TD N°1 : Programmation Dynamique Stochastique (PDS)

Rappel

Les problèmes d'optimisation auxquels nous nous intéressons comportent les éléments de base suivants:

- Un système dynamique à temps discret décrivant l'évolution de la variable d'état au fil des périodes.
- Une variable de décision influençant l'évolution précédente, dont les valeurs possibles dépendent de l'état du système.
- Une fonction coût, quand elle est additive elle a la forme : $E[g_{N+1}(x_{N+1}) + \sum_{k=1}^{N} g_k(x_k, u_{k,} w_k)]$.

Tous processus de décisions séquentielles ne font pas intervenir de paramètres aléatoires et, dans des nombreuses applications, l'état x_{k+1} est parfaitement déterminé dès que l'état courant x_k et la décision u_k sont connus. Il s'agit, dans ce cas, de systèmes de décisions déterministes qui peuvent être vus comme des cas particuliers des systèmes de décisions stochastiques, caractérisés par des perturbations artificielles ne pouvant prendre qu'une seule valeur.

Exercice 1 : (*Facteur d'actualisation***)**

Une compagnie a besoin d'utiliser une machine pour les cinq années à venir. Actuellement, elle dispose d'une machine neuve. A la fin de chaque année, la compagnie doit décider si elle veut garder son installation ou la revendre et la remplacer par une neuve. La durée de vie d'une machine ne dépasse pas trois ans et son prix d'achat P est de 500 000 fr. Les revenues annuels r(i), les coûts de maintenance c(i) et le prix de revente p(i) d'une machine dépendent de l'âge i de l'installation et sont données en table ci-dessous. La machine est revendue à la fin de la cinquième année.

Question : Quelle stratégie de remplacement doit adopter la compagnie si elle veut maximiser son profit pour les cinq prochaines années ?

i	0	1	2
r(i)	400 000	300 000	150 000
c(i)	50 000	70 000	110 000
p(i)	250 000	120 000	50 000

Tableau1. Données du problème de remplacement d'une machine.

i : âge de la machine en début d'année

Indication : On considère un facteur d'actualisation des coûts et des profits $\alpha = 0.9$, tel que un montant m disponible dans k années correspond à un montant $\alpha^k m$ disponible immédiatement.

Exercice 2 : (*Fonction MINIMAX*)

Un automobiliste doit se rendre de Séville à Strasbourg. Il a décidé d'effectuer le voyage en quatre jours et d'en profiter pour rendre visite à quelques amis. Afin de limiter les risques d'accident dus à la fatigue et de disposer d'un maximum de temps auprès de ses amis, il aimerait minimiser la durée de sa plus longue étape journalière. Notre conducteur a des amis à Madrid, Valence, Barcelone, Toulouse, Bordeaux, Lyon et Paris. Sur la figure ci-dessous, sont illustrés les différentes étapes ainsi que les estimations des temps de trajets.

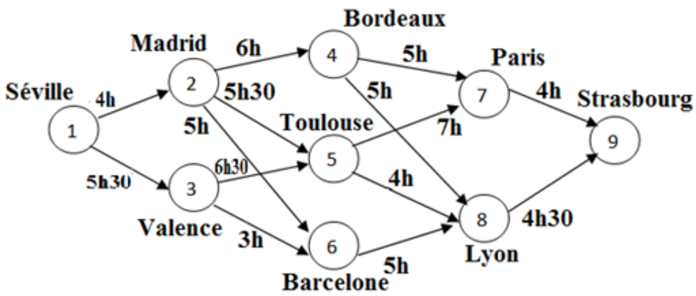

Figure 1. Description graphique du problème d'itinéraire de voyage.

Question : Trouver l'itinéraire de voyage optimal en appliquant l'algorithme de la PD.

Exercice 3 : (*Variables continues*)

Considérons un modèle de contrôle de la trajectoire d'un engin volant (avion, fusée, satellite, etc.). Supposons que le temps a été discrétisé en N intervalles numérotés de 1 à N. La vitesse de l'engin peut être modifiée au début de chaque intervalle et le but du contrôle est d'atteindre une vitesse cible V à la fin de l'intervalle N. Afin d'éviter des changements trop brusques, toute variation Δ_v de la vitesse est pénalisée par un coût égale à Δ_v^2 . Si, à la fin du processus, la vitesse v atteinte ne correspond pas à la vitesse V souhaitée, une pénalité égale à $4(V\text{-}v)^2$ est également encourue.

Question : Déterminer pour chaque intervalle k la vitesse de l'engin de manière à minimiser la somme des pénalités.

Indications : Considérez : N=2, la vitesse initiale=0 et la vitesse cible V=900.

Exercice 4 : (*Programmation dynamique stochastique***)**

Un joueur d'échec doit disputer une rencontre en deux parties et désire maximiser ses chances de gain. Chaque partie rapporte 1 point au vainqueur et 0 au perdant, à moins d'un match nul auquel cas les deux joueurs marquent un demi point. Si, à l'issue, des deux matchs, le score est égale à 1-1 les deux joueurs continuent à s'affronter jusqu'à ce que l'un d'eux remporte une partie et la rencontre par la même occasion. Connaissant bien les ouvertures pratiquées par son adversaire, le joueur a sélectionné les siennes afin de pouvoir choisir, pour chaque affrontement, entre deux styles de jeu :

1- un ***style agressif*** lui donnant une probabilité de gain $p_a > 0$ et une probabilité de perte $1 - p_a$.

2- un ***style passif*** lui donnant une probabilité de nul $p_p > 0$ et une probabilité de perte $1 - p_p$.

Question : déterminer une stratégie de sélection de style de jeu maximisant les chances du joueur à gagner la rencontre pour les probabilités : $p_a = 0.45 \; et \; p_p = 0.9$.

Indications : En cas d'égalité après les deux matchs, le joueur adopte un style de jeu agressif pour tous les matchs qui en suivent.

Le coût terminal exprime la probabilité de gain : $g_{N+1} = \begin{cases} 1 & \text{en cas de gain de la rencontre} \\ p_a & \text{au cas d'égalité après les deux parties} \\ 0 & \text{en cas de perte} \end{cases}$

De plus, nous avons que : $g_k(x_k, u_k, \omega_k) = 0, \forall \; k$