Examen d'optimisation stochastique

Vincent Leclère

10/11/2016

Name: ____________________

1 Partie Théorique

1. (6 points) On considère le système dynamique à valeur dans $\{0, 1, 2\}$ donné par

$$\boldsymbol{x}_{t+1} \equiv \boldsymbol{x}_t + \boldsymbol{u}_t + \boldsymbol{w}_{t+1}[3].$$

On suppose que $\boldsymbol{u}_t \in \mathbb{Z}$, $\boldsymbol{x}_0 = 0$, et que $(\boldsymbol{w}_t)_{t\in\mathbb{Z}}$ est une suite de variable aléatoire i.i.d de Bernoulli de paramètre 0.5 (i.e. $\mathbb{P}(w_t = 1) = 0.5$ et $\mathbb{P}(w_t = 0) = 0.5$). Par exemple si $u_0 = 2$, $\boldsymbol{x}_1$ vaut 2 avec probabilité 0.5 ($\boldsymbol{w}_1 = 0$) et 0 sinon ($\boldsymbol{w}_1 = 1$).

On souhaite résoudre le problème suivant

$$\begin{aligned} \min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\big[|u_0| + |\boldsymbol{u}_1| + \boldsymbol{x}_2^2\big] && (1) \\ s.t. \quad & \boldsymbol{x}_{t+1} \equiv \boldsymbol{x}_t + \boldsymbol{u}_t + \boldsymbol{w}_{t+1}[3] && (2) \\ & \boldsymbol{u}_t \in \mathbb{Z} && (3) \\ & u_0 \text{ deterministe} \qquad \sigma(\boldsymbol{u}_1) \subset \sigma(\boldsymbol{w}_1) && (4) \end{aligned}$$

(a) (1 point) Expliquer pourquoi le problème peut-être résolu par programmation dynamique ? Donner la structure d'information du problème.

> **Solution:** Les bruits sont indépendants (0.25) et supposé exogènes (0.25). Il s'agit d'un problème en décision-hasard (0.5).

(b) (1 point) Justifier que l'on peut se contenter d'étudier un nombre fini de contrôle u que l'on explicitera.

> **Solution:** Le système dynamique est à valeur dans $\mathbb{Z}/3\mathbb{Z}$ donc les contrôles u et $u+3$ produisent le même résultat (0.5). Comme l'on veut minimiser la valeur absolue des contrôles il suffit de considérer $\{-1, 0, 1\}$ (0.5).

(c) (2 points) Trouver la valeur du problème par programmation dynamique. Donner la stratégie optimale sous forme de "look-up" table

Solution: (1)

x	V_0	V_1	V_2
2	1	1.5	4
1	1.5	1.5	1
0	1	0.5	0

Donc la valeur du problème est $V(0) = 1$ (0.5).

La stratégie optimale est donnée par (1)

x	π_0	π_1
2	0	1
1	0	-1
0	0	0

(d) (2 points) Trouver la borne inférieure obtenue en considérant une solution anticipative.

Solution: Donnons une solution optimale (elle n'est pas toujours unique)

$\boldsymbol{w}_1$	$\boldsymbol{w}_2$	$\boldsymbol{u}_1$	$\boldsymbol{u}_2$	$\boldsymbol{x}_2$	coût total
0	0	0	0	0	0
0	1	0	-1	0	1
1	0	-1	0	0	1
1	1	0	1	0	1

La valeur de la solution anticipative est donc $0.75 < 1$.

2. (6 points) Considérons le problème suivant

$$\begin{aligned} \min_{x,\boldsymbol{y}} \quad & \mathbb{E}\Big[x^2 + \boldsymbol{y}^2\Big] \\ s.t. \quad & x + \boldsymbol{y} = \boldsymbol{d} \\ & \sigma(\boldsymbol{y}) \subset \sigma(\boldsymbol{d}) \end{aligned}$$

où $\boldsymbol{d}$ est une variable de Bernoulli de paramètre p (i.e. $\mathbb{P}(\boldsymbol{d} = 1) = p$ et $\mathbb{P}(\boldsymbol{d} = 0) = 1 - p$).

