
CAPÍTULO 4

O Problema do Caminho Mais Curto com Múltiplos Objectivos

1. Formulação do problema

Considere-se uma rede $\mathcal{G} = (\mathbf{N}, \mathbf{A}, \mathbf{C})$ com n nós e m arcos, em que $\mathbf{C} = (c^1, c^2, \dots, c^h)$ com $h \geq 1$ corresponde aos valores numéricos não negativos associados a cada arco. Ou seja, ao arco (i, j) estão associados os valores $c_{ij} = (c_{ij}^1, \dots, c_{ij}^h)$.

Sejam s e t dois nós da rede \mathcal{G} . Os valores numéricos associados ao caminho p de s para t na rede \mathcal{G} formam um h -uplo de somas, em que cada uma delas corresponde à soma dos valores associados aos arcos que pertencem àquele caminho. Ou seja,

$$c(p) = (c^1(p), c^2(p), \dots, c^h(p)) \text{ em que } c^w(p) = \sum_{(i,j) \in p} c_{ij}^w, \quad w \in \{1, 2, \dots, h\}.$$

O problema do caminho mais curto com vários (h) objectivos entre os nós s e t da rede \mathcal{G} , pode ser formulado da seguinte forma :

$$\text{Min } Z = (Z_1, \dots, Z_h)$$

sujeito a

$$\sum_{j \in \mathbf{N}} x_{sj} = 1$$

$$\sum_{i \in \mathbf{N}} x_{ij} - \sum_{l \in \mathbf{N}} x_{jl} = 0, \quad \forall j \in \mathbf{N} - \{s, t\}$$

$$\sum_{i \in \mathbf{N}} x_{it} = 1$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j)$$

onde,

$$Z_1 = \sum_{i \in N} \sum_{j \in N} c_{ij}^1 x_{ij}$$

...

$$Z_h = \sum_{i \in N} \sum_{j \in N} c_{ij}^h x_{ij}$$

$$x_{ij} = \begin{cases} 1, & \text{se o arco } (i, j) \text{ pertence ao caminho} \\ 0, & \text{caso contrário} \end{cases}.$$

2. Tipo de soluções do problema

Nos problemas de programação linear com vários objectivos, o conjunto das soluções admissíveis eficientes é um conjunto convexo contínuo, denominado por polítopo convexo (quando limitado tem o nome de poliedro convexo), o qual é construído à custa das restrições do problema. Desta forma, qualquer solução vértice eficiente pode ser determinada optimizando uma combinação convexa das funções objectivo (as soluções não vértices na fronteira eficiente podem depois ser caracterizadas à custa dos vértices).

No entanto, como os problemas de caminho mais curto multiobjectivo são de natureza discreta, a fronteira eficiente não é um conjunto convexo contínuo, podendo existir soluções não dominadas que não estejam situadas na fronteira do invólucro convexo (*contorno convexo*), mas sim no seu interior, nos chamados *desníveis de dualidade* ("duality gaps"). Às soluções não dominadas que pertencem ao contorno convexo, dá-se o nome de **soluções suportadas**, as quais podem ainda ser **extremas** (ou **vértices**) ou **não extremas**. Às soluções que se encontram nos desníveis de dualidade, dá-se o nome de **soluções não suportadas**.

Seja Z^\geq o invólucro convexo de $[Z \oplus \{z \in \mathfrak{R}^h : z \geq 0\}]$ (em que \oplus representa a adição de conjuntos e h o número de objectivos do problema).

1. Seja $z \in Z$ uma solução não dominada. Se z pertence à fronteira de Z^\geq então z é suportada; caso contrário, z é não suportada (é dominada por uma combinação convexa de outras soluções não dominadas suportadas).
2. Seja $z \in Z$ uma solução não dominada suportada. Se z não puder ser expressa como combinação convexa de outras soluções suportadas, então z é extrema; caso contrário, z é não extrema [25]. Ou seja, z é extrema se e só se existir $\alpha_k > 0$ com $k = 1, \dots, h$, tal que

$$\sum_{k=1}^h \alpha_k z_k < \sum_{k=1}^h \alpha_k w_k, \quad \forall w \in Z : z = (z_1, \dots, z_h) \neq (w_1, \dots, w_h) = w \quad [15].$$

O gráfico da Fig. 1 exemplifica os conceitos de solução não dominada suportada (extrema e não extrema) e não suportada :

- a) as soluções 1, 2 e 4 são não dominadas suportadas extremas (vértices),
- b) a solução 3 é não dominada suportada não extrema (é uma combinação convexa das soluções 2 e 4, pois encontra-se representada sobre a aresta $\overline{24}$),
- c) as soluções 5 e 6 são não dominadas não suportadas (dominadas pelas combinações convexas de 2 com 3 e de 3 com 4, respectivamente).

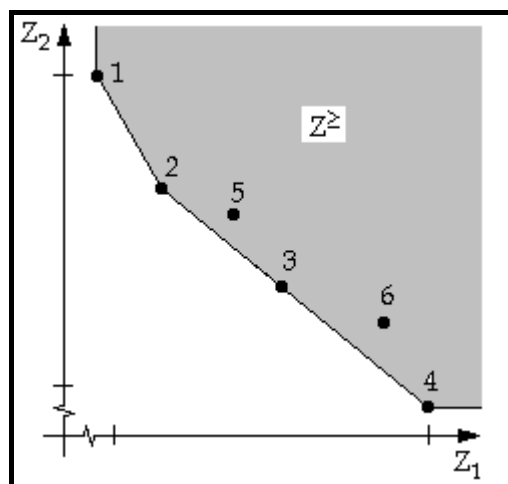


Fig. 1 – Tipo de soluções não dominadas : suportadas e não suportadas.

As imagens inversas das soluções não dominadas suportadas denominam-se por pontos eficientes suportados (no espaço das decisões) e das não dominadas não suportadas por pontos eficientes não suportados [25].

Os algoritmos que têm sido propostos até agora utilizam processos diferentes para determinar soluções não dominadas, consoante sejam suportadas extremas ou suportadas não extremas e não suportadas. No entanto, na pesquisa daquelas soluções, utiliza-se sempre o problema do caminho mais curto com um só objectivo.

O processo para determinar as soluções não dominadas suportadas extremas consiste em resolver um problema de caminho mais curto com um só objectivo, em que a função objectivo é uma combinação convexa das p funções objectivo do problema multiobjectivo.

Para se determinar as soluções não dominadas não suportadas têm sido apresentados, ao longo dos tempos, alguns processos diferentes, mas todos exigindo um elevado esforço computacional.

3. Um problema real de caminho mais curto com múltiplos objectivos

Muitos problemas reais podem ser modelados como um problema de caminho mais curto multiobjectivo, em que os dados do modelo são organizados recorrendo a uma estrutura em rede.

Como exemplo, veja-se um caso de circulação de veículos, mencionado em Rodrigues [21]. Pretende-se determinar o melhor caminho entre dois nós de uma rede de estradas (por exemplo, entre um ponto onde se carregam componentes que devem ser transportadas para uma fábrica localizada num outro ponto da rede), tendo em conta um compromisso entre dois objectivos : distância percorrida e tempo gasto. O tempo gasto num percurso pode ser influenciado por atrasos na entrega da mercadoria, uma vez que o tempo de circulação em cada artéria (arco da rede) pode variar muito, devido a problemas de congestão de tráfego, acidentes, obras de manutenção, filas de espera para pagamentos de portagens, etc..

