**DS R4.04- 2024**

**Durée : 1h30**

Accès autorisé au supports de cours, TDs et corrections et uniquement aux sites geogebra.org et Google Colab

**Exercice 1 : BikeWorks**

Une entreprise de fabrication de vélos, BikeWorks, produit deux types de vélos : les vélos de montagne et les vélos de route. Chaque vélo nécessite l'utilisation de trois départements pour l'assemblage : le département de cadre, le département de montage et le département de finition. Voici les détails de l'assemblage :

* Un vélo de montagne nécessite 4 heures au département de cadre, 2 heures au département de montage et 1 heure au département de finition.
* Un vélo de route nécessite 3 heures au département de cadre, 1 heure au département de montage et 2 heures au département de finition.

Les coûts des matériaux et des fournitures sont les suivants :

* Le coût des matériaux pour un vélo de montagne est de 50€, et pour un vélo de route est de 30€.
* Les frais généraux sont de 30€ par vélo de montagne et de 25€ par vélo de route.

Les heures de main-d'œuvre disponibles par semaine dans chaque département sont les suivantes :

* Département de cadre : 90 heures
* Département de montage : 40 heures
* Département de finition : 50 heures

Le prix de vente d’un vélo de montagne est 150€ l'unité et celui d’un vélo de route est, 120€ l'unité. L'entreprise souhaite établir un programme de production hebdomadaire qui maximisera ses bénéfices.

a) Ecrivez un programme linéaire ayant pour objectif de maximiser le bénéfice journalier de l’entreprise.

b) Résoudre ce problème graphiquement à l’aide du site geogebra.org. Il faut faire un imprime écran de l’espace de solutions réalisables et le délimiter.

c) Donner la solution optimale si elle existe.

d) Pour répondre à la demande du marché, l'entreprise doit produire au plus 20 vélos de montagne et au plus 18 vélos de route par semaine. L'entreprise est assurée de vendre toute sa production si cette condition est respectée. Que change cette contrainte dans la formulation de votre programme linéaire ?

e) Représenter le nouvel espace de solutions réalisables.

d) Donner la solution optimale si elle existe.

X = nb de vélos de montagne

Y = nb de vélos de route

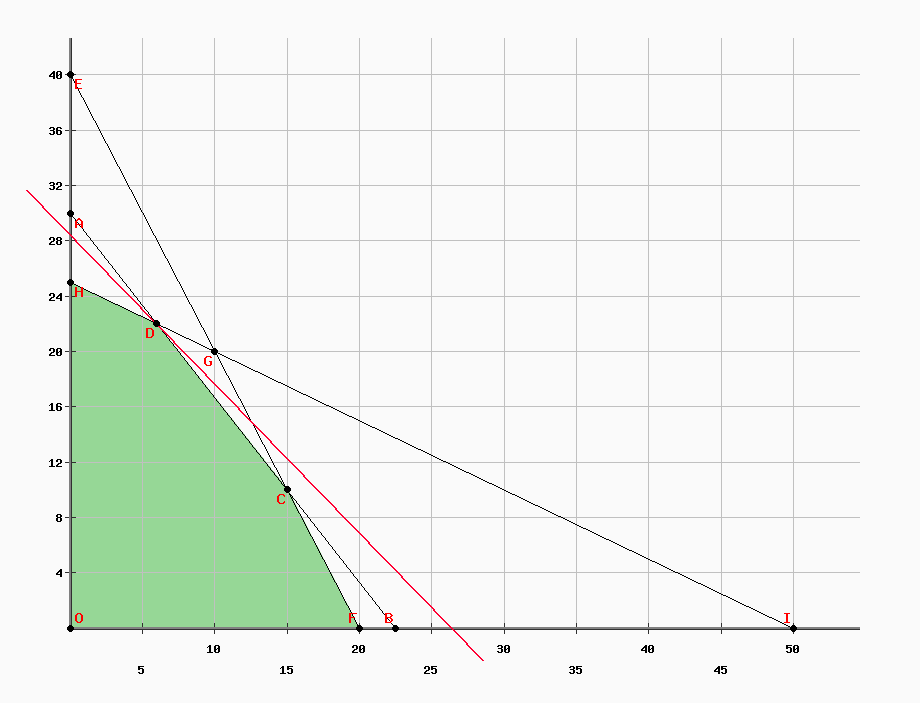
Max Z = 70x + 65y (150€ de vente vélo de montagne – 50 de matériaux – 30 de frais), (120€ de vente vélo de route – 30€ de matériaux – 25€ de frais)

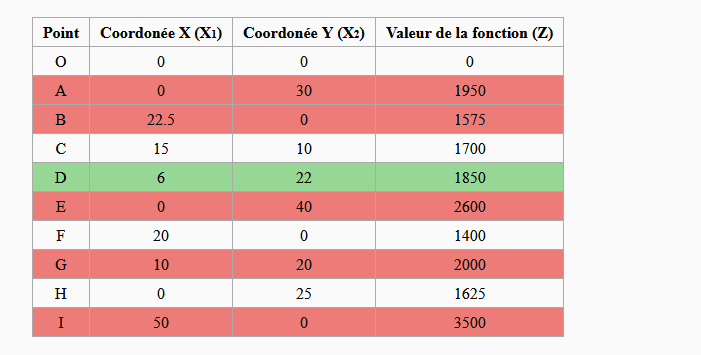
4x + 3y <= 90 (temps limité département cadre)