(a) (1 point) Justifier que ce problème est un "two-stage program", et l'écrire sous forme extensive.

Solution: x est une variable déterministe de première étape, $\boldsymbol{y}$ est choisie en connaissant l'ensemble des aléas.

$$\begin{aligned} \min_{x,y_1,y_2} \quad & x^2 + py_1^2 + (1-p)y_2^2 \\ s.t. \quad & x + y_1 = 1 \\ & x + y_2 = 0 \end{aligned}$$

(b) (2.5 points) Introduire une variable supplémentaire de première étape pour représenter la contrainte de non-anticipativité et dualiser cette contrainte. Justifier qu'il n'y a pas de saut de dualité et expliciter le problème dual sous forme décomposée (on doit obtenir une maximization de somme de problèmes de minimisation indépendants).

Solution: Le problème peut s'écrire (1)

$$\begin{aligned} \min_{x_1,x_2,y_1,y_2} \max_{\lambda \in \mathbb{R}} \quad & x^2 + py_1^2 + (1-p)y_2^2 + \lambda(x_1 - x_2) \\ s.t. \quad & x_1 + y_1 = 1 \\ & x_2 + y_2 = 0 \end{aligned}$$

Les contraintes étant linéaires elles sont qualifiées (0.25). La fonction objectif est convexe, il y a donc dualité forte (0.25). On peut réécrire le problème dual sous la forme. (1)

$$\max_{\lambda \in \mathbb{R}} \quad \begin{aligned}[t] \min_{x_1,y_1} & \Big\{p(x_1^2 + y_1^2) + \lambda x_1\Big\} \\ s.t. \quad & x_1 + y_1 = 1 \end{aligned} \; + \; \begin{aligned}[t] \min_{x_2,y_2} & \Big\{(1-p)(x_2^2 + y_2^2) - \lambda x_2\Big\} \\ s.t \quad & x_2 + y_2 = 0 \end{aligned}$$

(c) (2.5 points) Expliciter les deux variables de premières étapes en fonction du multiplicateur de la contrainte de non-anticipativité. En déduire le multiplicateur optimal, les variables primales optimales et la valeur du problème.

Solution: Le minimum de $p(x_1^2 + (1-x_1)^2) + \lambda x_1$ est atteint en $x_1 = \frac{2p-\lambda}{4p}$. Le minimum de $(1-p)2x_2^2 - \lambda x_2$ est atteint en $x_2 = \frac{\lambda}{4(1-p)}$. Pour le multiplicateur optimal $x_1 = x_2$, donc $\lambda^* = 2p(1-p)$. En conclusion on a

$$x^* = \frac{p}{2} \qquad y_1^* = 1 - \frac{p}{2} \qquad y_2^* = -\frac{p}{2}.$$

Le problème a pour valeur

$$\frac{p^2}{4} + p(1-\frac{p}{2})^2 + (1-p)\frac{p^2}{4} = \frac{p(p^2-3p+1)}{4}$$

2 Partie Modélisation

Rappel de probabilité: une variable aléatoire $\boldsymbol{X}$ suis une loi Binomiale $\mathcal{B}(n,p)$ si elle est à support dans $[\![0,n]\!]$, avec $\mathbb{P}(\boldsymbol{X}=k) = \binom{n}{k}p^k(1-p)^k$.

Dans les exercices suivant il sera parfois demandé de quelles bibliothèques vous avez besoin pour résoudre le problème. Par bibliothèque nous entendons des éléments de {algère linéaire, solveur linéaire, solveur PLNE, solveur convexe, générateur de nombres aléatoires }

L'objectif de l'examen n'étant pas d'évaluer vos connaissances en combinatoire vous pourrez définir des nombres avec des règles combinatoires du type : "soit A le nombre de vecteurs de tel espace à coordonnées entières vérifiant telle condition" quand la formule explicite est complexe.