Geralmente, pretende-se determinar uma solução cujos valores associados a cada objectivo se encontrem o mais próximo possível dos valores respectivos mínimos, de modo a rentabilizar o lucro da empresa responsável pela entrega das componentes a transportar. Neste caso, está-se perante um problema de caminho mais curto bi-objectivo.

Ao acrescentar-se mais um objectivo, por exemplo o número de artérias percorridas (ou quantidade de nós visitados), o problema transforma-se num problema de caminho mais curto tri-objectivo, se se pretender uma solução em que o número de nós a visitar seja também a menor possível.

Se se acrescentar o nível de perigosidade em termos de acidentes rodoviários (relacionado com o número de acidentes nas artérias), o problema transforma-se num de caminho mais curto tetra-objectivo, pois pretende-se uma solução cujo valor associado a este objectivo seja também o menor possível.

4. O método NISE

Uma das abordagens mais importantes no âmbito dos problemas de caminho mais curto com vários objectivos é o método NISE (“NonInferior Solutions Estimation”), que foi proposto por Cohon [9]. Este método destina-se a determinar uma aproximação do conjunto de soluções não inferiores e foi descrito, sobretudo, para problemas com duas funções objectivo.

Suponha-se que se pretende resolver um problema de caminho mais curto com dois objectivos, em que a função objectivo pode ser expressa da seguinte forma :

$$\text{Minimizar } Z = (Z_1, Z_2) .$$

As soluções não dominadas são determinadas utilizando o método da soma pesada das funções objectivo, reduzindo, desta forma, o problema original a um com apenas uma função objectivo. Para tal, efectua a soma das duas funções depois de se ter multiplicado uma delas por um parâmetro (peso) positivo w , ficando :

$$\text{Minimizar } Z(w) = w.Z_1 + Z_2 .$$

Desta forma, se o valor atribuído a w for superior a 1 então está-se a favorecer o primeiro objectivo em detrimento do segundo, acontecendo o inverso se w for inferior a 1, pois o peso que se está a atribuir ao segundo objectivo é constante e igual a 1.

Pode-se generalizar a utilização de pesos, atribuindo pesos a ambas as funções objectivo, obtendo-se :

$$\text{Minimizar } Z(w_1, w_2) = w_1.Z_1 + w_2.Z_2 .$$

No entanto, se se dividir as duas funções por um mesmo valor positivo (por exemplo, w_2), obtém-se a equação anterior, fazendo $w = w_1/w_2$. Desta forma, conclui-se que neste tipo de problemas o que interessa efectivamente é a relação entre os pesos atribuídos aos objectivos e não o seu valor absoluto.

O método NISE utiliza, em cada iteração, o seguinte critério para seleccionar os pesos : escolher valores de tal modo que a próxima solução esteja o mais distante possível da recta que passa pelas duas soluções previamente escolhidas.

O método começa por determinar as duas soluções não dominadas (vértices) que optimizam (minimizam) isoladamente cada uma das funções objectivo : as soluções A e B que minimizam Z_1 e Z_2 , respectivamente (Fig. 2).

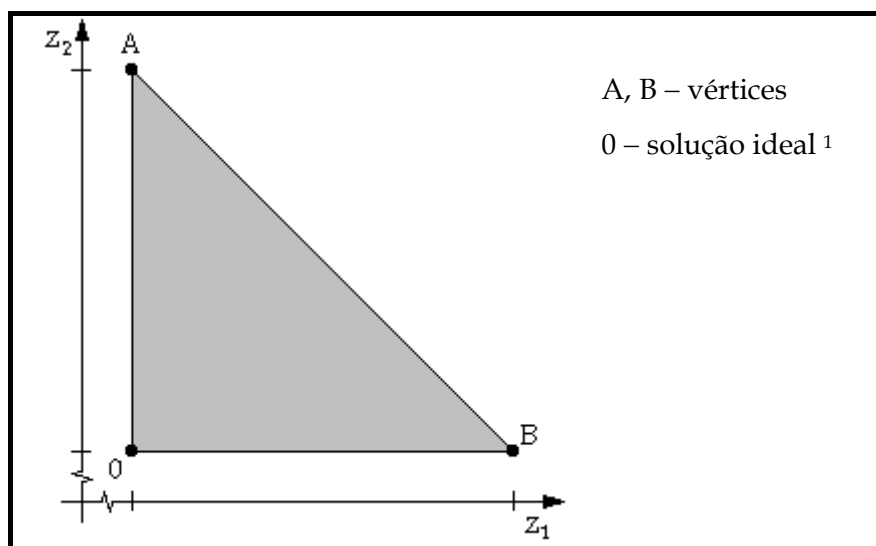


Fig. 2 – Método NISE : zona de pesquisa de soluções não dominadas (vértices).

A próxima solução não dominada a ser determinada, se existir, deve ser de tal modo que a sua representação deve situar-se o mais distante possível do segmento \overline{AB} , mas no interior do triângulo $\Delta(0,A,B)$. Como facilmente se pode verificar, não existem soluções à esquerda do segmento \overline{OA} , pois se existisse alguma solução nessa zona ela teria sido encontrada quando se optimizou a função objectivo Z_1 isoladamente. Da mesma maneira, se conclui que não podem existir soluções abaixo do segmento \overline{OB} .

Por outro lado, o segmento \overline{AB} servirá de base ao cálculo da próxima solução não dominada. Com efeito, aplicando a regra utilizada pelo método NISE para o cálculo dos pesos w_1 e w_2 , estes são tais que

$$-\frac{w_1}{w_2} = \alpha,$$

em que α é o declive da recta que contém o segmento \overline{AB} . A direcção da busca é perpendicular ao segmento \overline{AB} . Desta forma, qualquer nova solução admissível determinada por este método, só poderá situar-se no interior do triângulo $\Delta(0,A,B)$.

¹ Denomina-se por **solução ideal** aquela que optimizaria simultaneamente todas as funções objectivo (os seus valores obtêm-se considerando os valores das soluções que optimizam separadamente cada função objectivo), a qual é não admissível sempre que as funções objectivo estejam em conflito.

