

CHAPITRE 7

Tentative de réponse à la
NP-complétude

Hugo SALOU MPI*

Dernière mise à jour le 26 février 2023

Table des matières

0	Motivation	2
1	Problèmes d'optimisation	2
2	Algorithmes d'approximations	4
3	<i>Branch and Bound</i> — Séparation et évaluation	7
	Annexe A. Programmation dynamique	10

0 Motivation

On considère le problème **NP**-complet du *voyageur de commerce* : étant donné un graphe pondéré G quel est le tour de longueur¹ minimale *i.e.* quelle est la permutation de sommets telle que la longueur totale est minimale.

On se ramène à un problème de décision : étant donné une constante $K \in \mathbb{R}$, existe-t-il un chemin de longueur inférieure à K .

Un algorithme glouton, allant d'un sommet à son voisin le plus proche, ne permet pas de résoudre ce problème en complexité polynomiale.

On ne cherche plus le « tour optimal » mais on cherche une solution proche : on veut trouver une constante ρ telle que, quelque soit l'entrée, le chemin obtenu est de longueur inférieure à ρ fois la longueur optimale.

1 Problèmes d'optimisation

Dans un premier temps, on s'intéresse à un problème où l'on cherche à minimiser quelque chose. On réalise la transformation réalisée dans la partie précédente : étant donné un seuil K , on est ce que la valeur est inférieure à K . On se ramène donc à un problème de décision.

Définition : Soit $Q \subseteq \mathcal{E} \times \mathcal{S}$ un problème. Soit $\text{opt} \in \{\min, \max\}$. On dit que Q est un *problème d'optimisation* (*i.e.* problème de minimisation, maximisation), si pour toute entrée $e \in \mathcal{E}$, il existe

- un ensemble $\text{sol}(e)$ de solutions,
- une fonction $c_e : \text{sol}(e) \rightarrow \mathbb{R}^+$,

tels que $c_e^* = \text{opt}\{c_e(s) \mid s \in \text{sol}(e)\}$ est bien défini, et

$$\forall s \in \text{sol}(e), \quad (e, s) \in Q \implies c_e(s) = c_e^*.$$

On nomme :

- $\text{sol}(e)$ l'ensemble des solutions pour l'entrée e ,
- c_e la fonction objectif,
- c_e^* la valeur optimale (minimale ou maximale),
- pour une solution $s \in \text{sol}(e)$, $c_e(s)$ est appelée la valeur de la solution.

On appelle *solution optimale* une solution de valeur optimale.

EXEMPLE :

On considère le problème

- Entrée** : $G = (S, A)$ un graphe orienté fortement connexe, $s \in S$, et $p \in S$
- Sortie** : un plus court chemin (en nombre d'arcs) de s à p dans G .

Soit l'entrée ci-dessous.

1. poids des arrêtes total

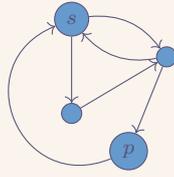


FIGURE 1 – Entrée du problème du plus court chemin

L'ensemble $\text{sol}(e)$ est l'ensemble (infini) des chemins de s à p , et $c_e(\gamma) = |\gamma|$ (la longueur du chemin γ). On vérifie qu'il existe un chemin de s à p , donc $\{c_e(s) \mid s \in \text{sol}(e)\}$ est une partie de \mathbb{N} non vide, elle admet donc un minimum.

Définition : Le problème de décision associé à un problème d'optimisation est le problème obtenu en ajoutant une constante aux entrées et en demandant en sortie s'il est possible de dépasser cette constante.

EXEMPLE :

Étant donné le problème d'optimisation

$$Q_O : \begin{cases} \text{Entrée} & : e \in \mathcal{E}_{Q_O} \\ \text{Sortie} & : \arg \text{opt}_{s \in \text{sol}(e)} c_e(s), \end{cases}$$

on définit le problème de décision associé

$$Q : \begin{cases} \text{Entrée} & : e \in \mathcal{E}_{Q_O}, K \in \mathbb{R}^+ \\ \text{Sortie} & : \text{existe-t-il } s \in \text{sol}(e) \text{ tel que } c_e(s) \bowtie K \end{cases}$$

avec $\bowtie = \leq$ si $\text{opt} = \min$ et $\bowtie = \geq$ si $\text{opt} = \max$.