2x + 1y <= 40 (temps limité département montage)

x + 2y <= 50 (temps limité département finition)

x, y >= 0 (nombre de vélos de montagne et de vélos de route produits doit être positif)

****

****

d. On ajoute x <= 20 et y <= 18

e. Avec l’algorithme Branch and Bound on peut donc trouver comme valeurs optimales : x = 9.0, y = 18.0 => Z = 1800.0

**Exercice 2 : A star**

Soit le graphe G et l’heuristique h(n) suivants :



|  |  |  |
| --- | --- | --- |
| Source | Destination | coût |
| A | B | 3 |
| A | C | 1 |
| A | D | 1 |
| B | E | 1 |
| B | F | 1 |
| C | F | 3 |
| C | G | 2 |
| D | G | 1 |
| D | H | 6 |
| E | H | 1 |
| F | H | 3 |
| G | H | 5 |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| sommet | A | B | C | D | E | F | G | H |
| heuristique | 4 | 2 | 2 | 2 | 1 | 3 | 4 | 0 |

Le nœud source est A et le nœud d’arrivée est H. Le coût est indiqué sur chaque arc du graphe et h(n) représente la fonction heuristique associée à chaque nœud du graphe.

1. Simulez l’exécution de l’algorithme A\* et donnez l’ensemble d’itérations avec le contenu des listes « Open » et « Closed » à chaque fois.
2. Donnez la solution finale de l’algorithme A\*

Correction :

Solution finale de l'algorithme A\*

|  |  |
| --- | --- |
| Nœuds à explorer - étapes | Nœuds visités et fermés |
| (A, 4, vide) | - |
| (D, 1+2, A), (C, 1+2, A), (B, 3+2, A) | (A, 4, vide) |
| (G, 1+1+4, D), (H, 1+6+0, D), (C, 1+2, A), (B, 3+2, A) | (D, 1+2, A) |

Le chemin trouvé par l'algorithme A\* est : A→D→H avec un coût total de 7.

1. L’heuristique h(n) est-elle admissible ? justifiez.

oui

**Exercice 3 :**

Soit l'ensemble des 10 points suivants :

Point 1:(1,5) Point 2:(4,8) Point 2:(4,8) Point 3:(7,2) Point 3:(7,2) Point 4:(3,6) Point 4:(3,6) Point 5:(6,4) Point 5:(6,4) Point 6:(2,9) Point 6:(2,9) Point 7:(5,2) Point 8:(2,3) Point 8:(2,3) Point 9:(7,7) Point 9:(7,7) Point 10:(3,9) Point 10:(3,9)

On veut répartir ces points en trois (3) clusters, en utilisant l'algorithme K-means. La distance d entre deux points a et b est calculée ainsi :

Travail à faire :

1. Donnez le principe de fonctionnement de l’algorithme K-means et des avantages/inconvénients

2. Appliquez K-means en choisissant comme centres initiaux des 3 clusters respectivement :

1. Centre de départ 1 : (2, 5)
2. Centre de départ 2 : (6, 7)
3. Centre de départ 3 : (4, 3)

Montrez toutes les étapes de calcul et la solution finale (les centre de Cluster et leurs compositions).

Correction :

### 1. Principe de fonctionnement de l’algorithme K-means

#### Principe de fonctionnement

L’algorithme K-means partitionne un ensemble de points en \( K \) clusters en minimisant la variance intra-cluster. Il fonctionne de la manière suivante :

1. \*\*Initialisation\*\* : Choisissez \( K \) centres initiaux (généralement aléatoires ou via une méthode spécifique).

2. \*\*Affectation des points\*\* : Affectez chaque point au cluster dont le centre est le plus proche en utilisant une distance donnée (ici, la distance de Manhattan).

3. \*\*Recalculation des centres\*\* : Pour chaque cluster, calculez le nouveau centre comme la moyenne des points assignés à ce cluster.

4. \*\*Convergence\*\* : Répétez les étapes 2 et 3 jusqu'à ce que les centres ne changent plus de manière significative ou jusqu'à un nombre maximal d'itérations.

#### Avantages

- \*\*Simplicité\*\* : Facile à comprendre et à implémenter.

- \*\*Rapidité\*\* : Converge généralement rapidement.

- \*\*Scalabilité\*\* : Peut gérer un grand nombre de points et de dimensions.

#### Inconvénients

- \*\*Sensibilité aux centres initiaux\*\* : Les résultats peuvent varier en fonction des centres initiaux choisis.