Como no cálculo dos pesos apenas interessa a relação entre eles, e sendo conhecido o declive da recta que contém o segmento \overline{AB} , basta igualar ao valor unitário um dos pesos para se calcular o outro. Neste caso, o declive da referida recta é dado por :

$$\alpha = \frac{Z_2(B) - Z_2(A)}{Z_1(B) - Z_1(A)}.$$

Como este declive é negativo (numerador negativo e denominador positivo) e os pesos devem ser escolhidos de forma que $-(w_1/w_2) = \alpha$, o quociente entre os pesos será um valor positivo. Isto é,

$$-\frac{w_1}{w_2} = \alpha = \frac{Z_2(B) - Z_2(A)}{Z_1(B) - Z_1(A)} = -\frac{Z_2(A) - Z_2(B)}{Z_1(B) - Z_1(A)} \Rightarrow \begin{cases} w_1 = Z_2(A) - Z_2(B) \\ w_2 = Z_1(B) - Z_1(A) \end{cases}.$$

Uma vez que é a relação entre os pesos que interessa, fazendo $w_2 = 1$, vem :

$$w_1 = \frac{Z_2(A) - Z_2(B)}{Z_1(B) - Z_1(A)}.$$

Resolvendo o problema cuja função objectivo é :

$$\text{Minimizar } Z(A, B) = w_1 \cdot Z_1 + Z_2$$

seria determinada a solução C e obter-se-ia uma nova aproximação para a região das soluções admissíveis não dominadas, zona a cheio na Fig. 3 — o triângulo $\Delta(0,0',0'')$ corresponde à região onde não se encontram soluções admissíveis, com o segmento $\overline{0'0''}$ paralelo a \overline{AB} .

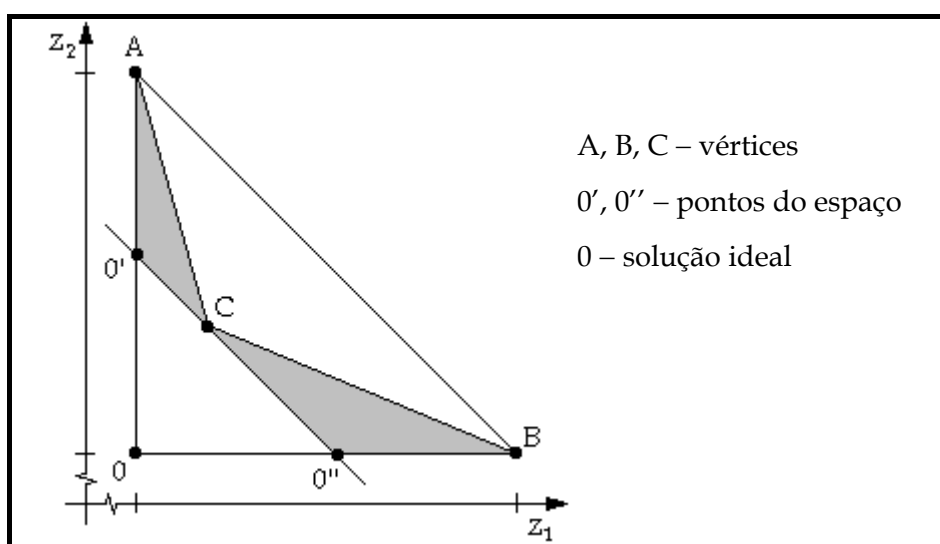


Fig. 3 – Método NISE : determinação de um vértice a partir de dois outros.

Se se prosseguir com o método, são possíveis duas novas direcções de busca : uma perpendicular ao segmento \overline{AC} , que percorrerá o triângulo $\Delta(0',A,C)$, e outra perpendicular ao segmento \overline{BC} , que percorrerá o triângulo $\Delta(0'',B,C)$.

Em termos genéricos, em cada iteração os valores dos pesos são os seguintes :

$$w_1 = \frac{Z_2(i) - Z_2(i+1)}{Z_1(i+1) - Z_1(i)}$$

$$w_2 = 1$$

em que i corresponde à i -ésima solução não dominada suportada extrema, tendo em conta que estas estão numeradas pela ordem decrescente do respectivo valor de Z_2 (a solução A é a que tem o maior valor de Z_2).

O método só termina, quando o máximo erro possível for inferior a determinado valor preestabelecido. Pela Fig. 3 pode-se verificar que as extensões das novas buscas estão relacionadas com as alturas dos triângulos (considera-se como bases dos triângulos os segmentos \overline{AC} e \overline{BC} , para cada caso); o erro máximo corresponde à maior altura.

5. O problema de caminho mais curto bi-objectivo

Nesta secção serão descritos alguns métodos para determinar soluções não dominadas suportadas extremas e não suportadas, assim como algumas abordagens interactivas para ajudar o AD a identificar a “melhor” solução de compromisso em problemas de caminho mais curto bi-objectivo.

5.1. Formulação do problema

Considere-se uma rede $G = (N, A, C)$ com n nós e m arcos, em que $C = (c^1, c^2)$. Ou seja, a cada arco (i, j) estão associados 2 valores não negativos, que correspondem ao “comprimento” do arco.

Sejam s e t dois nós da rede G . Os valores associados ao caminho p de s para t na rede G , são um par de somas, em que cada uma delas corresponde à soma dos valores associados aos arcos que pertencem àquele caminho; ou seja,

$$c(p) = (c^1(p), c^2(p)), \text{ em que, } c^1(p) = \sum_{(i,j) \in p} c_{ij}^1 \text{ e } c^2(p) = \sum_{(i,j) \in p} c_{ij}^2.$$

Seja P_i o conjunto de todos os caminhos do nó i para o nó t , em que $P_i \neq \emptyset, \forall i \in N$. Seja ainda P o conjunto de todos os caminhos de s para t . O problema do caminho mais curto com dois objectivos, entre os nós s e t , pode ser formulado da seguinte forma :

$$[\text{P1}] \quad \text{Min} \quad Z = (Z_1, Z_2) \quad (5.1)$$

sujeito a

$$\sum_{j \in N} x_{sj} = 1 \quad (5.2)$$

$$\sum_{i \in N} x_{ij} - \sum_{k \in N} x_{jk} = 0, \quad \forall j \in N - \{s, t\} \quad (5.3)$$

$$\sum_{i \in N} x_{it} = 1 \quad (5.4)$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j) \quad (5.5)$$

onde,

$$Z_1 = \sum_{i \in N} \sum_{j \in N} c_{ij}^1 x_{ij}$$

$$Z_2 = \sum_{i \in N} \sum_{j \in N} c_{ij}^2 x_{ij}$$

$$x_{ij} = \begin{cases} 1, & \text{se o arco } (i, j) \text{ pertence ao caminho} \\ 0, & \text{caso contrário} \end{cases} .$$

5.2. Representação do espaço dos objectivos

Todas as soluções admissíveis formam um conjunto convexo que se chama polígono convexo, uma vez que este conjunto é limitado (politopo convexo limitado). Desta forma, estas soluções encontram-se na sua fronteira (suportadas) e no seu interior (não suportadas). No entanto, como mostra a Fig. 4, enquanto que todas as soluções suportadas são não dominadas, as não suportadas são dos dois tipos. De facto, analisando a Fig. 4, verifica-se que qualquer solução que se encontre na região mais clara do invólucro convexo (solução não suportada) é dominada por soluções suportadas (por exemplo, a solução C domina qualquer solução que se encontre à direita da recta que contém o segmento \overline{CQ} e acima da recta que contém o segmento \overline{CR}).

