

Modèle stochastique sur horizon fini

Fabian Bastin

DIRO, Université de Montréal

IFT-6521 – Hiver 2013

Modèle stochastique sur horizon fini

A l'étape k , on observe l'état x_k et prend une décision $u_k \in U_k(x_k)$. Puis une variable aléatoire ω_k est générée selon une loi de probabilité $P_k(\cdot | x_k, u_k)$. On observe ω_k , on paye un coût $g_k(x_k, u_k, \omega_k)$, et l'état à la prochaine étape est $x_{k+1} = f_k(x_k, u_k, \omega_k)$.

à l'étape k , on a:

x_k = état du système à l'étape k ;

u_k = décision prise à l'étape k .

ω_k = perturbation (var. aléatoire) produite à l'étape k .

g_k = fonction de coût;

f_k = fonction de transition;

On cherche une politique admissible $\pi = (\mu_0, \dots, \mu_{N-1})$ qui minimise l'espérance mathématique du coût total,

$$\mathbb{E} \left[g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, \omega_k) \right].$$

On a

$$\begin{aligned} J_N(x) &= g_N(x) \quad \text{pour } x \in X_N, \\ J_k(x) &= \text{coût espéré total optimal de l'étape } k \\ &\quad \text{à la fin, si on est dans l'état } x \text{ à l'étape } k \\ &= \inf_{u \in U_k(x)} \mathbb{E}_{\omega_k} [g_k(x, u, \omega_k) + J_{k+1}(f_k(x, u, \omega_k))] \\ &\quad \text{pour } 0 \leq k \leq N-1, \quad x \in X_k, \end{aligned}$$

où ω_k suit la loi $P_k(\cdot | x, u)$, et

$$\mu_k^*(x) = \arg \min_{u \in U_k(x)} \mathbb{E}_{\omega_k} [g_k(x, u, \omega_k) + J_{k+1}(f_k(x, u, \omega_k))].$$

Exemple de gestion d'un inventaire

Nous pouvons appliquer l'algorithme de programmation dynamique si tout est discret et fini.

Reprenons un exemple de gestion d'inventaire (Exemple 1.3.2 de Bertsekas), en notant

x_k = niveau des stocks au début du mois k ,
avant de commander;

u_k = nombre de biens commandés (et reçus) au début du mois k ;

ω_k = nombre de bien demandés par les clients durant le
mois k . On suppose que les ω_k sont des variables
aléatoires discrètes indépendantes;

Nous supposons de plus que pas plus de deux unités ne peuvent être stockées:

$$x_k + u_k \leq 2.$$

La demande excédentaire ($\omega_k - x_k - u_k$) est perdue.

Exemple de gestion d'un inventaire (suite)

Le coût de stockage pour la période k est

$$(x_k + u_k - \omega_k)^2,$$

impliquant une pénalité à la fois pour l'excès d'inventaire et pour la demande non satisfaite à la fin de la période k .

Le coût de commande est de 1 par unité commandé.

Par conséquent,

$$g_k(x_k, u_k, \omega_k) = u_k + (x_k + u_k - \omega_k)^2.$$

Le coût terminal est supposé nul:

$$g(x_N) = 0.$$

Exemple de gestion d'un inventaire (suite)

On prendra $N = 3$, $x_0 = 0$, et, pour tout k ,

$$p(\omega_k) = \begin{cases} 0.1 & \text{si } \omega_k = 0, \\ 0.7 & \text{si } \omega_k = 1, \\ 0.2 & \text{si } \omega_k = 2. \end{cases}$$

Nous pouvons démarrer l'algorithme de programmation dynamique avec

$$J_3(x_3) = 0.$$

Pour $k = 0, 1$ ou 2 , nous avons la récurrence

$$J_k(x_k) = \min_{u_k} E_{\omega_k}[u_k + (x_k + u_k - \omega_k)^2 + J_{k+1}(\max\{0, x_k + u_k - \omega_k\})],$$

sous les contraintes

$$u_k \in \{0, 1, 2\},$$

$$u_k \in [0, 2 - x_k].$$

Exemple de gestion d'un inventaire: périodes 2

Considérons les trois états possibles.

Tout d'abord,

$$\begin{aligned} J_2(0) &= \min_{u_2} E_{\omega_2}[u_2 + (u_2 - \omega_2)^2] \\ &= \min_{u_2} u_2 + 0.1(u_2)^2 + 0.7(u_2 - 1)^2 + 0.2(u_2 - 2)^2. \end{aligned}$$

En comparant explicitement les valeurs pour $u_2 = 0, 1, 2$, nous obtenons

$$J_2(0) = 1.3, \quad \mu_2^*(0) = 1.$$

De même,

$$\begin{aligned} J_2(1) &= 0.3, \quad \mu_2^*(1) = 0, \\ J_2(2) &= 1.1, \quad \mu_2^*(2) = 0. \end{aligned}$$

Exemple de gestion d'un inventaire: période 1

De la même manière,

$$J_1(0) = \min_{u_1} E_{\omega_1}[u_1 + (u_1 - \omega_1)^2 + J_2(\max\{0, u_1 - \omega_1\})].$$

Il est facile de calculer

$$J_1(0) = \min\{2.8, 2.5, 3.68\} = 2.5, \mu_1^*(0) = 1.$$

De même,

$$J_1(1) = 1.5, \mu_2^*(1) = 0,$$

$$J_1(2) = 1.68, \mu_2^*(2) = 0.$$

Exemple de gestion d'un inventaire: période 0

De la même manière, on peut montrer

$$J_0(0) = 3.7, \mu_0^*(0) = 1.$$

Par conséquent, la politique optimale est de commander une unité si le stock est vide, et de ne rien commander sinon. Le coût espéré à l'étape initiale est 3.7.

Objectifs du chapitre

L'énumération devient vite lourde.

Dans ce chapitre, on va examiner des modèles (ou exemples) plus particuliers et caractériser la forme de la politique optimale μ_k^* et de la fonction de coût anticipé J_k .

Système linéaire à coût quadratique

x_k = vecteur de dimension n ,

u_k = vecteur de dimension m ,

w_k = vecteur de dimension n , $\mathbb{E}[w_k] = 0$, $\text{Var}[w_k] < \infty$,
indép. de (x_k, u_k) , et les w_k sont indépendants.

$x_{k+1} = f_k(x_k, u_k, w_k) = A_k x_k + B_k u_k + w_k,$

$g_N(x_N) = x'_N Q_N x_N,$

$g_k(x_k, u_k) = x'_k Q_k x_k + u'_k R_k u_k,$

où A_k , B_k , Q_k , et R_k sont des matrices, les Q_k sont symétriques et
définies semi-positives, les R_k sont symétriques et définies positives.

Le modèle se généralise facilement au cas où $\mathbb{E}[w_k] \neq 0$ et où les formes quadratiques sont décentrées. Le coût total pourrait ainsi avoir la forme

$$E \left[(x_N - \bar{x}_N)' Q_N (x_N - \bar{x}_N) + \sum_{k=0}^{N-1} ((x_k - \bar{x}_k)' Q_k (x_k - \bar{x}_k) + u_k' R_k u_k) \right].$$

On ne le fera pas pour éviter d'alourdir la notation. On se ramène au cas centré par simple changement de variable.

Les **équations de la PD** deviennent:

$$\begin{aligned} J_N(x_N) &= g_N(x_N) = x_N' Q_N x_N, \\ J_k(x_k) &= \min_{u_k} \mathbb{E} [x_k' Q_k x_k + u_k' R_k u_k \\ &\quad + J_{k+1}(A_k x_k + B_k u_k + w_k)], \quad k < N. \end{aligned}$$

On va montrer que les fonctions J_k sont quadratiques en x_k et u_k , et que $\mu_k^*(x_k)$ est linéaire en x_k , et ne dépend pas des lois de probabilité des w_k . Ainsi, pour ce modèle, on peut remplacer les w_k par leurs espérances sans changer la politique optimale! C'est le principe du modèle déterministe équivalent.