EXEMPLE :

Avec l'exemple précédent (plus court chemin), le problème de décision associé est

$$\begin{cases} \text{Entrée} & : G = (S, A) \text{ un graphe orienté fortement connexe, } (p, s) \in S^2, \text{ et } K \in \mathbb{R}^+ \\ \text{Sortie} & : \text{Existe-t-il un chemin de } s \text{ à } p \text{ dans } G \text{ de longueur inférieure ou égale à } K? \end{cases}$$

EXEMPLE :

On considère le problème KNAPSACK de décision défini comme

$$\begin{cases} \text{Entrée} & : \text{Un entier } n \in \mathbb{N}, (p_1, \dots, p_n) \in (\mathbb{N}^*)^n, (v_1, \dots, v_n) \in (\mathbb{N}^*)^n, P \in \mathbb{N} \text{ et } K \in \mathbb{N} \\ \text{Sortie} & : \text{Existe-t-il } I \subseteq \llbracket 1, n \rrbracket \text{ telle que } \sum_{i \in I} p_i \leq P \text{ et } \sum_{i \in I} v_i \geq K? \end{cases}$$

Le problème d'optimisation associé est KNAPSACK_O défini comme

$$\begin{cases} \text{Entrée} & : \text{Un entier } n \in \mathbb{N}, (p_1, \dots, p_n) \in (\mathbb{N}^*)^n, (v_1, \dots, v_n) \in (\mathbb{N}^*)^n, \text{ et } P \in \mathbb{N} \\ \text{Sortie} & : I \subseteq \llbracket 1, n \rrbracket \text{ tel que } \sum_{i \in I} p_i \leq P \text{ et maximisant } \sum_{i \in I} v_i. \end{cases}$$

REMARQUE :

Soit Q_O un problème d'optimisation et Q le problème de décision associé. Étant donné un algorithme \mathcal{A}_O pour Q_O , on fabrique l'algorithme \mathcal{A} suivant résolvant Q .

Algorithme 1 Solution à un problème de seuil

Entrée e une entrée de Q_O et K un seuil

1: **retourner** $c_e(\mathcal{A}_O) \bowtie K$ \triangleright où \bowtie est \geq pour si opt est max, et \leq si opt est min

Ainsi, le problème Q_O est plus difficile à résoudre que le problème Q de décision associé. Alors, lorsque le problème de décision Q associé à un problème d'optimisation S_O est **NP**-difficile, c'est mal engagé.

2 Algorithmes d'approximations

REMARQUE (Vocabulaire) :

On fixe dans la suite un problème d'optimisation Q , on note $\text{OPT}(e)$ la valeur optimale pour une entrée e .

Définition (Algorithme d'approximation pour un problème de maximisation) : On dit d'un algorithme $\mathcal{A} : \mathcal{E}_Q \rightarrow \mathbb{R}^+$ qu'il approxime un problème Q de maximisation avec un ratio d'approximation $\rho < 1$ dès lors que

$$\forall e \in \mathcal{E}_Q, \quad \mathcal{A}(e) \geq \rho \cdot \text{OPT}(e).$$

On dit alors que l'algorithme \mathcal{A} est une ρ -approximation (standard).

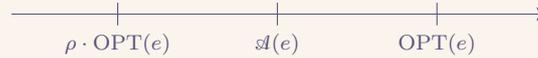


FIGURE 2 – Algorithme d'approximation pour un problème de maximisation

Définition (Algorithme d'approximation pour un problème de minimisation) : On dit d'un algorithme $\mathcal{A} : \mathcal{E}_Q \rightarrow \mathbb{R}^+$ qu'il approxime un problème Q de minimisation avec un ratio d'approximation $\rho > 1$ dès lors que

$$\forall e \in \mathcal{E}_Q, \quad \mathcal{A}(e) \leq \rho \cdot \text{OPT}(e).$$

On dit alors que l'algorithme \mathcal{A} est une ρ -approximation (standard).

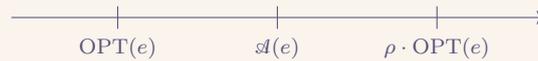


FIGURE 3 – Algorithme d'approximation pour un problème de minimisation

Dans la définition suivante, on suppose connu une fonction $\text{Pire}(e)$ donnant la pire valeur de solution pour une entrée e .