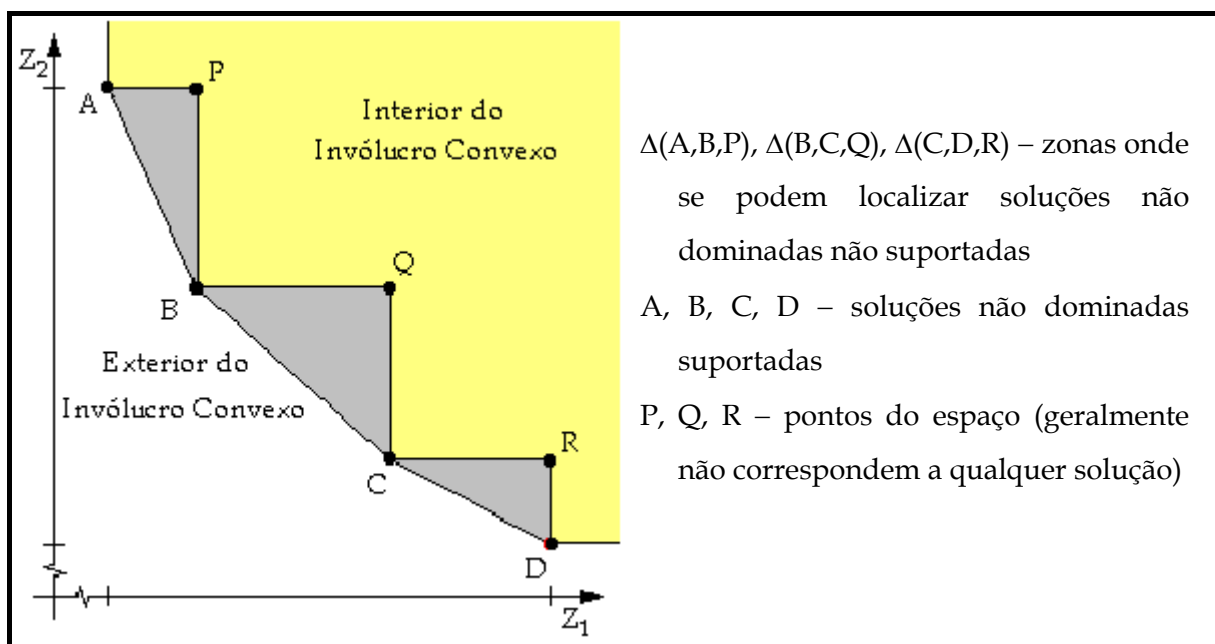


Fig. 4 – Bi-objectivo : representação das soluções no espaço dos objectivos.

Uma *zona de desnível de dualidade* é definida por dois vértices adjacentes e por um ponto formada pelos valores máximos das duas funções objectivo, relativamente àquelas soluções, formando um triângulo (por exemplo, $\Delta(A,B,P)$, $\Delta(B,C,Q)$ e $\Delta(C,D,R)$ formam três zonas de desníveis de dualidade na Fig. 4).

5.3. Métodos para determinar soluções não dominadas

Um processo usual é gerar todo o conjunto de soluções não dominadas para que o AD possa *a posteriori* seleccionar a “melhor” solução existente naquele conjunto. No entanto, esta abordagem torna-se complicada, pois este conjunto pode ser muito extenso e, portanto, muito difícil de determinar.

Uma outra forma de determinar as soluções não dominadas, consiste em identificar, em fases distintas e com processos diferentes, as soluções não dominadas suportadas extremas e as não suportadas. Podem, então, considerar-se dois tipos de métodos para determinar estas soluções, cada um correspondendo a cada tipo de soluções.

5.3.1. Métodos para determinar soluções não dominadas suportadas extremas

Nesta secção serão descritos quatro métodos para determinar soluções não dominadas suportadas extremas, dos quais três foram apresentados em 1985 por Henig [15].

5.3.1.1. Método de Henig adaptado a partir de Shin e Hartley

Este método consiste em gerar uma sequência finita $\{\beta_k\}$, $1 = \beta_1 \geq \beta_k > \beta_{k+1} \geq 0$ em paralelo com a lista $\{(c_k^1, c_k^2)\}$, onde (c_k^1, c_k^2) é o valor do vértice associado a cada $\beta \in (\beta_k, \beta_{k+1})$. Este método é uma adaptação do utilizado no algoritmo desenvolvido por Shin [24] e Hartley [14], para um modelo de programação dinâmica estocástica.

Começa-se por determinar o caminho $p^* \in P$, tal que $c_1^1 = c^1(p^*) \leq c^1(p)$ para todo o $p \in P$ e $c_1^2 = c^2(p^*) \leq c^2(p)$ com $p \in P$ e $c^1(p^*) = c^1(p)$. Este caminho é o *lexicograficamente mais curto* relativamente a $\beta_1 = 1$ e pode ser determinado aplicando duas vezes um algoritmo de caminho mais curto : uma para determinar todos os caminhos mais curtos relativamente ao primeiro objectivo e em seguida, entre todos aqueles caminhos, determinar o mais curto relativamente ao segundo objectivo.

Depois, determina-se a árvore dos caminhos mais curtos de qualquer nó para t , relativamente ao primeiro objectivo ($c^1(p_j)$, $j \in N$) e, naquela árvore, determinam-se também os valores dos caminhos mais curtos para t em relação ao segundo objectivo ($c^2(p_j)$, $j \in N$).

Dado β_k suponhamos que o valor do caminho mais curto de j para t , associado a β_k é

$$(c^1(p_j), c^2(p_j)).$$

Seja

$$\alpha = \min \frac{-(c_{ij}^1 + c^1(p_j) - c^1(p_i))}{(c_{ij}^2 + c^2(p_j) - c^2(p_i))} \quad (5.6)$$

em que o mínimo é considerado sobre todos os $(i, j) \in A$, tal que o denominador é negativo.

Seja (i^*, j^*) o arco associado ao valor mínimo da equação (5.6) e actualize-se a árvore dos caminhos mais curtos para t , substituindo o arco com origem em i^* pelo arco (i^*, j^*) . Desta forma, obtêm-se os caminhos mais curtos até t associados a

$$\beta_{k+1} = \frac{1}{1 + \alpha}.$$

Este processo continua até

$$c_{ij}^2 + c^2(p_j) - c^2(p_i) \geq 0, \quad \forall (i, j) \in A$$

indicando que foi determinado o caminho lexicograficamente mais curto associado a $\beta = 0$.

5.3.1.2. Método de Henig semelhante ao de Denardo

Este método consiste em gerar uma sequência finita $\{\beta_k\}$ com $0 = \beta_0 < \beta_k < \beta_1 = 1$ e a lista que lhe é associada $\{(c_k^1, c_k^2)\}$ dos valores dos caminhos não dominados extremos. Este método assemelha-se ao utilizado por Denardo [11] para determinar um multiplicador de Lagrange, numa rede com uma restrição adicional.

Inicialmente, determinam-se os caminhos lexicograficamente mais curtos de s para t , associados a $\beta_0 = 0$ e $\beta_1 = 1$. Estes valores denominam-se por (c_0^1, c_0^2) e (c_1^1, c_1^2) .

Seja $L = \{(c_0^1, c_0^2), (c_1^1, c_1^2)\}$ e $S = \emptyset$. Na iteração principal, calcula-se

$$\alpha = \frac{c_{\beta 2}^1 - c_{\beta 1}^1}{c_{\beta 1}^2 - c_{\beta 2}^2}$$

onde $(c_{\beta 1}^1, c_{\beta 1}^2)$ e $(c_{\beta 2}^1, c_{\beta 2}^2)$ são os dois primeiros elementos de L .