Posons $K_N = Q_N$, et pour $k < N$,

$$\begin{aligned} L_k &= -(B'_k K_{k+1} B_k + R_k)^{-1} B'_k K_{k+1} A_k, \\ K_k &= A'_k (K_{k+1} - K_{k+1} B_k (B'_k K_{k+1} B_k + R_k)^{-1} B'_k K_{k+1}) A_k + Q_k \\ &= A'_k K_{k+1} (A_k + B_k L_k) + Q_k. \end{aligned} \quad (\text{Eq. de Riccati})$$

La proposition qui suit donne la forme de la politique optimale, qui est très facile à calculer et à implanter via une rétroaction ("feedback") linéaire, avec matrices de gain L_k .

Proposition. Les matrices K_k sont symétriques et définies semi-positives et les L_k sont donc bien définies.

La politique optimale et la fonction de coût optimal J_k , pour $0 \leq k < N$, sont donnés par

$$\mu_k^*(x_k) = L_k x_k,$$

$$J_k(x_k) = x_k' K_k x_k + \sum_{j=k}^{N-1} \mathbb{E}[w_j' K_{j+1} w_j].$$

Preuve. Par induction arrière sur k .

Pour $k = N - 1$, en utilisant le fait que $\mathbb{E}[w_{N-1}] = 0$, et $K_N = Q_N$,

$$\begin{aligned} J_{N-1}(x_{N-1}) &= \min_{u_{N-1}} \mathbb{E} [x'_{N-1} Q_{N-1} x_{N-1} + u'_{N-1} R_{N-1} u_{N-1} \\ &\quad + (A_{N-1} x_{N-1} + B_{N-1} u_{N-1} + w_{N-1})' \\ &\quad K_N (A_{N-1} x_{N-1} + B_{N-1} u_{N-1} + w_{N-1})] \\ &= x'_{N-1} Q_{N-1} x_{N-1} \\ &\quad + x'_{N-1} A'_{N-1} K_N A_{N-1} x_{N-1} + \mathbb{E}[w'_{N-1} K_N w_{N-1}] \\ &\quad + \min_{u_{N-1}} [u'_{N-1} R_{N-1} u_{N-1} + u'_{N-1} B'_{N-1} K_N B_{N-1} u_{N-1} \\ &\quad + 2x'_{N-1} A'_{N-1} K_N B_{N-1} u_{N-1}]. \end{aligned}$$

En mettant à 0 la dérivée p.r. à u_{N-1} , on obtient

$$2R_{N-1}u_{N-1} + 2B'_{N-1}K_N B_{N-1} u_{N-1} + 2(x'_{N-1} A'_{N-1} K_N B_{N-1})' = 0$$

Puisque R_{N-1} est définie positive et $B'_{N-1}K_NB_{N-1}$ est définie semi-positive, cela implique que la commande optimale est la valeur de u_{N-1} qui satisfait cette équation, i.e.,

$$\begin{aligned}\mu_{N-1}^*(x_k) &= -(R_{N-1} + B'_{N-1}K_NB_{N-1})^{-1}B'_{N-1}K_NA_{N-1}x_{N-1} \\ &= L_{N-1}x_{N-1}.\end{aligned}$$

En remplaçant u_{N-1} par $\mu_{N-1}^*(x_{N-1})$ dans l'expression pour J_{N-1} , on obtient.

$$J_{N-1}(x_{N-1}) = x'_{N-1}K_{N-1}x_{N-1} + \mathbb{E}[w'_{N-1}K_Nw_{N-1}].$$

Pour le voir, développons l'expression de $J_{N-1}(x_{N-1})$.

En remplaçant u_{N-1} par la décision optimale, nous obtenons

$$\begin{aligned} J_{N-1}(x_{N-1}) &= x'_{N-1} Q_{N-1} x_{N-1} \\ &+ x'_{N-1} A'_{N-1} K_N A_{N-1} x_{N-1} + \mathbb{E}[w'_{N-1} K_N w_{N-1}] \\ &+ [x'_{N-1} L'_{N-1} R_{N-1} L_{N-1} x_{N-1} + x'_{N-1} L'_{N-1} B'_{N-1} K_N B_{N-1} L_{N-1} x_{N-1} \\ &+ 2x'_{N-1} A'_{N-1} K_N B_{N-1} L_{N-1} x_{N-1}] \\ &= x'_{N-1} Q_{N-1} x_{N-1} + \mathbb{E}[w'_{N-1} K_N w_{N-1}] \\ &+ x'_{N-1} (A'_{N-1} K_N A_{N-1} + L'_{N-1} R_{N-1} L_{N-1} + L'_{N-1} B'_{N-1} K_N B_{N-1} L_{N-1} \\ &+ 2A'_{N-1} K_N B_{N-1} L_{N-1}) x_{N-1} \end{aligned}$$

Or,

$$\begin{aligned} L'_{N-1} (R_{N-1} + B_{N-1} K_N B_{N-1}) L_{N-1} + 2A'_{N-1} K_N B_{N-1} L_{N-1} \\ = A'_{N-1} K'_N B_{N-1} [(B'_{N-1} K_N B_{N-1} + R_{N-1})^{-1}]' \\ (R_{N-1} + B_{N-1} K_N B_{N-1}) L_{N-1} + 2A'_{N-1} K_N B_{N-1} L_{N-1} \\ = A'_{N-1} K_N B_{N-1} L_{N-1} \end{aligned}$$

Dès lors

$$\begin{aligned} J_{N-1}(x_{N-1}) &= x'_{N-1} (Q_{N-1} + A'_{N-1} K_N A_{N-1} \\ &\quad + A'_{N-1} K_N B_{N-1} L_{N-1}) x_{N-1} + \mathbb{E}[w'_{N-1} K_N w_{N-1}] \\ &= x'_{N-1} K_{N-1} x_{N-1} + \mathbb{E}[w'_{N-1} K_N w_{N-1}]. \end{aligned}$$

La matrice K_{N-1} est symétrique et elle est définie semi-positive, puisque cette dernière expression est ≥ 0 car Q_{N-1} , R_{N-1} , et K_N sont définies semi-positives. Cela prouve le résultat pour $k = N - 1$.

On suppose maintenant que le résultat tient pour $k + 1$ et on prouve que cela implique qu'il tient pour k . On a

$$J_{k+1}(x_{k+1}) = x'_{k+1} K_{k+1} x_{k+1} + \sum_{j=k+1}^{N-1} \mathbb{E}[w'_j K_{j+1} w_j]$$

où la somme est simplement une constante qui n'a pas d'influence sur la politique optimale.

On peut refaire exactement le même raisonnement que pour $k + 1 = N$, en remplaçant N par $k + 1$, et on obtient le résultat. \square

État partiellement observé.

Un résultat semblable tient aussi si on ne peut pas observer l'état x_k au complet, mais seulement une transformation linéaire

$$z_k = C_k x_k + v_k$$

où les v_k sont des vecteurs aléatoires indépendants entre eux et indépendants des w_k . La politique optimale a alors la forme

$$\mu_k^*(I_k) = L_k \mathbb{E}[x_k \mid I_k]$$

où $I_k = (u_0, \dots, u_{k-1}, z_0, \dots, z_k)$ est l'information disponible à l'étape k .

Modèle stationnaire sur horizon infini.

Supposons que $(A_k, B_k, Q_k, R_k) = (\textcolor{red}{A}, \textcolor{red}{B}, \textcolor{red}{Q}, \textcolor{red}{R})$ pour tout k .