3. (11 points) Un directeur d'usine dispose d'une chaîne de montage qui peut produire 10 types possibles de voitures (indicées par i). Il dispose d'un garage pouvant acceuillir (gratuitement) jusqu'à 500 voitures. Au début de chaque mois il reçoit les commandes (nombre de voitures de chaque type) à livrer à la fin du mois, avec un maximum de 100 de chaque type. Le coût de production est donné par une fonction linéaire par morceau $c_t(u)$ où $u \in \mathbb{R}^{10}$ représente le vecteurs des quantités à produire durant le mois. A là fin de l'année les voitures en stock sont valorisées au prix unitaire π_i, $i \in [\![1,10]\!]$. La demande en chaque type de voiture est modélisée par une loi binomiale $\mathcal{B}(100, p_{i,t})$ toutes étant indépendantes (en temps et en types de voitures). L'objectif est de minimiser l'espérance du coût annuel de production. Au début de l'année on suppose qu'il dispose de 50 voiture de chaque type dans son garage.

(a) (2 points) Préciser le système dynamique contrôlé correspondant au problème. Spécifier son espace d'état, son espace de contrôle, sa dynamique et les bruits l'affectant. Quelle est la structure de décision : boucle ouverte, hasard-décision ou décision-hasard ? Quel est le coût instantané ? Quel est le coût final ?

Solution: Le système dynamique est $\boldsymbol{x}_t \in [\![0,100]\!]^{10}$, où chaque coordonnée $\boldsymbol{x}_t^i$ représente le nombre de voiture de type i en stock au début du mois t. Il suit la dynamique $\boldsymbol{x}_{t+1}^i = \boldsymbol{x}_t^i + \boldsymbol{u}_t^i - \boldsymbol{d}_t^i$ où $\boldsymbol{u}_t$ représente la prodution (contrôle) du mois t et $\boldsymbol{d}_t^i$ la demande (bruit) du mois t. (1)
Le problème est en hasard-décision (0.5). Le coût instantané est $c_t(\boldsymbol{u})$, le coût final est $K(x_{12}) = \sum_{i=1}^{10} \pi_i x_{12}^i = \pi^T x_{12}$. (0.5)

(b) (2 points) Formuler le problème d'optimisation sous forme mathématique.

Solution:

$$\begin{aligned}
\min_{\boldsymbol{x}\geq 0,\boldsymbol{u}\geq 0} \quad & \mathbb{E}\Big[\sum_{t=1}^{12} c_t(\boldsymbol{u}_t) + \pi^T \boldsymbol{x}_{12}\Big] \\
s.t. \quad & \boldsymbol{x}_{t+1}^i = \boldsymbol{x}_t^i + \boldsymbol{u}_t^i - \boldsymbol{d}_t^i && \forall i \quad \forall t \\
& \boldsymbol{x}_0^i = 50 && \forall i \\
& \sum_{i=1}^{10} x_t^i \leq 500 && \forall i \\
& \sigma(\boldsymbol{p}_t) \subset \sigma(\boldsymbol{d}_0, \dots, \boldsymbol{d}_t) && \forall t
\end{aligned}$$

(c) (1 point) De quelle(s) biliothèque(s) a-t-on besoin pour résoudre le problème par programmation stochastique ? Combien de feuilles dans l'arbre seront nécessaires ? Est-ce raisonnable ?

Solution: Il y a 11 vecteurs aléatoires de demande $\boldsymbol{d}_1$, chacun avec 10 coordonnées et 100 valeurs possibles, tous étant indépendants. Le problème peut donc être résolu par programmation stochastique à l'aide d'un solveur PLNE (0.5) et de $1000^{11} = 10^{33}$ scénarios (0.5). Ce nombre est bien trop important pour imaginer résoudre le problème en pratique.