Seja $(c_{\beta}^1, c_{\beta}^2)$ o valor do caminho mais curto associado a $\beta = 1/(1+\alpha)$. Se $\beta \times c_{\beta}^1 + (1-\beta) \times c_{\beta}^2 = \beta \times c_{\beta 1}^1 + (1-\beta) \times c_{\beta 1}^2$, então transfere-se $(c_{\beta 1}^1, c_{\beta 1}^2)$ de L para S . Caso contrário, insere-se $(c_{\beta}^1, c_{\beta}^2)$ em segundo lugar na lista L . Em ambos os casos, a iteração principal é repetida até L ser singular. Assim, $S \cup L$ contém os valores dos vértices do problema.

5.3.1.3. Método de Henig que utiliza a operação Ext

Seja G_i o conjunto de valores dos caminhos não dominados suportados extremos de i para t . Note-se que se o arco (i, j) e o caminho $p_j \in P_j$ formam um caminho mais curto de i para t , associado a β , então p_j é um caminho mais curto de j para t , associado a β . Isto verifica a seguinte equação :

$$G_i = \text{Ext} \{ \bar{G}_i \}, \quad i \in \mathbf{N}, \quad G_t = \{ (0, 0) \}$$

onde

$$\bar{G}_i = \{ (c_{ij}^1, c_{ij}^2) + G_j : j \in \mathbf{N} \}.$$

A operação **Ext** refere-se à escolha dos pontos não dominados extremos, que pertencem ao contorno convexo de \bar{G}_i .

O processo consiste em determinar os conjuntos G_i , com $i \in \mathbf{N}$, em que i varia entre t e s . No fim, G_s contém os valores dos caminhos não dominados extremos do problema.

5.3.1.4. Versão do método NISE apresentada por Cohon

Esta versão, que foi apresentada por Cohon [9] em 1978, resolve, em cada iteração, um problema de caminho mais curto com um único objectivo, cuja função objectivo é uma combinação convexa das funções objectivo originais. Desta forma, cada solução determinada é não dominada para o problema P1. Portanto, este método consiste na resolução do seguinte problema (que é P1 modificado) :

$$\text{Min } Z = w_1 \cdot Z_1 + w_2 \cdot Z_2 \quad (5.7)$$

sujeito a

$$(5.2) - (5.5)$$

em que

$$w_1 = \frac{|Z_2(A) - Z_2(B)|}{|Z_1(A) - Z_1(B)|}$$

$$w_2 = 1$$

$Z_i(x)$ é o valor do i -ésima função objectivo associado à solução x

$Z(A)$ e $Z(B)$ são soluções não dominadas adjacentes.

Em cada iteração do método é calculado o erro máximo para a aproximação, o qual é utilizado para determinar o esquema de ponderação a ser usado para determinar a próxima solução não inferior. Por outro lado, pode-se controlar a exactidão da aproximação, através de um critério de erro predeterminado.

5.3.2. Métodos para determinar soluções não dominadas não suportadas

Se para determinar as soluções não dominadas suportadas existem alguns métodos, para determinar as não suportadas já isso não acontece. De facto, existem muito poucos métodos para determinar estas últimas soluções, tornando-se a sua identificação o “calcanhar de Aquiles” dos problemas de caminho mais curto com vários objectivos.

Apesar disto, será feito referência a dois métodos : um que se baseia no método NISE com restrições e o outro no problema dos k caminhos mais curtos.

5.3.2.1. Método NISE com restrições adicionais

O método NISE, como foi apresentado antes, apenas consegue determinar as soluções não dominadas suportadas extremas. Desta forma, este método consegue determinar as soluções A , B , C e D , mas não as soluções que possam existir nos interiores dos triângulos $\Delta(A,B,P)$, $\Delta(B,C,Q)$ e $\Delta(C,D,R)$ — Fig. 4.

No entanto, Current et al. [10] apresentaram um processo para determinar soluções não dominadas que se encontrem nas zonas de desníveis de dualidade, utilizando uma técnica que consiste em resolver um problema de caminho mais curto com um só objectivo, com restrições adicionais. Neste caso, recomenda-se, sempre que possível, a redução da zona de desnível de dualidade a pesquisar.

Suponha-se, por exemplo, que se pretende pesquisar soluções não dominadas na zona de desnível de dualidade representada pelo triângulo $\Delta(B,C,Q)$ na Fig. 4. Esta zona de desnível de dualidade encontra-se representada isoladamente na Fig. 5.

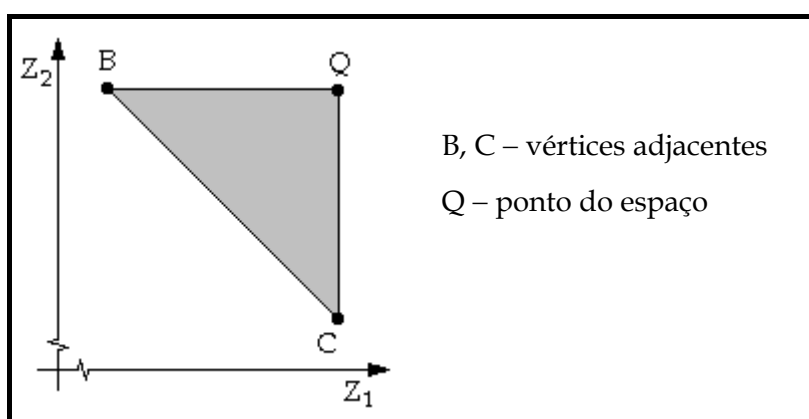


Fig. 5 – Bi-objectivo : representação de uma Zona de Desnível de Dualidade.

Para tal, pode-se resolver, por exemplo, o seguinte problema de caminho mais curto :

$$[\text{P2}] \quad \text{Min } Z = w_1 \cdot Z_1 + w_2 \cdot Z_2 \quad (5.8)$$

sujeito a

$$(5.2) - (5.5)$$

$$Z_1 \leq Z_1(C) - \varepsilon \quad (5.9)$$

em que

$$w_1 = \frac{|Z_2(Q') - Z_2(B)|}{|Z_1(Q') - Z_1(B)|} \text{ é o valor absoluto do declive do segmento } \overline{BQ'}$$

(onde Q' é um ponto pertencente ao segmento \overline{CQ})

$$w_2 = 1$$

ε = um pequeno valor positivo.

Este problema corresponde a minimizar a função de somas pesadas, para a qual a linha que passa por B e Q' é uma linha de custo constante; ou seja, o gradiente desta função é perpendicular à linha referida.

Note-se que se poderia resolver outros problemas, bastando para tal alterar as restrições impostas : $Z_2 \leq Z_2(B) - \varepsilon$ (em vez da restrição (5.9)) e associar w_1 ao declive do segmento $\overline{CQ''}$ (onde Q'' é um ponto pertencente ao segmento \overline{BQ}).