Lorsque $k \rightarrow -\infty$, on s'attend à ce que $K_k \rightarrow K$ où K satisfait l'équation de Riccati algébrique

$$\textcolor{red}{K} = A'(K - KB(B'KB + R)^{-1}B'K)A + Q. \quad (1)$$

La politique optimale sur horizon infini sera alors stationnaire:

$$\mu^*(x) = Lx$$

où

$$\textcolor{red}{L} = -(B'KB + R)^{-1}B'KA.$$

On peut prouver que cela est vrai sous les conditions qui suivent.
D'abord quelques définitions.

Décomposons $\textcolor{red}{Q} = \textcolor{red}{C}'\textcolor{red}{C}$ où $\textcolor{red}{C}$ est $r \times n$ si $\textcolor{red}{r} = \text{rang}(Q)$.

(Définition 4.1.1 DPOC) La paire (A, B) est dite **contrôlable** si la matrice $n \times nm$

$$(B, AB, A^2B, \dots, A^{n-1}B)$$

est de rang n . La paire (A, C) est dite **observable** si la matrice $nm \times n$

$$\begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} = (C', A'C', \dots, (A^{n-1})'C')'$$

est de rang n . En d'autres termes (A', C') est contrôlable.

Contrôlabilité

La **contrôlabilité** implique qu'en l'absence de bruit (si $w_k = 0$), à partir de n'importe quelle position initiale, on peut ramener le système à 0 en n étapes par un choix approprié de u_0, \dots, u_{n-1} :

$$x_n = A^n x_0 + (B, AB, \dots, A^{n-1} B) \begin{pmatrix} u_{n-1} \\ \vdots \\ u_0 \end{pmatrix},$$

et donc

$$x_n - A^n x_0 = (B, AB, \dots, A^{n-1} B) \begin{pmatrix} u_{n-1} \\ \vdots \\ u_0 \end{pmatrix}.$$

Le caractère contrôlable garantit l'existence d'une solution au système ainsi construit (on peut isoler une sous-matrice $n \times n$ inversible).

Observabilité

L'observabilité implique qu'en observant Cx_0, \dots, Cx_{n-1} , on peut en déduire l'état initial x_0 du système $x_{k+1} = Ax_k$, car

$$\begin{pmatrix} Cx_0 \\ \vdots \\ Cx_{n-1} \end{pmatrix} = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{pmatrix} x_0.$$

L'observabilité est équivalente à la propriété qu'en l'absence de contrôle, $x_k \rightarrow 0$ si $Cx_k \rightarrow 0$.

Cela implique aussi que si $x'_k Q x_k \rightarrow 0$, alors $x_k \rightarrow 0$, comme $Q = CC'$.

En l'absence de bruit, on a

$$x_k = (A + BL)x_{k-1} = (A + BL)^k x_0. \quad (2)$$

Ce système en boucle fermée est dit **stable** si et seulement si $x_k \rightarrow 0$ lorsque $k \rightarrow \infty$, i.e., si et seulement si toutes les valeurs propres de la matrice $A + BL$ (qui sont en général des nombres complexes) sont **strictement** à l'intérieur du cercle unité.

Proposition. Si (A, B) est contrôlable et (A, C) est observable, alors

$$\lim_{k \rightarrow -\infty} K_k = K$$

où K est une matrice définie positive qui est l'unique solution de (1) dans l'espace des matrices définies semi-positives, et le système (2) est stable.

Preuve: Voir DPOC, Proposition 4.4.1.

Matrices aléatoires.

Supposons que les matrices A_k et B_k ne sont plus constantes, mais aléatoires. La politique optimale est encore de la forme

$$\mu_k^*(x_k) = L_k x_k,$$

si on prend soin de remplacer les formes quadratiques $A'_k K_{k+1} A_k$, $A'_k K_{k+1} B_k$, etc., par leurs **espérances conditionnelles** $\mathbb{E}[A'_k K_{k+1} A_k | K_{k+1}]$, $\mathbb{E}[A'_k K_{k+1} B_k | K_{k+1}]$, etc., dans les expressions pour K_k et L_k .

Par ailleurs, dans le cas stationnaire, le système est stable seulement si la mesure d'incertitude

$$T = \mathbb{E}[A^2] \mathbb{E}[B^2] - (\mathbb{E}[A] \mathbb{E}[B])^2$$

n'est pas trop grande. Si T dépasse un certain seuil (trop d'incertitude), K_k diverge lorsque $k \rightarrow -\infty$ et l'optimisation sur horizon infini n'a plus de sens.

Modèle d'inventaire

- x_k = Niveau d'inventaire au début du mois k ,
avant de commander (entier ou réel);
 u_k = Quantité commandée au début du mois k ;
 $y_k = x_k + u_k$ = Niveau d'inventaire après avoir commandé;
 w_k = Demande durant le mois k (sont indép.);
 $K + cu$ = Coût d'une commande de taille u ;
 $r(x_{k+1})$ = Coût d'inventaire payé à la fin du mois k ;

L'algorithme de la PD vu au chapitre 1 devient très coûteux lorsque le nombre de valeurs de x_k possibles est très grand.

Hypothèse: Supposons que $K = 0$ et que r est une fonction continue, convexe et telle que

$$\lim_{x \rightarrow \infty} r(x) = \lim_{x \rightarrow -\infty} (cx + r(x)) = \infty.$$

Exemple: $r(x) = p \max(0, -x) + h \max(0, x)$ où $p > c$ et $h > 0$.

On suppose que les inventaires négatifs sont permis:

$$x_{k+1} = x_k + u_k - w_k = y_k - w_k.$$

On n'a pas à tenir compte des revenus de vente. On a alors

$$\begin{aligned} J_N(x_N) &= 0; \\ J_k(x_k) &= \text{coût espéré total optimal pour les mois } k \text{ à } N \\ &= \min_{u_k \geq 0} (cu_k + \mathbb{E}[r(x_k + u_k - w_k) + J_{k+1}(x_k + u_k - w_k)]) \\ &= \min_{y_k \geq x_k} (cy_k + \mathbb{E}[r(y_k - w_k) + J_{k+1}(y_k - w_k)]) - cx_k \\ &= \min_{y_k \geq x_k} G_k(y_k) - cx_k \end{aligned}$$

où

$$G_k(y) = cy + \mathbb{E}[r(y - w_k) + J_{k+1}(y - w_k)]$$

est le coût espéré pour k à N si $x_k = 0$ et on commande y .

La proposition qui suit montre que la politique optimale est de toujours ramener le niveau d'inventaire à la valeur S_k qui minimise G_k si on est en dessous et de ne rien commander si on est au dessus.

Il suffit donc de trouver le min pour le cas où $x_k = 0$.

Proposition. Pour chaque $k < N$, G_k et J_k sont des fonctions convexes telles que $G_k(y) \rightarrow \infty$ et $J_k(y) \rightarrow \infty$ lorsque $y \rightarrow \pm\infty$. La politique optimale est définie par

$$\mu_k^*(x) = \max(0, S_k - x)$$

où S_k est la valeur de y qui minimise $G_k(y)$.

Preuve. Par induction arrière sur k .

Pour $k = N - 1$, $G_{N-1}(y) = cy + \mathbb{E}[r(y - w_{N-1})]$ est convexe car r est convexe, et $G_{N-1}(y) \rightarrow \infty$ quand $|y| \rightarrow \infty$ en supposant que la loi de w_{N-1} est "raisonnable". (Pour être rigoureux, il faudrait ajouter des hypothèses sur cette loi et divers détails techniques. Bertsekas suppose que les w_k prennent leurs valeurs dans un intervalle borné, mais on peut être moins restrictif.)

