.

$$\max \quad \sum_{v \in V^-} p_v \cdot y_v \quad (2-39)$$

$$\sum_{l=0}^L \sum_{a \in \delta^-(v)} x_a^l - y_v = 0, \quad \forall v \in V^- \quad (2-40)$$

$$\sum_{a \in \delta^-(v)} x_a^l - \sum_{a \in \delta^+(v)} x_a^l = 0, \quad \forall l = 0, \dots, L \quad \forall v \in V^- \quad (2-41)$$

$$x_{(i,w_{ij})}^l = x_{(w_{ij},j)}^{(l-l(i,j))}, \quad \forall l = 0, \dots, L \quad \forall (i,j) \in A \quad (2-42)$$

$$\sum_{l=0}^L \sum_{a \in \delta^+(0)} x_a^l = m, \quad (2-43)$$

$$\sum_{a \in \delta^-(n+1)} x_a^0 = m, \quad (2-44)$$

$$y_v \in \{0, 1\}, \quad \forall v \in V^- \quad (2-45)$$

$$x_a^l \in \{0, 1\}, \quad \forall a \in A_1 \cup A_2 \quad \forall l = 0, \dots, L \quad (2-46)$$

Essa formulação pode ser vista como uma formulação baseada em fluxo. A conservação de fluxo é garantida pelas restrições (2-41) sobre os vértices originais, enquanto que as restrições (2-42) têm a mesma função no que diz respeito aos vértices intermediários. Elas também impõem o consumo de unidades de tempo associado ao percurso dos arcos. Finalmente, restrições (2-43) e (2-44) determinam o número de rotas.

### 2.3.4

#### Formulação de Geração de Colunas

Como o (pseudo-polinomial) número de variáveis na formulação apresentada pode vir a ser muito grande, foi aplicada uma decomposição de Dantzig-Wolfe. O problema mestre considera somente as restrições que mantêm o controle dos vértices visitados e a restrição que assegura o número de rotas  $m$ . As colunas representam as rotas que já são geradas garantindo a conservação de fluxo, evidentemente. Para soluções inteiras da segunda formulação, as colunas são rotas elementares. Por outro lado, quando a relaxação linear é considerada,

a formulação menos compacta é equivalente a essa formulação de geração de colunas, quando as rotas podem também ser não-elementares, ou caminhos no grafo.

Todas as possíveis colunas podem ser expressas em termos dos seus arcos indexados pelo seu instante inicial da rota, independente da rota ser elementar ou não-elementar. O coeficiente  $g_a^{lj}$  indica que o arco  $a$  iniciando com  $l$  unidades de tempo sobrando é utilizado na rota  $j$ . O conjunto  $Q$  representa todas as possíveis rotas. Finalmente,  $\lambda_j$  representam as variáveis que indicam se a uma rota  $j$  (elementar ou não elementar) é selecionada. Pode-se portanto, escrever:

$$\sum_{j \in Q} g_a^{lj} \cdot \lambda_j = x_a^l \quad \forall a \in A \quad \forall l = 0, \dots, L \quad (2-47)$$

A formulação de geração de colunas é então obtida pela substituição das variáveis  $x_a^l$  nas restrições (2-40) e (2-44) de conformidade com a equação acima (2-47). A nova formulação é dada por:

$$\max \quad \sum_{v \in V^-} p_v \cdot y_v \quad (2-48)$$

$$\sum_{a \in \delta^-(v)} \sum_{j \in Q} \sum_{l=0}^L g_a^{lj} \cdot \lambda_j = y_v, \quad \forall v \in V^- \quad (2-49)$$

$$\sum_{a \in \delta^+(0)} \sum_{j \in Q} g_a^{0j} \cdot \lambda_j = m, \quad (2-50)$$

$$y_v \in \{0, 1\} \quad \forall v \in V^- \quad (2-51)$$

$$\lambda_j \in \{0, 1\}, \quad \forall j \in Q \quad (2-52)$$

Restrições (2-49) garantem que se um vértice é visitado, alguma rota selecionada deve visitá-lo. As restrições (2-50) forçam que  $m$  rotas partam do depósito inicial.

Esta formulação é utilizada na fase geração de colunas do algoritmo BCP, como problema mestre restrito.

## 2.4

### Conclusão

Este capítulo apresentou as três formulações propostas para o TOP. A primeira delas, é semelhante à formulação de Vansteenwegen et al. (VSBO09). A segunda, é uma formulação estendida. Essa formulação é útil pois os cortes utilizados no algoritmo BCP proposto neste trabalho são descritos em termos das variáveis de arcos desta formulação. A terceira formulação é baseada em rotas não elementares. Essa formulação é relaxada e usada como problema mestre restrito na fase de geração de colunas. Os cortes gerados em termos

das variáveis de arcos da segunda formulação são traduzidos e adicionados à terceira formulação.