Suponha que ao resolver-se o problema P2, com $Q' = Q$, foi determinada a solução E (Fig. 6). Então são formadas novas zonas de pesquisa, pertencentes à zona de desnível de dualidade analisada, de acordo com a solução encontrada (região a cheio na Fig. 6). Por outro lado, abaixo da recta que passa por E, paralela ao segmento \overline{BQ} , não foram encontradas soluções e qualquer solução que se encontre no rectângulo que contém E e Q é dominada por E.

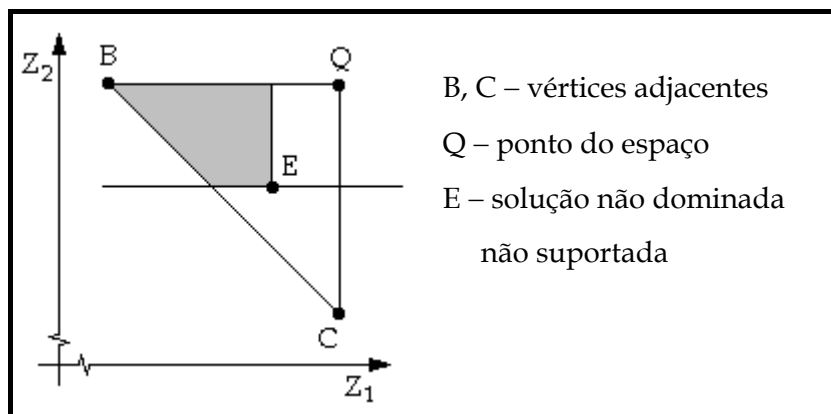


Fig. 6 – Bi-objectivo : actualização de uma Zona de Desnível de Dualidade.

Note-se que as soluções determinadas ao resolver P2 encontram-se no interior da referida zona de desnível de dualidade, uma vez que :

- abaixo da recta que contém o segmento \overline{BC} não existem soluções (B e C são vértices adjacentes),
- antes de ser identificada qualquer solução acima do segmento \overline{BQ} , terá que ser determinada a solução B (portanto, esta solução serve de “vigia”),
- P2 não determina soluções que se encontrem à direita da recta que passa pelo segmento \overline{CQ} nem a solução C, devido à restrição (5.9).

Refira-se que pode continuar-se a pesquisa de soluções nesta zona de desnível de dualidade, resolvendo o problema P2 em que a restrição (5.9) é substituída por $Z_1 \leq Z_1(E) - \varepsilon$.

5.3.2.2. Utilizando algoritmos dos k caminhos mais curtos

Estes métodos são utilizados para determinar as soluções não dominadas que se encontram no interior de uma dada zona de desnível de dualidade (não suportadas). A função objectivo utilizada é a mesma que é utilizado no método NISE com restrições, e que é construída a partir dos dois vértices que caracterizam aquela zona de desnível de dualidade.

Os algoritmos utilizados para determinar os k caminhos mais curtos são extremamente eficientes computacionalmente e identificam, muito rapidamente, todas as soluções admissíveis, segundo uma dada direcção de pesquisa. Desta forma, são determinadas todas as soluções não dominadas não suportadas de uma qualquer zona de desnível de dualidade.

Este processo de pesquisa resolve um problema de determinação dos k caminhos mais curtos correspondente ao seguinte problema :

$$\text{Min } Z = w_1 \cdot Z_1 + w_2 \cdot Z_2 \quad (5.10)$$

sujeito a

$$(5.2) - (5.5)$$

em que

$$w_1 = \frac{|Z_2(A) - Z_2(B)|}{|Z_1(A) - Z_1(B)|}$$

$$w_2 = 1$$

$Z(A)$ e $Z(B)$ são os vértices que definem a zona de desnível de dualidade.

Este processo é executado até se determinar um caminho, cujo valor da função objectivo seja maior ou igual do que o valor da função objectivo associado ao ponto do espaço correspondente ao terceiro vértice do triângulo que representa a zona de desnível de dualidade escolhida.

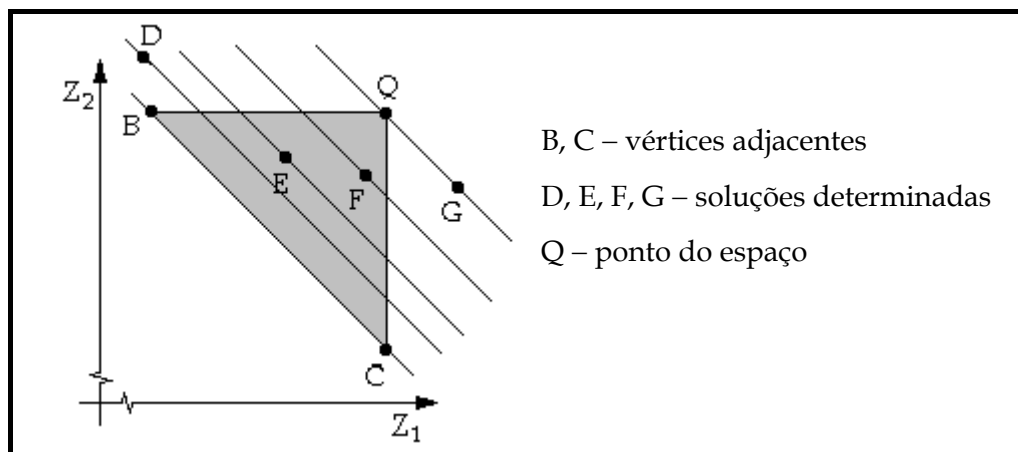


Fig. 7 – Bi-objectivo : utilização dos k caminhos mais curtos.

Para o exemplo apresentado na Fig. 7 o referido vértice é o ponto Q — esta zona de desnível de dualidade é representada pelo triângulo $\Delta(B,C,Q)$.

Ao aplicar-se o processo descrito à zona de desnível de dualidade representada na Fig. 7, em que w_1 é o valor absoluto do declive do segmento \overline{BC} , são determinadas as soluções B, C (ou C, B), D, E, F e G, por esta ordem. O processo termina quando se determina a solução G, pois o valor da função objectivo correspondente a esta solução é igual ao do ponto Q. Analisando as características das soluções determinadas, verifica-se que :

- B e C são soluções não dominadas suportadas extremas,
- E e F são não dominadas não suportadas (pertencem à zona de desnível de dualidade),
- D não pertence a esta zona de desnível de dualidade, mas à definida por A e B (Fig. 4).
- G é dominada pela solução C.

5.4. Algoritmos interactivos para determinar a “melhor” solução de compromisso

Vários autores, como Hansen [13], Clímaco e Martins [5], Martins [18] e Henig [15] apresentaram algoritmos para determinar todas as soluções não dominadas de um problema de caminho mais curto bi-objectivo. No entanto, todos eles reconhecem a complexidade computacional do problema, e por isso, com excepção para Martins, propuseram algoritmos para determinar uma aproximação do conjunto de soluções não dominadas. Por outro lado, nenhum destes autores recomendou a interacção directa com o AD.

De uma maneira geral, os algoritmos existentes consistem em duas fases, as quais correspondem aos dois tipos de soluções não dominadas existentes (Fig. 4) :

1ª fase : determinar alguns vértices (ou todos caso o AD deseje), minimizando combinações convexas das duas funções objectivo envolvidas,

2ª fase : determinar algumas soluções não dominadas não suportadas, correspondentes a zonas de desníveis de dualidade onde se localizam as soluções mais de acordo com as preferências do AD.