Définition : On dit qu'un algorithme $\mathcal{A} : \mathcal{E}_Q \rightarrow \mathbb{R}$ est une ρ -approximation différentielle dès lors que

$$\frac{|\mathcal{A}(e) - \text{Pire}(e)|}{|\text{Pire}(e) - \text{OPT}(e)|} \geq \rho.$$

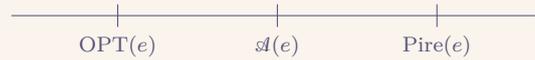


FIGURE 4 – ρ -approximation différentielle

REMARQUE :

Dans le cadre d'un problème de minimisation, on ne calcule pas $\text{OPT}(e)$ en général. On minore $\text{OPT}(e)$, et alors

$$\frac{\mathcal{A}(e)}{\text{OPT}(e)} \leq \underbrace{\frac{\mathcal{A}(e)}{B}}_{\rho}.$$

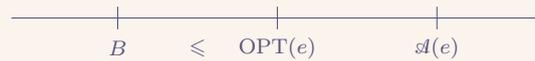


FIGURE 5 – Calcul de $\text{OPT}(e)$

EXEMPLE :

Soit $C \in \mathbb{N}^*$. On rappelle le problème **STABLE**, restreint à un graphe de degré maximal C :

STABLE : $\left\{ \begin{array}{l} \text{Entrée} : G = (S, A) \text{ une graphe tel que } 0 \neq \Delta(G) \leq C \\ \text{Sortie} : \text{un stable de } G \text{ de cardinal maximal} \end{array} \right.$

où $X \subseteq S$ est un stable si pour tout $(u, v) \in X^2$, $\{u, v\} \notin A$, et $\Delta(G) = \max_{v \in S} \deg_G(v)$ est le degré du graphe.

Algorithme 2 Algorithme gloutin de recherche de stables

Entrée $G = (S, A)$ un graphe

- 1: $S' \leftarrow \emptyset$
- 2: **tant que** $S \neq \emptyset$ **faire**
- 3: $v^* = \arg \min_{v \in S} \deg_G(v)$ \triangleright les degrés sont modifiés à chaque itération
- 4: $S' \leftarrow S' \cup \{v^*\}$
- 5: $S \leftarrow S \setminus (\{v^*\} \cup \text{voisins}(v^*))$
- 6: $A \leftarrow$ restriction de A à S
- 7: **retourner** S'

Cet algorithme n'est pas correct, la figure ci-après en est un contre-exemple.

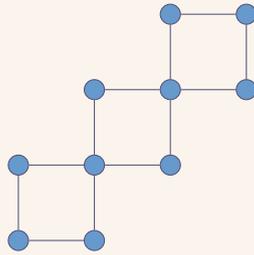


FIGURE 6 – Contre-exemple à l'algorithme 2

Propriété : L'algorithme 2 est une $\frac{1}{\Delta(G)}$ -approximation.

Preuve :

Soit S' la réponse de l'algorithme. Soit S^* la solution optimale. On a, par terminaison de l'algorithme

$$\forall v \in S \setminus S', \exists v' \in S', \{v, v'\} \in A$$

car l'algorithme s'arrête. En particulier, $\forall v^* \in S^* \setminus S', \exists v' \in S', \{v, v'\} \in A$. Or, S^* est stable donc, si $v^* \in S^*$ et $v' \in S'$ tels que $\{v^*, v'\} \in A$, alors $v' \notin S^*$. D'où,

$$\forall v^* \in S^* \setminus S', \exists v' \in S' \setminus S^*, \{v^*, v'\} \in A.$$

Par définition de degré d'un graphe, on a $|S^* \setminus S'| \leq \Delta(G) |S' \setminus S^*|$ donc

$$\begin{aligned} |S^*| &= |S^* \cap S'| + |S^* \setminus S'| \\ &\leq |S^* \cap S'| + \Delta(G) |S' \setminus S^*| \\ &\leq \Delta(G) |S^* \cap S'| + \Delta(G) |S' \setminus S^*| \\ &\leq \Delta(G) |S'|. \end{aligned}$$

□

REMARQUE :

Cette preuve ne fait pas d'hypothèses sur le résultat de l'algorithme. Tout algorithme répondant au problème est une $\frac{1}{\Delta(G)}$ -approximation.

EXEMPLE :

On appelle *couverture par sommets* d'un graphe $G = (S, A)$ la donnée d'un ensemble $X \subseteq S$ tel que

$$\forall \{u, v\} \in A, \quad u \in X \text{ ou } v \in X.$$

EXEMPLE :

L'ensemble ● est une couverture par sommets du graphe ci-dessous.

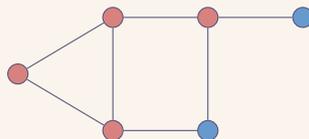


FIGURE 7 – Exemple de couverture par sommets

On considère le problème

Entrée : $G = (S, A)$ un graphe
Sortie : une couverture de cardinal maximal.