Il y a donc une valeur finie et unique de y qui minimise $G_{N-1}(y)$. Appelons-la S_{N-1} . Si $S_{N-1} \geq x_{N-1}$, cette valeur est la valeur optimale de y_{N-1} , i.e., la valeur optimale de u_{N-1} est $S_{N-1} - x_{N-1}$. On aura $J_{N-1}(x_{N-1}) = (S_{N-1} - x_{N-1})c + G_{N-1}(S_{N-1})$.

Si $S_{N-1} < x_{N-1}$, on est bloqué par la contrainte $y_{N-1} \geq x_{N-1} > S_{N-1}$. On prendra $y_{N-1} = x_{N-1}$, i.e., $u_{N-1} = 0$. Cela donne $\mu_{N-1}^*(x) = \max(0, S_{N-1} - x)$.

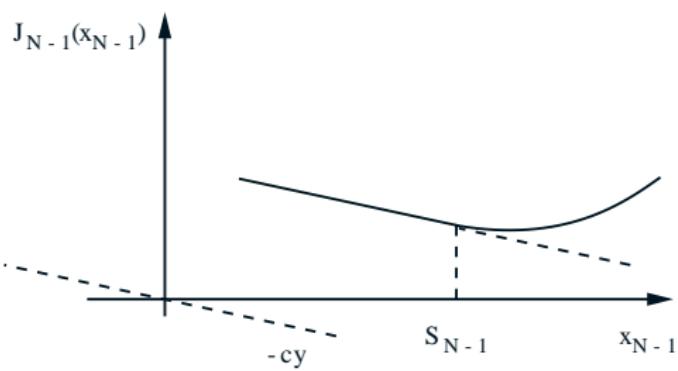
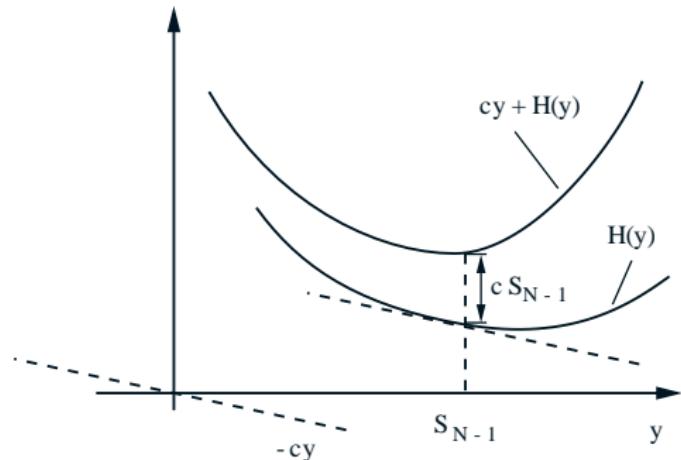
Il reste à montrer les propriétés de J_{N-1} . On a

$$J_{N-1}(x) = \begin{cases} (S_{N-1} - x)c + \mathbb{E}[r(S_{N-1} - w_{N-1})] & \text{if } x < S_{N-1}, \\ \mathbb{E}[r(x - w_{N-1})] & \text{if } x \geq S_{N-1}, \end{cases}$$

qui est continue, convexe, et tend vers l'infini lorsque $|x| \rightarrow \infty$, grâce aux propriétés de r .

Si on suppose que les propriétés tiennent pour $k+1$, on peut montrer par les mêmes arguments et en utilisant aussi l'hypothèse que J_{k+1} est convexe, que cela implique qu'elles tiennent pour k . \square

Exercice: complétez les détails de la preuve pour passer de $k+1$ à k .



Coût fixe positif.

Supposons maintenant qu'il y a un coût fixe $K > 0$ pour commander. On définit G_k de la même façon:

$G_k(y) = \text{coût espéré pour } k \text{ à } N \text{ si } x_k = 0, \text{ on a déjà payé le coût fixe, et on commande } y.$ On a

$$\begin{aligned} J_k(x_k) &= \min \left[G_k(x_k), \min_{u_k \geq 0} (K + G_k(x_k + u_k)) \right] - cx_k \\ &= \min \left[G_k(x_k), \min_{y_k \geq x_k} (K + G_k(y_k)) \right] - cx_k. \end{aligned}$$

Cette fonction J_k n'est pas convexe. Et G_k n'est pas nécessairement convexe non plus. Mais on peut prouver la forme de la politique optimale en utilisant une notion plus générale de convexité.

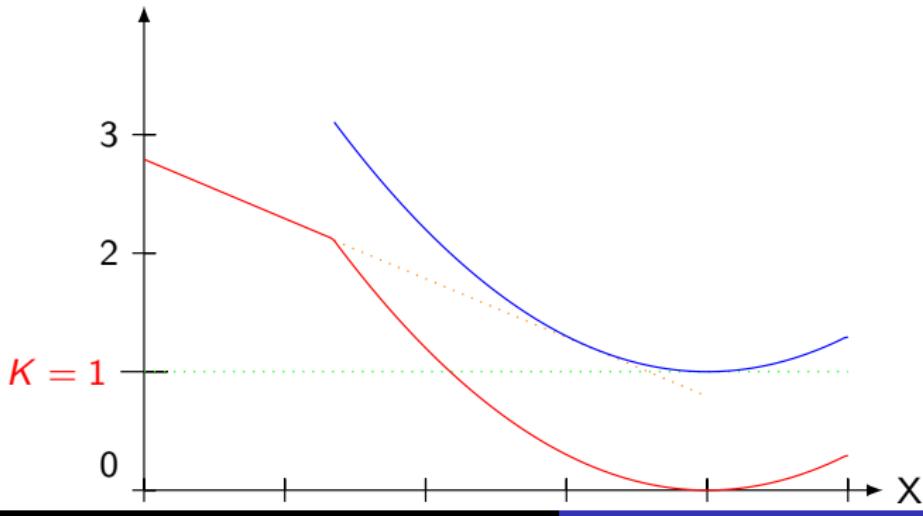
Ce genre de technique peut être utile en général pour déterminer et prouver la forme de la politique optimale dans le cas où il y a des coûts fixes pour les transactions.

K -convexité

Définition. (4.2.1, DPOC) Une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ est K -convexe, pour $K \geq 0$, si

$$K + g(z + y) \geq g(y) + z \left(\frac{g(y) - g(y - b)}{b} \right), \quad \forall z \geq 0, \quad b > 0, \quad y.$$

Ci-dessous, g (en rouge) est K -convexe pour $K = 1$, mais pas pour $K < 1$.

