

Prof. Pierre L'Ecuyer

DEVOIR 3

Devoir à remettre le *lundi 9 mars 2015, au début du cours, 9h30*. DPOC désigne le livre de Bertsekas, “Dynamic Programming and Optimal Control”. Dans chaque cas, il est important de bien définir (mathématiquement et clairement) tous les symboles utilisés. Dans la correction, les explications claires et détaillées comptent davantage que les résultats.

1. Ce problème généralise un exemple vu en classe, au Chapitre 4. On a le système linéaire

$$x_{k+1} = A_k x_k + B_k u_k, \quad k = 0, 1, \dots, N-1,$$

où x_k et u_k sont des vecteurs de dimensions n et m , les A_k sont des matrices $n \times n$ connues, et les B_k sont des matrices aléatoires indépendantes $n \times m$ de lois de probabilité connues qui ne dépendent pas de (x_k, u_k) . Chaque u_k peut être choisi dans \mathbb{R} (aucune contrainte) après avoir observé le vecteur x_k .

On veut maximiser $\mathbb{E}[U(c^\top x_N)]$ où c est un vecteur connu et U est une fonction d'utilité concave et deux fois continument différentiable, telle que

$$-\frac{U'(y)}{U''(y)} = a + by.$$

Montrez que la décision u_k optimale est une fonction linéaire de x_k pour chaque k (i.e., dans la politique optimale, chaque μ_k est une fonction linéaire).

Suggestion: Réduire le problème à un problème en une dimension via un changement de variable de la forme $y_k = c^\top A_{N-1} \cdots A_k x_k$ pour l'état, et quelque chose de semblable pour la décision, puis utiliser les résultats de la section 4.3 de DPOC.

2. DPOC (2005), Exercice 4.29. Remplacez 5.1 par 4.5.1 dans la question.

3. DPOC (2005), Exercice 5.14.

4. Vous disposez d'un montant initial $x_0 > 0$ à investir, sur une période de N mois. Après k mois, pour $k = 0, \dots, N-1$, votre capital est x_k et vous décidez d'un montant u_k à investir dans des actifs risqués. Ce montant doit satisfaire $0 \leq u_k \leq x_k$. Avec probabilité $p > 1/2$, le montant investi devient $(1+c)u_k$ au début du mois suivant, et avec probabilité $1-p$ il devient $(1-d)u_k$, où c et d sont des constantes positives telles que $pc > (1-p)d$. Votre objectif est de maximiser $\mathbb{E}[\ln x_N]$, votre utilité espérée après N étapes, en supposant que votre fonction d'utilité est logarithmique.

Prouvez (pensez à l'induction) que la politique optimale est définie par $u_k = bx_k$ et que la fonction d'utilité espérée optimale à partir de l'étape k est

$$J_k(x_k) = C_k + \ln x_k$$

où

$$b = \frac{(pc - (1-p)d)}{cd} \quad \text{et} \quad C_k = (N - k)[p \ln(1 + cb) + (1 - p) \ln(1 - db)]$$

pour $k = 0, \dots, N - 1$. Selon cette politique, on investit toujours une fraction constante de notre capital, et cette fraction b ne dépend pas du numéro d'étape.