Ce problème peut-être résolu à l'aide du calcul de couplages maximal.

Algorithme 3 Calcul d'un couplage maximal (COUPLAGE_{MAXIMAL})

Entrée $G = (S, A)$ un graphe

Sortie C un couplage maximal, *non nécessairement maximum*

```
1:  $C \leftarrow \emptyset$ 
2: tant que  $\exists \{u, v\} \in A$ ,  $u$  libre dans  $C$  et  $v$  libre dans  $C$  faire
3:   Soit  $\{u, v\}$  une telle arête.
4:    $C \leftarrow C \cup \{\{u, v\}\}$ 
5: retourner  $C$ 
```

L'algorithme retourne un couplage maximal d'après la négation de la condition de boucle. On répond donc au problème avec l'algorithme ci-dessous.

Algorithme 4 Approximation de couverture par sommets

Entrée $G = (S, A)$ un graphe

Sortie Une couverture par sommets

```
1:  $C \leftarrow \text{COUPLAGE}_{\text{MAXIMAL}}(G)$ 
2: retourner  $\{u \in S \mid \exists v \in S, \{u, v\} \in C\}$ .
```

L'algorithme retourne une couverture : soit X la valeur retournée pour une entrée $G = (S, A)$. Soit $\{u, v\} \in A$. Si $u \notin X$ et $v \notin X$, alors le couplage C calculé pour l'algorithme n'est pas maximal, on peut y ajouter $\{u, v\}$. Montrons que l'algorithme \mathcal{A} est une 2-approximation du problème « couverture par sommets. »

$$\forall G \in \mathcal{E}, \quad \mathcal{A}(G) \leq 2 \text{OPT}(G).$$

Soit $G \in \mathcal{E}$. Soit X la couverture par sommets calculé par \mathcal{A} sur G , et C le couplage calculé par cet algorithme. Soit X^* la couverture par sommets optimale. Soit donc

$$\varphi : C \longrightarrow X^*$$
$$\{u, v\} \longmapsto \begin{cases} u & \text{si } u \in X^* \\ v & \text{si } v \in X^* \end{cases}$$

Soit $(c_1, c_2) \in C^2$, tels que $\varphi(c_1) = \varphi(c_2)$, alors c_1 et c_2 partagent un sommet, ce qui est absurde (c.f. définition de couplage). Donc φ est injective, d'où $|C| \leq |X^*|$. Or, $\mathcal{A}(X) = |X| = 2|C| \leq 2|X^*| \leq 2 \text{OPT}(X)$.

3 Branch and Bound — Séparation et évaluation

Branch and Bound n'est pas un algorithme, mais une famille d'algorithmes, similairement aux algorithmes diviser pour régner. Ces algorithmes répondent à des problèmes de maximisation. Les algorithmes *Branch and Bound* sont des algorithmes enrichis de trois fonctions :

- une fonction branch de branchement, i.e. découpage en sous-problèmes,
- une fonction valeur donnant un résultat, pas forcément optimal, i.e. elle associe une solution partielle à une solution,
- une fonction bound donnant un majorant de la solution optimale, complétant cette solution partielle.

Avec les deux dernières fonctions, on borne la valeur de la solution optimale.

EXEMPLE (PL et PLNE) :
c.f. DM4.

EXEMPLE :
 On considère le problème **KNAPSACK** :

Entrée : $n \in \mathbb{N}$, $(v_i)_{i \in [1, n]} \in (\mathbb{N}^*)^n$, $(w_i)_{i \in [1, n]} \in (\mathbb{N}^*)^n$, et $P \in \mathbb{N}^*$
Sortie : une allocation I maximale d'objets de somme de poids $(w_i)_{i \in I}$ inférieure ou égale à P .

Dans toute la suite de l'exemple, les $(v_i)_{i \in [1, n]}$ et $(w_i)_{i \in [1, n]}$ sont triés par v_i/w_i décroissants.

Tentative 1. Algorithme glouton.