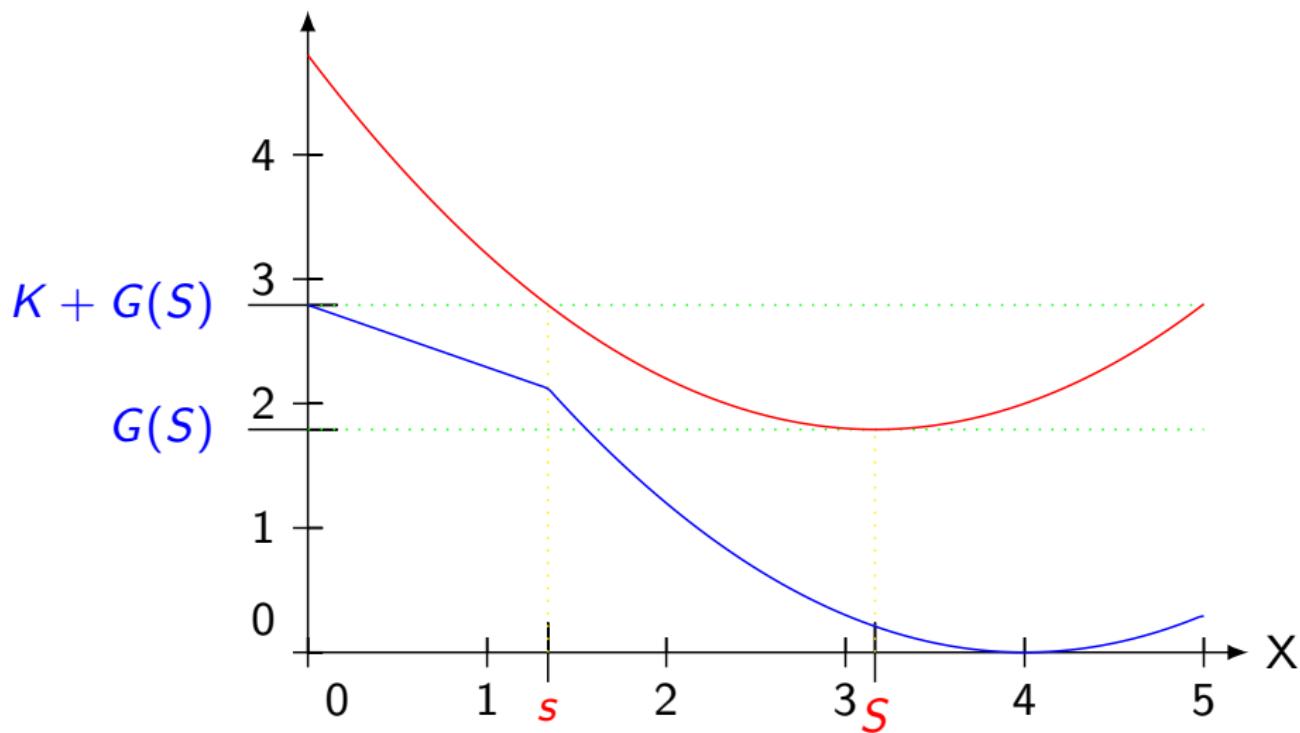


Une conséquence importante est qu'une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$ est **K -convexe**, pour $K \geq 0$, si

$$\frac{K + g(y + b_2) - g(y)}{b_2} \geq \frac{g(y) - g(y - b_1)}{b_1} \quad \forall y, \forall b_1 > 0, \forall b_2 > 0.$$

Lemme 4.2.1 (page 167).

- (a) Si g est convexe, alors g est K -convexe pour tout $K \geq 0$.
- (b) Si g_1 est K -convexe et g_2 L -convexe, alors
 $\alpha g_1 + \beta g_2$ est $(\alpha K + \beta L)$ -convexe.
- (c) Si g est K -convexe et W une v.a. aléatoire, alors
 $h(y) = \mathbb{E}[|g(y - W)|]$ est K -convexe si $h(y) < \infty$ pour tout y .
- (d) Si g est continue et K -convexe et si $\lim_{y \rightarrow \pm\infty} g(y) = \infty$, alors
il existe deux constantes $s \leq S$ telles que:
 - (i) $g(y) \geq g(S)$ pour tout y ;
 - (ii) $g(y) > g(s) = g(S) + K$ pour $y < s$;
 - (iii) $g(y)$ est décroissante pour $y < s$;
 - (iv) $g(y) \leq g(z) + K$ si $s \leq y \leq z$.



Proposition. Pour chaque $k < N$, G_k et J_k sont des fonctions continues et K -convexes telles que $G_k(y) \rightarrow \infty$ et $J_k(y) \rightarrow \infty$ lorsque $y \rightarrow \pm\infty$. La politique optimale est une politique de type (s, S) non stationnaire, définie par

$$\mu_k^*(x_k) = \begin{cases} S_k - x_k & \text{si } x_k < s_k; \\ 0 & \text{si } x_k \geq s_k; \end{cases}$$

où S_k est la valeur de y qui minimise $G_k(y)$ et s_k est la plus petite valeur de y telle que $G_k(y) = K + G_k(S_k)$.

Intérêt du résultat: simplifie beaucoup les calculs!

Preuve du lemme. (a), (b), (c) découlent directement de la définition. Prouvons (d).

(i): Puisque $g(y) \rightarrow \infty$ lorsque $|y| \rightarrow \infty$, g possède un minimum. Soit S un endroit où il est atteint.

(ii): Soit $s \leq S$ le plus petit y tel que $g(y) = g(S) + K$. Pour $y < s$, on a par la K -convexité

$$\frac{K + g(S) - g(s)}{S - s} \geq \frac{g(s) - g(y)}{s - y}.$$

Puisque $K + g(S) - g(s) = 0$, cela donne $g(y) \geq g(s) = K + g(S)$. Mais par définition de s , on a $g(y) > g(S) + K = g(s)$.

(iii) Pour $y_1 < y_2 < s$, on a de (i) et la K -convexité

$$0 > \frac{K + g(S) - g(y_2)}{S - y_2} \geq \frac{g(y_2) - g(y_1)}{y_2 - y_1}$$

et donc $g(y_1) > g(y_2)$.

(iv) On veut montrer que $K + g(z) \geq g(y)$ si $s \leq y \leq z$.

Si $y = s$ ou $y = z$, facile à vérifier.

Si $y > S$, alors par la K -convexité,

$$\frac{K + g(z) - g(y)}{z - y} \geq \frac{g(y) - g(S)}{y - S} \geq 0.$$

Si $s < y < S$, alors

$$\frac{g(s) - g(y)}{S - y} = \frac{K + g(S) - g(y)}{S - y} \geq \frac{g(y) - g(s)}{y - s}$$

qui implique

$$(g(s) - g(y))(y - s) \geq (g(y) - g(s))(S - y)$$

et donc

$$(g(s) - g(y))(S - s) \geq 0.$$

Ainsi

$$K + g(z) \geq K + g(S) = g(s) \geq g(y).$$

Plan de preuve de la proposition. Par induction sur k , comme d'habitude.

Pour $k = N - 1$, G_{N-1} est convexe, et donc K -convexe. On a

$$\begin{aligned} J_{N-1}(x) &= \min \left[G_{N-1}(x), \min_{y \geq x} (K + G_{N-1}(y)) \right] - cx \\ &= \begin{cases} K + G_{N-1}(s_{N-1}) - cx & \text{si } x < s_{N-1}, \\ G_{N-1}(x) - cx & \text{si } x \geq s_{N-1}. \end{cases} \end{aligned}$$

La politique μ_{N-1}^* qui fait atteindre le min a la forme voulue.

Pour montrer que J_{N-1} est K -convexe, on considère 3 cas:

(1) $y \geq s_{N-1}$, (2) $y < y + b_2 \leq s_{N-1}$, et (3) $y < s_{N-1} < y + b_2$.

Dans chaque cas, on montre que la définition de K -convexité tient.

On montre facilement que G_{N-1} est continue, car r l'est, et dès lors J_{N-1} est continue. $G_{N-1}(y) \rightarrow \infty$ et $J_{N-1}(y) \rightarrow \infty$ lorsque $y \rightarrow \pm\infty$ de la même façon que pour le cas où $K = 0$. On a donc le résultat pour $k = N - 1$.

On montre ensuite que si on suppose le résultat vrai pour $k + 1$, cela implique qu'il tient aussi pour k . \square

Un modèle de gestion de portefeuille.

x_0 = capital initial à investir;

x_k = capital de l'investisseur au début de la période k ;

n types d'actifs risqués et un type d'actif sans risque;

s_k = taux de rendement sans risque pour la période k ;

$e_k = (e_{k,1}, \dots, e_{k,n})$ = taux de rendement des n actifs risqués pour la période k . Ce sont des vecteurs aléatoires indépendants, dont la loi est telle que les espérances que nous allons considérer sont finies;

$u_k = (u_{k,1}, \dots, u_{k,n})$ = montants investis dans les différents actifs risqués à la période k ;

Le capital évolue selon:

$$\begin{aligned}x_{k+1} &= \sum_{i=1}^n e_{k,i} u_{k,i} + s_k(x_k - u_{k,1} - \dots - u_{k,n}) \\&= s_k x_k + \sum_{i=1}^n (e_{k,i} - s_k) u_{k,i}.\end{aligned}$$

L'investisseur a une fonction d'utilité U , et veut maximiser $\mathbb{E}[U(x_N)]$, l'espérance de l'utilité de son capital final.