Desta forma, os algoritmos distinguem-se uns dos outros, pelas técnicas utilizadas em cada fase, podendo alguns deles utilizar a mesma técnica na mesma fase.

Todos os algoritmos que irão ser analisados baseiam-se em sistemas interactivos de apoio à decisão, e têm por objectivo reduzir o esforço computacional necessário à identificação de uma solução final pois, desta forma, não é necessário que se determinem todas as soluções não dominadas do problema, mas apenas um subconjunto, o qual irá conter a “melhor” solução de compromisso.

5.4.1. Algoritmo de Current, ReVelle e Cohon

Este algoritmo, que foi apresentada por Current et al. [10] em 1988, utiliza, na primeira fase do processo, a versão do algoritmo NISE, que corresponde ao método apresentado por Cohon (ver secção 5.3.1.4), e na segunda fase, também o algoritmo NISE, mas agora correspondente à versão com restrições adicionais (ver secção 5.3.2.1).

No entanto, o algoritmo NISE foi modificado de forma a identificar um subconjunto do conjunto de soluções não dominadas através da interacção com o AD. Desta forma, o algoritmo concentra a procura de soluções eficientes em áreas, do conjunto das soluções não dominadas, que parecem mais prometedoras ao AD.

Em cada iteração o AD é interrogado para indicar, caso o deseje, se alguma região do conjunto não inferior pode ser eliminada em posteriores pesquisas (por não lhe interessar), para estabelecer a direcção na pesquisa de novas soluções no Contorno Convexo ou para escolher as Zonas de Desníveis de Dualidade a serem analisadas. Desta forma, as soluções que pertencem às regiões eliminadas não irão ser consideradas. A eliminação de regiões pode ser baseada no valor de uma das funções objectivo ou na importância dos compromissos entre os objectivos na região.

Na primeira fase do algoritmo aplica-se o método NISE apresentado por Cohon, podendo-se determinar todos os vértices do problema, caso o AD o deseje. Nas duas primeiras iterações, são identificados os vértices que optimizam separadamente cada função objectivo. Nas restantes, aplica-se o algoritmo NISE associado aos dois vértices adjacentes escolhidos pelo AD. Esta fase termina quando todos os vértices forem encontradas, ou então, por indicação do AD, por achar que as soluções que possam existir nas regiões não analisadas não têm interesse para ele (decisão baseada na sua experiência).

Note-se que o algoritmo NISE permite conhecer, numa dada região, se não existem mais vértices, o que acontece quando a solução encontrada é um dos vértices que serviu para construir a função objectivo a minimizar, o que significa que aqueles vértices definem uma Zona de Desnível de Dualidade.

Para identificar as soluções não dominadas que possam existir numa determinada Zona de Desnível de Dualidade, este algoritmo utiliza uma versão do método NISE com restrições adicionais. No entanto, como a pesquisa de soluções nestas regiões requer um elevado esforço computacional, é recomendada a redução da Zona de Desnível de Dualidade, eliminando a região que represente, para o AD, compromissos indesejáveis entre os objectivos (Fig. 8).

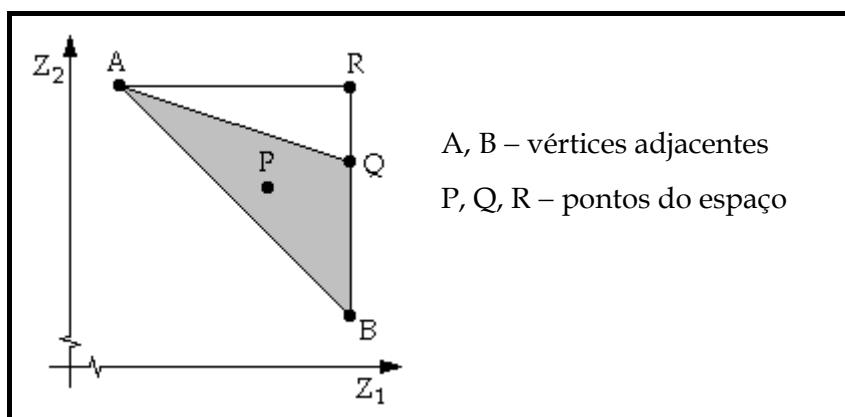


Fig. 8 – Bi-objectivo : redução de uma Zona de Desnível de Dualidade.

Analisando a Fig. 8, pode-se concluir que o aumento necessário no objectivo Z_1 para diminuir (melhorar) em uma unidade o objectivo Z_2 , é maior ao deslocar-se de A para qualquer ponto P da zona de desnível de dualidade, do que ao deslocar-se de A para B, já que o valor absoluto do declive do segmento \overline{AP} será menor que o do segmento \overline{AB} . Assim sendo, o compromisso de Z_1 para Z_2 é limitado pelos declives dos segmentos \overline{AR} e \overline{AB} .

Desta forma, pode-se reduzir a região de pesquisa nesta zona de desnível de dualidade, através da indicação, por parte do AD, do compromisso máximo aceitável de Z_1 para Z_2 . Supondo-se, por exemplo, que o AD indica um compromisso máximo aceitável de 3, então qualquer solução que se encontre acima do segmento \overline{AQ} pode ser eliminada, uma vez que o declive desta linha é $-1/3$. Ou seja, qualquer solução que se encontre acima daquela linha, terá um compromisso superior a 3.

Para determinar a solução não dominada abaixo do segmento \overline{AQ} (em $\Delta(A,B,Q)$), deve ser aplicado o algoritmo correspondente ao método NISE com restrições associado às soluções A e B. Este problema corresponde a minimizar a função de somas pesadas, para a qual a linha que passa pelo segmento \overline{AQ} é uma linha de custo constante; ou seja, o gradiente desta função é perpendicular ao segmento referido.

Como resultado da resolução deste problema foi determinada a solução C, a partir da qual se actualizam as regiões de interesse onde podem existir mais soluções não dominadas, pois C domina qualquer uma que se situe no rectângulo em que C e R são dois dos seus vértices; isto é, qualquer solução x tal que $Z_1(x) \geq Z_1(C)$ e $Z_2(x) \geq Z_2(C)$ é dominada pela solução C. Por outro lado, não existem soluções abaixo da linha paralela ao segmento \overline{AQ} que passa por C (Fig. 9).