Algorithme 5 Algorithme glouton $\mathcal{G}^{\mathbb{N}}$ répondant au problème **KNAPSACK**

```

1:  $I \leftarrow \emptyset$ 
2:  $S \leftarrow 0$ 
3: pour  $i \in [1, n]$  faire
4:   si  $S + w_i \leq P$  alors
5:      $I \leftarrow I \cup \{i\}$ 
6:      $S \leftarrow S + w_i$ 
7: retourner  $I$ 

```

Cet algorithme ne donne pas toujours une solution optimale, voici un contre-exemple : $P = 3$, $(v_i) = (1, 2)$ et $(w_i) = (1, 3)$. L'algorithme renvoie $\mathcal{G}^{\mathbb{N}}(E_1) = 1$ mais la solution optimale $\text{OPT}(E_1) = 2$, pour l'entrée E_1 . On peut compléter une solution partielle à l'aide de cet algorithme, on a donc défini la fonction valeur :

Tentative 2. On résout le problème associé dans \mathbb{R} , définit ci-dessous :

$\text{KNAPSACK}_{\mathbb{R}} \left\{ \begin{array}{l} \text{Entrée} : n \in \mathbb{N}, (v_i)_{i \in [1, n]} \in (\mathbb{N}^*)^n, (w_i)_{i \in [1, n]} \in (\mathbb{N}^*)^n, \text{ et } P \in \mathbb{N}^* \\ \text{Sortie} : \arg \max_{x \in S} \sum_{i=1}^n x_i v_i \end{array} \right.$

où $S = \left\{ (x_i)_{i \in [1, n]} \in [0, 1]^n \mid \sum_{i=1}^n x_i w_i \leq P \right\}$. On résout ce problème à l'aide d'un algorithme glouton.

Algorithme 6 Algorithme glouton $\mathcal{G}^{\mathbb{R}}$ répondant au problème **KNAPSACK_R**

```

1:  $S \leftarrow 0$ 
2:  $i \leftarrow 1$ 
3:  $x \leftarrow (0)_{i \in [1, n]}$ 
4: tant que  $i \leq n$  et  $S + w_i < P$  faire
5:    $S \leftarrow S + w_i$ 
6:    $x_i \leftarrow 1$ 
7:    $i \leftarrow i + 1$ 
8: si  $i \leq n$  alors
9:    $x_i \leftarrow \frac{P - S}{w_i}$ 
10:   $S \leftarrow P$ 
11: retourner  $x$ 

```

En notant $\text{OPT}(e)$ une solution optimale à **KNAPSACK**, et $\text{OPT}^{\mathbb{R}}(e)$ une solution optimale à **KNAPSACK_R**, on a $\forall e \in \mathcal{E}, \text{OPT}(e) \leq \text{OPT}^{\mathbb{R}}(e)$. De plus, à faire à la maison, le glouton $\mathcal{G}^{\mathbb{R}}$ donne la solution optimale : on a $\text{OPT}^{\mathbb{R}}(e) = \mathcal{G}^{\mathbb{R}}(e)$.

On branche sur la partie fractionnaire. On considère l'entrée définie dans la table ci-après, avec $P = 20$.

i	1	2	3	4	5	6
v_i	13	16	19	24	3	5
w_i	6	8	10	14	2	5
$\approx v_i/w_i$	2,2	2	1,9	1,7	1,5	1

TABLE 1 – Entrée du problème KNAPSACK

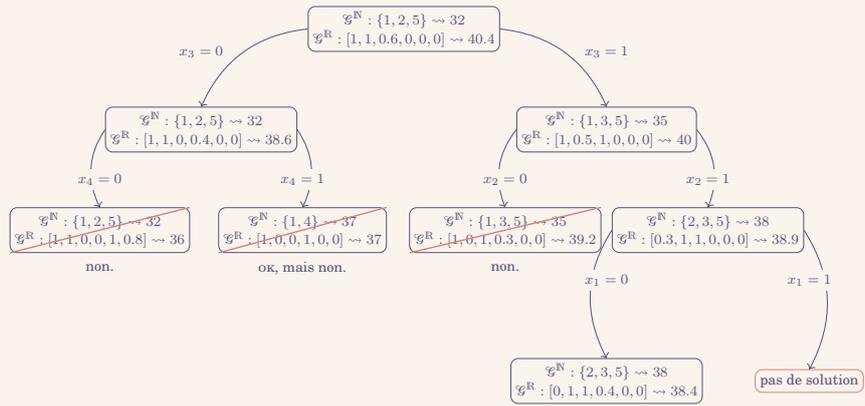


FIGURE 8 – Stratégie *branch and bound* appliquée au problème KNAPSACK

Annexe A. Programmation dynamique

On rappelle le problème KNAPSACK :

$$\begin{cases} \text{Entrée} & : n \in \mathbb{N}, w \in (\mathbb{N}^*)^n, v \in (\mathbb{N}^*)^n, P \in \mathbb{N} \\ \text{Sortie} & : \max_{x \in \{0,1\}^n} \{ \langle x, v \rangle \mid \langle x, w \rangle \leq P \}. \end{cases}$$