Hypothèses: On suppose que U est concave et appartient à C^2 , i.e. est deux fois continûment différentiable. continue, que $\mathbb{E}[U(x_N)] < \infty$, et que U satisfait:

$$-U'(x)/U''(x) = a + bx.$$

Par exemple, les fonctions suivantes satisfont ces conditions:

$$U(x) = 1 - e^{-x/a}, \quad U(x) = \ln(x + a).$$

Les équations de la PD:

$$\begin{aligned} J_N(x_N) &= U(x_N), \\ J_k(x_k) &= \max_{\pi} \mathbb{E}[U(x_N) | x_k] \\ &= \max_{u_{k,1}, \dots, u_{k,n}} \mathbb{E} \left[J_{k+1} \left(s_k x_k + \sum_{i=1}^n (e_{k,i} - s_k) u_{k,i} \right) \right]. \end{aligned} \quad (3)$$

Proposition. Pour $0 \leq k < N$, on a

$$\mu_k^*(x_k) = \left(\frac{a}{s_{N-1} \cdots s_{k+1}} + b s_k x_k \right) \alpha_k$$

où $\alpha_k = (\alpha_{k,1}, \dots, \alpha_{k,n})$ dépend de la loi de probabilité de e_k mais pas de x_k . (On peut déterminer α_k via (3).) De plus, J_k satisfait

$$-\frac{J'_k(x)}{J''_k(x)} = \frac{a}{s_{N-1} \cdots s_k} + bx.$$

Preuve. (Pas complètement rigoureuse, comme dans DPOC.)

On suppose que la politique optimale existe et est différentiable en x_k (vrai si U et les lois des e_k sont suffisamment régulières).

Supposons qu'un portfolio optimal existe et est de la forme

$$\mu_k^*(x) = \alpha_k(x) \left(\frac{a}{s_{N-1} \cdots s_{k+1}} + b s_k x \right),$$

où les $\alpha_k(x) = (\alpha_{k,1}(x), \dots, \alpha_{k,n}(x))$ sont des vecteurs de fonctions différentiables.

En particulier,

$$\mu_{N-1}^*(x) = \alpha_{N-1}(x)(a + b s_{N-1} x).$$

On va montrer que $d\alpha_{k,i}(x)/dx = 0$, ce qui impliquera que $\alpha_{k,i}(x)$ est constante en x .

La preuve se fait par induction arrière sur k .

Prenons $k = N - 1$, $x_{N-1} = x$ et $s_{N-1} = s$.

Puisque $\mu_{N-1}^*(x)$ est le portefeuille optimal, on a

$$\begin{aligned} 0 &= \frac{d\mathbb{E}[U(x_N)]}{du_{N-1,i}} \Big|_{u_{N-1}=\mu_{N-1}^*(x)} \\ &= \frac{d}{du_{N-1,i}} \mathbb{E} \left[U \left(sx + \sum_{j=1}^n (e_{N-1,j} - s) u_{N-1,j} \right) \right] \\ &= \mathbb{E} \left[(e_{N-1,i} - s) U' \left(sx + \sum_{j=1}^n (e_{N-1,j} - s) \alpha_{N-1,j}(x)(a + bsx) \right) \right] \end{aligned}$$

en supposant que l'on peut échanger la dérivée et l'espérance (le théorème de convergence dominée de Lebesgue donne des conditions suffisantes pour cela).

En dérivant la dernière équation par rapport à x , et en supposant encore que l'on peut passer la dérivée à l'intérieur de \mathbb{E} , on obtient

$$0 = \mathbb{E} \left[(e_{N-1,i} - s) U''(x_N) \left(s + \sum_{j=1}^n (e_{N-1,j} - s) \left(\alpha_{N-1,j}(x) b s + \frac{d\alpha_{N-1,j}}{dx}(x)(a + b s x) \right) \right) \right] \quad (4)$$

$$= \mathbb{E} \left[U''(x_N)(a + b s x) \sum_{j=1}^n (e_{N-1,i} - s)(e_{N-1,j} - s) \frac{d\alpha_{N-1,j}(x)}{dx} \right] \quad (4)$$

$$+ \mathbb{E} \left[U''(x_N)(e_{N-1,i} - s)s \left(1 + \sum_{j=1}^n (e_{N-1,j} - s)\alpha_{N-1,j}(x)b \right) \right]. \quad (5)$$

Mais on sait que

$$\begin{aligned} -\frac{U'(x_N)}{U''(x_N)} &= a + b x_N = a + b \left(s x + \sum_{j=1}^n (e_{N-1,j} - s)\alpha_{N-1,j}(x)(a + b s x) \right) \\ &= (a + b s x) \left(1 + \sum_{j=1}^n (e_{N-1,j} - s)\alpha_{N-1,j}(x)b \right) \end{aligned} \quad (6)$$

Si on isole $U''(x_N)$ dans cette expression et on remplace dans le second membre de (5), on obtient

$$-\mathbb{E}[U'(x_N)(e_{N-1,i} - s)]s/(a + bsx),$$

et on a vu plus haut que cette espérance vaut 0.

On obtient ainsi, pour $i = 1, \dots, n$,

$$\mathbb{E} \left[(a + bsx) U''(x_N) \sum_{j=1}^n (e_{N-1,i} - s)(e_{N-1,j} - s) \frac{d\alpha_{N-1,j}(x)}{dx} \right] = 0.$$

ce système de n équations s'écrit sous forme matricielle:

$$(a + bsx)\mathbf{M}\nabla_x\alpha_{N-1}(x) = 0,$$

où $\nabla_x\alpha_{N-1}(x)$ est un vecteur dont l'élément j est $d\alpha_{N-1,j}(x)/dx$, et \mathbf{M} est une matrice $n \times n$ dont l'élément (i,j) est $\mathbb{E}[U''(x_N)(e_{N-1,i} - s)(e_{N-1,j} - s)]$.

Si on suppose que \mathbf{M} est inversible (ce qui est vrai sauf dans les cas dégénérés), on obtient que $d\alpha_{N-1,j}(x)/dx = 0$ pour tout j .

Montrons maintenant que

$$-\frac{J'_{N-1}(x)}{J''_{N-1}(x)} = \frac{a}{s} + bx. \quad (7)$$

Notant $J(x) = J_{N-1}(x) = \mathbb{E}[U(x_N)]$, On a

$$J(x) = \mathbb{E} \left[U \left(\left(1 + \sum_{j=1}^n (e_{N-1,j} - s) \alpha_{N-1,j} b \right) sx + \sum_{j=1}^n (e_{N-1,j} - s) \alpha_{N-1,j} a \right) \right]$$

On dérive 2 fois par rapport à x pour obtenir $J'(x)$ et $J''(x)$:

$$\begin{aligned} J'(x) &= \mathbb{E} \left[U'(x_N) \left(1 + \sum_{j=1}^n (e_{N-1,j} - s) \alpha_{N-1,j} b \right) s \right] \\ J''(x) &= \mathbb{E} \left[U''(x_N) \left(1 + \sum_{j=1}^n (e_{N-1,j} - s) \alpha_{N-1,j} b \right)^2 s^2 \right]. \end{aligned}$$

Puis, si on exprime $U''(x)$ en fonction de $U'(x)$ via (6) et si on remplace dans $J''(x)$, puis on divise l'expression de $J'(x)$ par cette expression de $J''(x)$, on obtient (7).