(d) (1 point) Aujourd'hui le directeur de l'usine dispose d'un outil résolvant efficacement la version déterministe du problème (où l'on suppose connaître la demande pour toute l'année). Comment estimer une borne inférieure du coût de production annuel espéré ? A l'aide de quelle(s) bibliothèque(s) ?

Solution: On peut estimer la valeur de la solution anticipative. Pour cela on tire un certain nombre N de scénarios (générateur de nombre aléatoire) (0.5), et on calcule la valeur de la version déterministe sur chaque scénario (outil déterministe, e.g. PLNE solveur). La moyenne de ces valeurs est une estimation (de Monte-Carlo) d'une bonne inférieure du coût de production annuel espéré. (0.5)

(e) (2 points) Écrire et justifier l'équation de programmation dynamique associée au problème.

Solution: Puisque les bruits $\boldsymbol{d}_t$ sont exogènes et indépendants (0.5) en temps on peut écrire une équation de programmation dynamique. (1.5)

$$\begin{aligned}
V_T(x) &= \pi^T x \\
\hat{V}_t(x,d) &= \min_{u\in [\![0,100]\!]^{10}} \Big\{ c_t(u) + V_{t+1}\big(x+u-d\big) + \mathbb{I}_{\|x+u-d\|_1 \leq 500} \Big\} \\
V_t(x) &= \mathbb{E}\Big[\hat{V}_t(x,\boldsymbol{d})\Big]
\end{aligned}$$

(f) (1 point) Si l'on dispose de la fonction de Bellman associée au problème, comment trouver la stratégie optimale ?

Solution: Si V_{t+1} est la fonction de Bellman associée au problème, alors la stratégie optimale à l'instant t en l'état x, connaissant la demande d est donnée par

$$\pi_t^*(x,d) \in \operatorname*{arg\,min}_{u\in [\![0,100]\!]^{10}} \Big\{ c_t(u) + V_{t+1}\big(x+u-d\big) + \mathbb{I}_{\|x+u-d\|_1 \leq 500} \Big\}$$

(g) (1 point) Peut-on résoudre le problème par programmation dynamique ? Si oui combien d'opérations élémentaires (évaluation de fonction, somme, produit, test d'appartenance à un intervalle) sont nécessaires ? Si non quelle approximation est nécessaire ?

> **Solution:** Il existe une équation de programmation dynamique. De plus états, controles et support des bruits sont de cardinal finis donc on peut utiliser la programmation dynamique (0.5). Cela nécessite (0.5) $\underbrace{11}_{T} \times \underbrace{A}_{|\mathbb{X}|} \times \underbrace{100^{10}}_{|\mathbb{U}|} \times \underbrace{100^{10}}_{\mathbb{W}} \times 4$ opérations, où $A := \big\{x \in \mathbb{N}^{10} \mid \|x\|_1 \leq 500\big\}$.

(h) (1 point) Le directeur de l'usine vous demande de trouver une bonne stratégie de gestion de son stock de voiture, quelle type d'approche suggérez vous ?

> **Solution:** La programmation dynamique discrète peut être utilisée en théorie mais pas en pratique (dimension trop importante) (0.5). Il faut donc se tourner vers des méthodes de type programmation dynamique approximée (0.5).