On pose

$$\text{SAD}(n, w, v, P) = \max_{x \in \{0,1\}^n} \{ \langle x, v \rangle \mid \langle x, w \rangle \leq P \},$$

et

$$\text{sol}(n, w, v, P) = \{ \langle x, v \rangle \mid \langle x, w \rangle \leq P, x \in \{0, 1\}^n \}.$$

Lorsque $y \in \mathbb{R}^n$, avec $y = (y_1, \dots, y_n)$, on note $\mathbb{R}^{n-1} \ni \tilde{y} = (y_2, y_3, \dots, 0)$. Ainsi, si $n > 0$,

$$\begin{aligned} \text{sol}(n, w, v, P) &= \{ \langle x, v \rangle \mid \langle x, w \rangle \leq P, x \in \{0, 1\}^n \text{ et } x_1 = 0 \} \\ &\quad \cup \{ \langle x, v \rangle \mid \langle x, w \rangle \leq P, x \in \{0, 1\}^n \text{ et } x_1 = 1 \} \\ &= \{ \langle \tilde{x}, \tilde{v} \rangle \mid \langle \tilde{x}, \tilde{w} \rangle \leq P, \tilde{x} \in \{0, 1\}^n \text{ et } x_1 = 0 \} \\ &\quad \cup \{ v_1 + \langle \tilde{x}, \tilde{v} \rangle \mid \langle \tilde{x}, \tilde{w} \rangle \leq P - w_1, x \in \{0, 1\}^n \text{ et } x_1 = 1 \} \\ &= \{ \langle y, \tilde{v} \rangle \mid \langle y, \tilde{w} \rangle \leq P \text{ et } y \in \{0, 1\}^{n-1} \} \\ &\quad \cup \{ v_1 + \langle y, \tilde{v} \rangle \mid \langle y, \tilde{w} \rangle \leq P - w_1 \text{ et } y \in \{0, 1\}^{n-1} \} \end{aligned}$$

D'où, par passage au max, si $n > 0$,

$$\begin{aligned} \text{SAD}(n, w, v, P) &= \max(\\ &\quad \max\{ \langle y, \tilde{v} \rangle \mid \langle y, \tilde{w} \rangle \leq P \text{ et } y \in \{0, 1\}^{n-1} \} \\ &\quad v_1 + \max\{ \langle y, \tilde{v} \rangle \mid \langle y, \tilde{w} \rangle \leq P - w_1 \text{ et } y \in \{0, 1\}^{n-1} \} \\ & \quad) = \max(\text{SAD}(n-1, \tilde{w}, \tilde{v}, P), v_1 + \text{SAD}(n-1, \tilde{w}, \tilde{v}, P - w_1)). \end{aligned}$$

Si $n = 0$, alors $\text{SAD}(0, v, w, P) = 0$.

REMARQUE :

Si on le code *tel quel*, il y aura $\Theta(2^n)$ appels récurrents. Mais, on a $(n+1)(P+1)$ sous-problèmes.

Notons alors, pour n, v, w, P fixés, $(s_{i,j})_{\substack{i \in \llbracket 1, n \rrbracket \\ j \in \llbracket 0, P \rrbracket}}$ tel que

$$s_{i,j} = \text{SAD}(n-i, v|_{\llbracket i+1, n \rrbracket}, w|_{\llbracket i+1, n \rrbracket}, j).$$

On a alors $\text{SAD}(n, v, w, P) = s_{0,P}$. Ainsi, pour $j \in \llbracket 0, P \rrbracket$, $s_{n,j} = 0$; pour $i \in \llbracket 0, n \rrbracket$, $s_{i,0} = 0$;

$$s_{i,j} = \max(s_{i+1,j}, v_{i+1} + s_{i+1,j-w_{i+1}});$$

et, si $w_{i+1} > j$, alors $s_{i,j} = s_{i+1,j}$.

La complexité de remplissage de la matrice est en $\Theta(nP)$ en temps et en espace. On n'a pas prouvé $\mathbf{P} = \mathbf{NP}$, la taille de l'entrée est

- pour un entier n : $\log_2(n)$,
- pour un tableau de n entiers : $n \log_2(n)$,
- pour un tableau de n entiers : $n \log_2(n)$,
- pour un entier P : $\log_2(P)$.

Vis à vis de la taille de l'entrée, la complexité de remplissage est **exponentielle**.