Induction: on suppose le résultat vrai pour $k + 1$, et on montre que cela implique qu'il est vrai pour k . Les détails sont semblables au cas $k = N - 1$. □

Remarque: La politique optimale à l'étape k est la même que si c'était la dernière étape et que la fonction d'utilité était J_{k+1} . Cette fonction J_{k+1} a exactement la même forme que la fonction d'utilité, avec a remplacé par a par $a/(s_{N-1} \cdots s_k)$ dans l'expression qui la définit. La politique optimale aura donc la même forme que si on était à la dernière étape, avec a remplacé par $a/(s_{N-1} \cdots s_k)$.

Si $a = 0$, ou si $s_k = 1$ pour tout k , alors la politique optimale est la même à toutes les étapes, et ne tient compte que de l'étape courante. On agit toujours comme si on était à l'étape $N - 1$. Cela s'appelle une politique **myope**.

Si $a \neq 0$, cela équivaut à dire que sous les hypothèses de ce modèle, on peut optimiser à l'étape k en faisant l'hypothèse qu'à partir de la prochaine étape tout notre argent sera investi sans risque. En effet, à la dernière étape, aucune décision d'investissement dans des actifs risqués n'était prise.

Cela revient à maximiser, par rapport à $u_k = (u_{k,1}, \dots, u_{k,n})$,

$$\mathbb{E} \left[U \left(s_{k+1} \cdots s_{N-1} \left(s_k x_k + \sum_{i=1}^n (e_{k,i} - s_k) u_{k,i} \right) \right) \right].$$

Il s'agit d'une politique **partiellement myope**.

Si $s_k > 1$ et $k \rightarrow -\infty$, on retrouve une stratégie myope.

Supposons qu'à chaque étape k , on doit décider si on arrête le système (et on encaisse un revenu ou un coût) ou si on continue. L'espace d'états X_k peut être partitionné en deux: les états où il est optimal de s'arrêter et ceux où il est optimal de continuer.

Exemple: vente d'un actif.

Durant chaque période k , $0 \leq k \leq N - 1$, on reçoit une offre w_k que l'on peut accepter à la fin de la période (au temps $k + 1$) si aucune offre n'a encore été acceptée.

On suppose que les w_k sont des v.a. indépendantes, et indépendantes de nos décisions (pas nécessairement i.i.d.).

On pose $u_k = 1$ si on vend à l'étape k , $u_k = 0$ sinon.

L'état du système à l'étape $k + 1$ sera

$$x_{k+1} = \begin{cases} w_k & \text{si } u_0 = \dots = u_k = 0; \\ \Delta & \text{sinon;} \end{cases}$$

où Δ est l'état dans lequel on a déjà vendu.

On pose $x_0 = u_0 = 0$.

Les décisions admissibles pour $k \geq 1$ sont $u_k = 0$ si $x_k = \Delta$, $u_N = 1$ si $x_N \neq \Delta$, et $u_k \in \{0, 1\}$ sinon.

Si $u_k = 1$, on reçoit $x_k = w_{k-1}$ au temps k , que l'on place au taux d'intérêt r pour les $N - k$ périodes restantes. Le revenu à l'étape k , actualisé au temps N , est donc

$$g_k(x_k, u_k, w_k) = \begin{cases} (1 + r)^{N-k} x_k & \text{si } u_k = 1; \\ 0 & \text{sinon.} \end{cases}$$

Soit $J_k(x_k)$ le **revenu espéré optimal** à partir du temps k jusqu'à la fin, actualisé au temps N , si on est dans l'état x_k . On obtient:

$$J_k(x_k) = \begin{cases} 0 & \text{si } x_k = \Delta; \\ x_N & \text{si } k = N \text{ et } x_N \neq \Delta; \\ \max((1+r)^{N-k}x_k, \mathbb{E}[J_{k+1}(w_k)]) & \text{si } k < N \text{ et } x_k \neq \Delta. \end{cases}$$

La **décision optimale**, lorsqu'on a encore le choix, est de vendre (accepter l'offre $x_k = w_{k-1}$) si et seulement si

$$x_k \geq \alpha_k \stackrel{\text{def}}{=} \frac{\mathbb{E}[J_{k+1}(w_k)]}{(1+r)^{N-k}}.$$

La **politique optimale** est déterminée par ces seuils $\alpha_1, \dots, \alpha_{N-1}$.

On pose $\alpha_N = 0$. Pour $x_k \neq \Delta$, le revenu total espéré actualisé au temps k est

$$\begin{aligned} V_k(x_k) &= J_k(x_k)/(1+r)^{N-k} \\ &= \begin{cases} x_N & \text{si } k = N; \\ \max(x_k, \mathbb{E}[V_{k+1}(w_k)]/(1+r)) & \text{si } k < N \end{cases} \\ &= \max(x_k, \alpha_k). \end{aligned}$$

On a

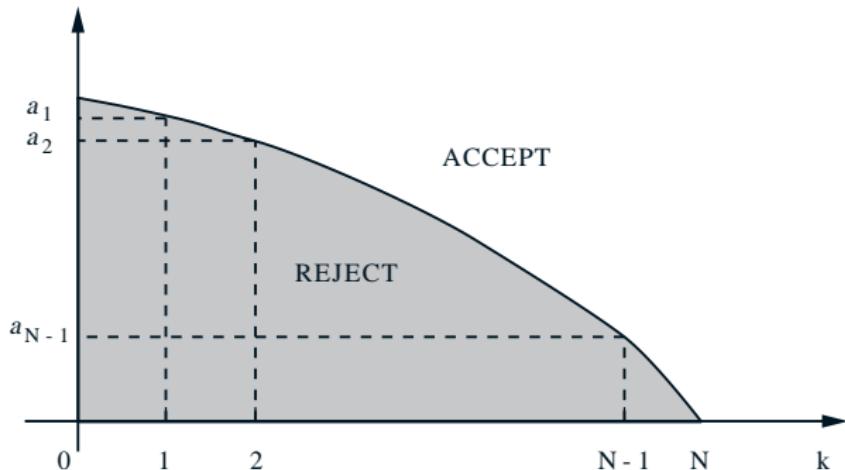
$$\begin{aligned} \alpha_k &= \mathbb{E}[J_{k+1}(w_k)]/(1+r)^{N-k} = \mathbb{E}[V_{k+1}(w_k)]/(1+r); \\ (1+r)\alpha_k &= \mathbb{E}[V_{k+1}(x_{k+1})] \\ &= \mathbb{E}[\max(x_{k+1}, \alpha_{k+1})] \\ &= \mathbb{E}[\max(w_k, \alpha_{k+1})] \\ &= \mathbb{E}[w_k \mathbb{I}[w_k > \alpha_{k+1}]] + \mathbb{E}[\alpha_{k+1} \mathbb{I}[w_k \leq \alpha_{k+1}]] \\ &= \mathbb{E}[w_k \mathbb{I}[w_k > \alpha_{k+1}]] + \alpha_{k+1} \mathbb{P}[w_k \leq \alpha_{k+1}], \end{aligned} \tag{8}$$

ce qui nous permet de calculer les α_k par récurrence.

On a $\alpha_N = 0$, $\alpha_{N-1} = \mathbb{E}[w_{N-1}]/(1+r)$, etc.

Peut-on caractériser davantage cette politique? Intuitivement, moins il reste de temps, moins on devrait être exigeant, car il nous reste moins d'opportunités; donc α_k devrait diminuer avec k .

Proposition. Si les w_k sont i.i.d., alors $\alpha_k \geq \alpha_{k+1}$ pour tout k .