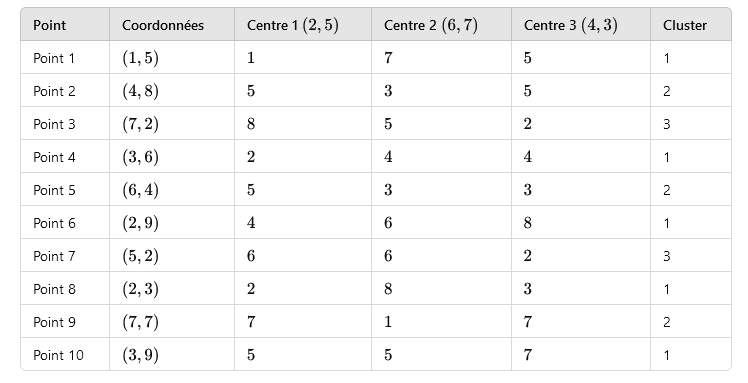
- \*\*Convergence locale\*\* : Peut converger vers des minima locaux.

- \*\*Nécessité de spécifier K\*\* : Il faut choisir le nombre de clusters à l'avance.

- \*\*Clusters sphériques\*\* : Fonctionne mieux pour des clusters de forme sphérique avec une taille similaire.

### 2. Application de K-means

Calcul des distances de chaque point aux centres initiaux :



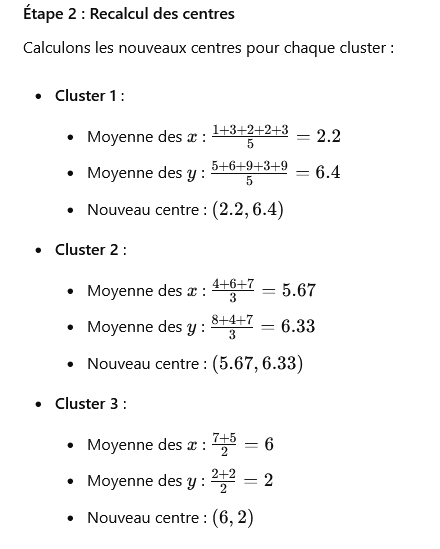
Répartition initiale :

- \*\*Cluster 1\*\* : \((1,5), (3,6), (2,9), (2,3), (3,9)\)

- \*\*Cluster 2\*\* : \((4,8), (6,4), (7,7)\)

- \*\*Cluster 3\*\* : \((7,2), (5,2)\)

#### Étape 2 : Recalcul des centres



#### Étape 3 : Réaffectation des points

Nous recalculons les distances des points aux nouveaux centres et réaffectons les points aux clusters.



Répartition après réaffectation :

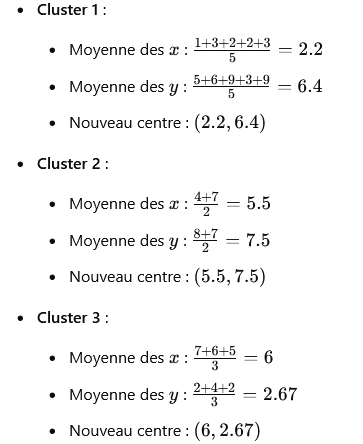
- \*\*Cluster 1\*\* : \((1,5), (3,6), (2,9), (2,3), (3,9)\)

- \*\*Cluster 2\*\* : \((4,8), (7,7)\)

- \*\*Cluster 3\*\* : \((7,2), (6,4), (5,2)\)

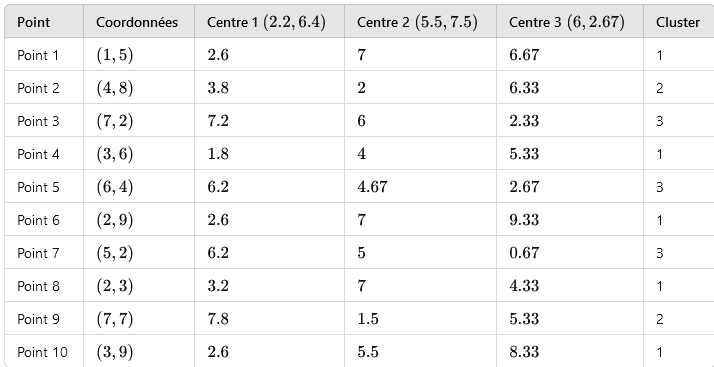
#### Étape 4 : Recalculation des centres

Calculons les nouveaux centres pour chaque cluster :



#### Réaffectation finale

Vérifions la convergence en réaffectant les points une dernière fois.



Répartition finale :

- \*\*Cluster 1\*\* : \((1,5), (3,6), (2,9), (2,3), (3,9)\)

- \*\*Cluster 2\*\* : \((4,8), (7,7)\)

- \*\*Cluster 3\*\* : \((7,2), (6,4), (5,2)\)

#### Centres finaux des clusters

- \*\*Cluster 1\*\* : Centre \((2.2, 6.4)\)

- \*\*Cluster 2\*\* : Centre \((5.5, 7.5)\)

- \*\*Cluster 3\*\* : Centre \((6, 2.67)\)

Nous avons donc les clusters finaux et leurs centres, ainsi que la répartition des points.