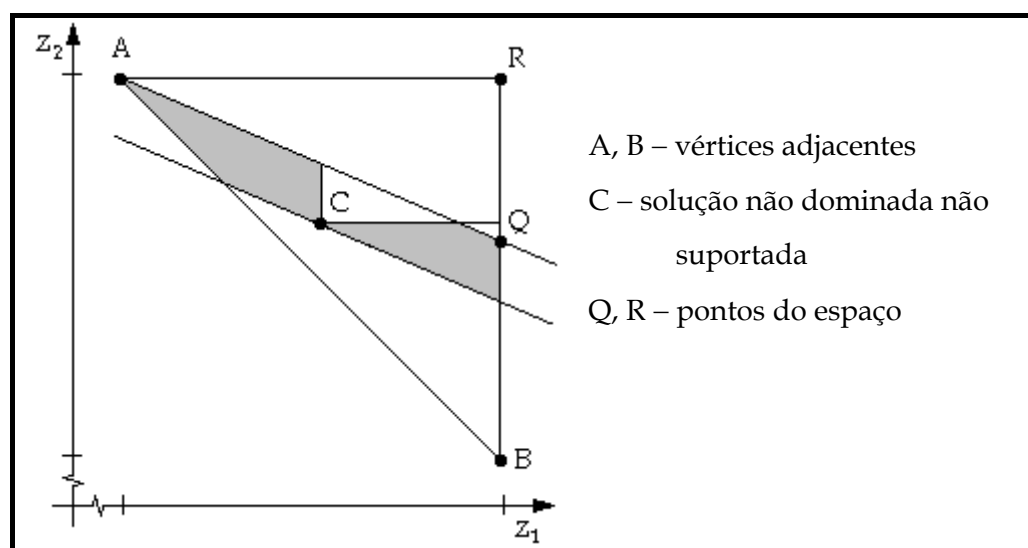


Fig. 9 – Bi-objectivo : actualização de uma Zona de Desnível de Dualidade.

Para encontrar uma outra solução, aplica-se novamente o algoritmo NISE com restrições, correspondente ao problema P2, em que a restrição (5.9) é substituída por $Z_1 \leq Z_1(C) - \varepsilon$, caso se pretenda analisar a região à esquerda de C, ou então, o mesmo problema P2 com $w_1 = 0$ (logo, $w_2 = 1$), se for pretensão analisar a região à direita de C. No primeiro caso, o resultado é a determinação da solução A, o que provoca o término do processo, já que esta solução encontra-se sobre o segmento \overline{AQ} (tem um compromisso igual a 3).

5.4.2. Algoritmo de Clímaco e Rodrigues

Este algoritmo, que foi apresentado por Clímaco e Rodrigues [8] em 1988, surgiu na sequência de duas abordagens associadas ao problema de caminho mais curto bi-critério :

- 1) Clímaco e Martins [5], Martins [17] e Clímaco [6] desenvolveram um método para gerar todos os caminhos não dominados, baseado num algoritmo para determinar os k caminhos mais curtos e num adequado teste de não dominância;
- 2) Current et al. [10] apresentaram um método interactivo como alternativa aos métodos que geram todos os caminhos não dominados.

Este algoritmo foi apresentado em duas versões, correspondendo a duas maneiras diferentes de analisar as zonas de desníveis de dualidade : uma utiliza apenas parte daquela zona (tal como em [10]) e a outra utiliza toda a zona.

A primeira fase do processo é comum às duas versões e, tal como no método proposto por Current et al. [10], utilizam a mesma versão do algoritmo NISE, apresentado por Cohon, para determinar alguns ou todos os vértices.

Na segunda fase do processo, é utilizado o algoritmo para determinar os k caminhos mais curtos descrito por Martins [17], para determinar as soluções não dominadas existentes nas zonas de desníveis de dualidade. Para tal, são apresentados dois processos para o fazer, associados a cada versão do algoritmo :

1^a) tal como acontece no método de Current et al. [10], é solicitado ao AD que fixe o compromisso máximo aceitável de Z_1 para Z_2 . Por exemplo, se apenas se está interessado em soluções contidas no triângulo $\Delta(A,B,Q)$, utiliza-se o algoritmo referido cuja função objectivo a minimizar é uma soma pesada, para a qual a linha que passa pelo segmento \overline{AQ} é de custo constante (gradiente da função perpendicular ao segmento \overline{AQ}) — Fig. 8.

2^a) funciona da mesma maneira que a versão anterior, mas considera-se agora toda a zona de desnível de dualidade (o AD não impõe qualquer compromisso). Por exemplo, se se está interessado em todas as soluções contidas no triângulo $\Delta(A,B,R)$, utiliza-se o mesmo algoritmo, mas agora a função objectivo utilizada é uma soma pesada, para a qual a linha que pelo segmento \overline{AB} é de custo constante (gradiente perpendicular a \overline{AB}) — Fig. 8.

Excluindo o caso em que a pesquisa de soluções termina quando o AD desejar, esta pesquisa só termina, na primeira versão, quando for encontrada a solução A , pois é a última solução dentro da região pretendida (Fig. 8), e na segunda, quando se determina uma solução que se encontre na zona onde se tem a certeza que não existem mais soluções não dominadas, o que acontece quando se passar para além da linha carregada (Fig. 10) — esta linha é paralela ao segmento \overline{AB} e contém o ponto pertencente a alguma região de interesse que se encontra a maior distância do segmento \overline{AB} .

Desta forma, enquanto na primeira versão a solução D nunca é determinada, na segunda isso já acontece (Fig. 10). No entanto, a segunda versão parece ser mais adequada quando se pretende pesquisar em pequenas zonas de desníveis de dualidade, uma vez que neste caso, o tempo que se ganha ao reduzir a área de pesquisa é insignificante, atendendo aos tempos conseguidos com os algoritmos existentes para determinar os k caminhos mais curtos na determinação de tais caminhos.

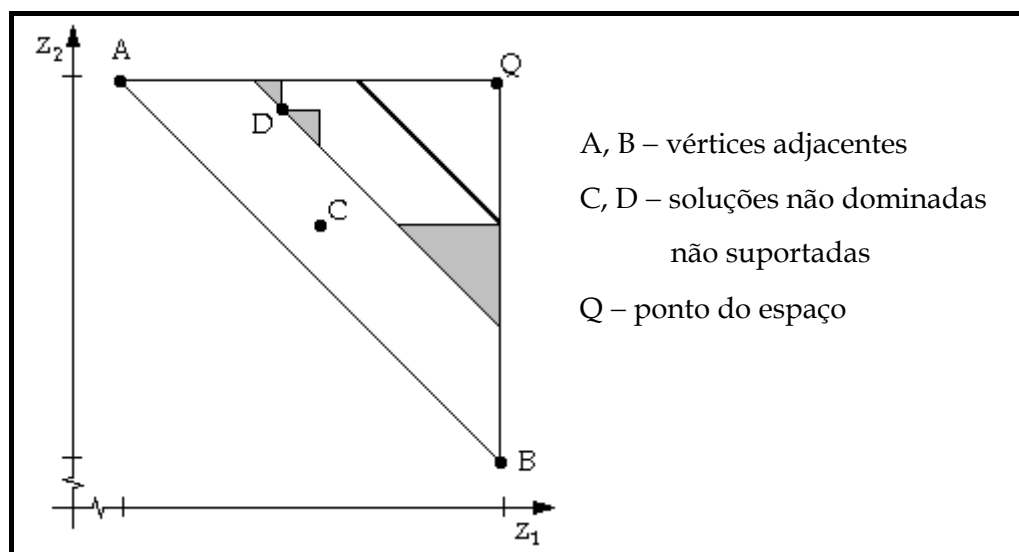


Fig. 10 – Bi-objectivo : soluções de uma Zona de Desnível de Dualidade.