4. (7 points) Considérons un problème de contrôle de mobile en dimension 3. Soit $\boldsymbol{x}_t \in \mathbb{R}^3$ la position du mobile à l'instant t et $\boldsymbol{u}_t \in \mathbb{R}^4$ le vecteur de contrôles que vous pouvez appliquer. On suppose qu'entre le contrôle que vous voulez appliqué et le contrôle réellement appliqué au mobile il y a une erreur $\boldsymbol{w}_t \in \mathbb{R}^4$. On suppose que l'erreur est i.i.d et telle que $\mathbb{E}\big[\boldsymbol{w}_t\big] = 0$. Le mobile évolue suivant une dynamique linéaire: $\boldsymbol{x}_{t+1} = A\boldsymbol{x}_t + B(\boldsymbol{u}_t + \boldsymbol{w}_t)$. De plus on suppose que le mobile est 0 à $t = 0$. Nous cherchons à minimiser sur 10 pas de temps l'espérance de la somme des carrées des écarts à une trajectoire $(x_t^{ref})_{t\in[\![0,10]\!]}$ donnée: $\|\boldsymbol{x}_t - x_t^{ref}\|^2$ et des carrés des contrôles appliqués $\|\boldsymbol{u}_t\|^2$.

Aujourd'hui le problème est résolu par programmation dynamique en $1'$ en se fixant des bornes raisonnables sur $\boldsymbol{x}_t$ et $\boldsymbol{u}_t$ et en choisissant un pas de discrétisation pour l'état et le contrôle de 0.1.

(a) (2 points) Écrire le problème sous forme d'un problème d'optimisation stochastique. Quelle est la structure d'information ?

> **Solution:**
>
> $$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{u}} \quad & \mathbb{E}\Big[\sum_{t=0}^{9} \|\boldsymbol{x}_t - x_t^{ref}\|_2^2 + \|u_t\|_2^2 + \|\boldsymbol{x}_{10} - x_{10}^{ref}\|_2^2\Big] \\ s.t. \quad & \boldsymbol{x}_{t+1} = A\boldsymbol{x}_t + B(\boldsymbol{u}_t + \boldsymbol{w}_t) \qquad \forall t \\ & \boldsymbol{x}_0 = 0 \\ & \sigma(\boldsymbol{u}_t) \subset \sigma(\boldsymbol{w}_0, \dots, \boldsymbol{w}_{t-1}) \end{aligned}$$
>
> Nous sommes en décision-hasard.

(b) (1 point) Écrire l'équation de programmation dynamique associée au problème.

> **Solution:**
>
> $$\begin{aligned} V_{10}(x) &= \|\boldsymbol{x}_{10} - x_{10}^{ref}\|_2^2 \\ V_t(x) &= \min_{u\in\mathbb{R}^4} \mathbb{E}\Big[\|\boldsymbol{x}_t - x_t^{ref}\|_2^2 + \|u_t\|_2^2 + V_{t+1}\big(A\boldsymbol{x}_t + B(\boldsymbol{u}_t + \boldsymbol{w}_t)\big)\Big] \end{aligned}$$

(c) (1 point) La qualité de la solution n'étant pas satisfaisante, on vous demande de diviser par 2 le pas de discrétisation (état et contrôle) pour avoir un contrôle plus fin. Combien de temps prendra la résolution du problème ?

Solution: Il y a 3 dimension d'état et 4 dimension de contrôle. Diviser par 2 le pas consiste à multiplier par 2^7 (en réalité un peu moins) le temps de calcul soit plus de $2h$.

(d) (3 points) On vous demande de trouver un moyen plus rapide et au moins aussi précis de résoudre le problème, que proposez vous ? De quelles bibliothèque aurez-vous besoin ? Quels sont les avantages / inconvénients de votre suggestion ?

Solution: Le système suis une dynamique linéaire pour un coût quadratique avec des matrices symmétriques définie positive. Le bruit qui affecte la dynamique est $B\boldsymbol{w}_t$ qui est centré et indépendant en temps. Nous sommes donc dans un cas de problème linéaire quadratique pour lequel on peut résoudre exactement l'équation de programmation dynamique. Pour cela nous avons seulement besoin d'une librairie d'algèbre linéaire. La méthode sera plus rapide, plus précise (solution exacte) et ne nécessiteras pas de trouver de bornes a priori. En revanche si l'on souhaite ajouter des contraintes (e.g. des bornes sur les controles) ou changer la fonction coût la méthode ne pourra plus être adaptée.