Proposition. Si les w_k sont i.i.d., alors $\alpha_k \geq \alpha_{k+1}$ pour tout k .

Preuve. On note par w une v.a. qui a la même loi que les w_k .

Il suffit de montrer que $V_k(x) \geq V_{k+1}(x)$ pour $x \geq 0$ et $1 \leq k \leq N - 1$.

Cela se fait aisément par **induction sur k** .

Pour $k = N - 1$, on a $V_{N-1}(x) \geq x = V_N(x)$.

Si on suppose que $V_{k+1}(x) \geq V_{k+2}(x)$ pour tout x , alors

$$\begin{aligned}V_k(x) &= \max(x, \mathbb{E}[V_{k+1}(w)]/(1+r)) \\&\geq \max(x, \mathbb{E}[V_{k+2}(w)]/(1+r)) \\&= V_{k+1}(x).\end{aligned}$$

Que se passe-t-il lorsque $N \rightarrow \infty$?

Ou de façon équivalente, lorsque $k \rightarrow -\infty$?

Si on peut borner la suite des α_k , cela montrera que cette suite converge lorsque $k \rightarrow -\infty$. Par (8),

$$(1+r)\alpha_k \leq \mathbb{E}[w] + \alpha_{k+1}.$$

D'où l'on tire

$$\begin{aligned}\alpha_k &\leq \frac{\mathbb{E}[w] + \alpha_{k+1}}{1+r} \\ &\leq \frac{\mathbb{E}[w]}{1+r} + \frac{\mathbb{E}[w] + \alpha_{k+2}}{(1+r)^2} \\ &\leq \dots \\ &\leq \mathbb{E}[w] \sum_{i=1}^{\infty} \frac{1}{(1+r)^i} \\ &= \mathbb{E}[w] \frac{1+r}{r} < \infty.\end{aligned}$$

en supposant $r > 0$.

Ainsi, lorsque $k \rightarrow -\infty$, $\alpha_k \rightarrow \bar{\alpha}$ pour une constante $\bar{\alpha}$ qui satisfait

$$\begin{aligned}(1+r)\bar{\alpha} &= \mathbb{E}[\max(w, \bar{\alpha})] \\ &= \mathbb{E}[w \mathbb{I}[w > \bar{\alpha}]] + \bar{\alpha} \mathbb{P}[w \leq \bar{\alpha}].\end{aligned}$$

Lorsque l'horizon tend vers l'infini, la politique optimale est une politique stationnaire déterminée par le seuil $\bar{\alpha}$.

Dans ce modèle, $r > 0$ fait qu'il devient plus attrayant de vendre plus tôt, à prix égal. Et si on avait $r = 0$ pour le modèle sur horizon infini? Dans ce cas, à moins que l'on atteigne la valeur maximale que w peut prendre (ce qui est impossible pour plusieurs lois de prob.), on attendra toujours indéfiniment...

Achat avant une date limite.

Au lieu de vendre un actif, on veut (ou on doit) l'acheter au plus tard au début de la période N . L'état x_k à l'étape k est Δ si on a déjà acheté, et $x_k = w_{k-1}$ si on n'a pas encore acheté, où w_{k-1} est le **prix du marché** pour le produit dont on a besoin, au début de la période k .

Ici **on minimise** au lieu de maximiser.

Le **coût espéré total optimal** à partir du temps k , actualisé au temps k , si on est dans l'état $x_k \neq \Delta$, est

$$V_k(x_k) = \begin{cases} x_N & \text{si } k = N; \\ \min(x_k, \mathbb{E}[V_{k+1}(w_k)]/(1+r)) & \text{si } k < N \end{cases}$$
$$= \min(x_k, \alpha_k)$$

où $\alpha_N = \infty$, $\alpha_{N-1} = \mathbb{E}[w_{N-1}] / (1+r)$, etc.

On a

$$(1+r)\alpha_k = \mathbb{E}[V_{k+1}(w_k)] \quad (9)$$

$$= \mathbb{E}[\min(w_k, \alpha_{k+1})]$$

$$= \mathbb{E}[w_k \mathbb{I}[w_k \leq \alpha_{k+1}]] + \alpha_{k+1} \mathbb{P}[w_k > \alpha_{k+1}], \quad (10)$$

ce qui nous permet de calculer les α_k par récurrence.

La **politique optimale** est ainsi déterminée par des **seuils** α_k comme pour le problème du vendeur: on **achète si et seulement si** $x_k \leq \alpha_k$. On peut montrer, de la même façon:

Proposition. Si les w_k sont i.i.d., alors $\alpha_k \leq \alpha_{k+1}$ pour tout k .

Prix corrélés.

On peut généraliser le modèle au cas où les w_k sont dépendants.
Supposons, par exemple, que $x_0 = 0$ et

$$x_{k+1} = w_k = \lambda x_k + \xi_k, \quad 0 \leq k \leq N-1,$$

où $\lambda \in [0, 1)$ est une constante et les ξ_k sont des v.a. i.i.d. à valeur dans $[0, \infty)$, avec $\bar{\xi} = \mathbb{E}[\xi_k] > 0$. Pour simplifier, supposons que $r = 0$. On a alors, pour le problème de l'acheteur,

$$J_k(x_k) = \begin{cases} x_N & \text{si } k = N \text{ et } x_N \neq \Delta; \\ \min(x_k, \mathbb{E}[J_{k+1}(\lambda x_k + \xi_k)]) & \text{si } k < N \text{ et } x_k \neq \Delta. \end{cases}$$

Proposition. Pour $k \leq N - 1$, $J_k(x)$ est croissante et concave en x et on a $J_k(x) \leq J_{k+1}(x)$ pour tout x . La décision optimale à l'étape k est d'acheter si $x_k \leq \alpha_k$ et d'attendre sinon, où α_k est l'unique valeur positive qui satisfait

$$\alpha_k = \mathbb{E}[J_{k+1}(\lambda\alpha_k + \xi_k)].$$

De plus, on a $\alpha_k \leq \alpha_{k+1}$.

Preuve. Par induction sur k .

Pour $k = N - 1$, on a

$$J_{N-1}(x) = \min(x, \lambda x + \bar{\xi}) \leq x = J_N(x)$$

et le résultat tient avec $\alpha_{N-1} = \lambda\alpha_{N-1} + \bar{\xi}$, i.e., $\alpha_{N-1} = \bar{\xi}/(1 - \lambda)$.

Supposons que le résultat tient pour $k + 1$ et montrons qu'il tient alors aussi pour k . On a

$$\begin{aligned} J_k(x) &= \min(x, \mathbb{E}[J_{k+1}(\lambda x + \xi_k)]) \\ &= \min(x, \mathbb{E}[J_{k+1}(\lambda x + \xi_{k+1})]) \\ &\leq \min(x, \mathbb{E}[J_{k+2}(\lambda x + \xi_{k+1})]) = J_{k+1}(x) \end{aligned}$$

et $J_k(x)$ est croissante et concave en x car $J_{k+1}(x)$ l'est.

Par ailleurs, pour $x = 0$, $\mathbb{E}[J_{k+1}(\xi_k)] > 0$, ce qui implique que $J_k(x) = x$ pour x proche de 0. De plus,

$J_k(x) = \varphi(x) \stackrel{\text{def}}{=} \mathbb{E}[J_{k+1}(\lambda x + \xi_k)]$ pour $x \geq \alpha_{k+1}$, car

$$\alpha_{k+1} = \mathbb{E}[J_{k+2}(\lambda \alpha_{k+1} + \xi_{k+1})] \geq \mathbb{E}[J_{k+1}(\lambda \alpha_{k+1} + \xi_k)].$$

Puisque $J_k(x)$ est concave, cela implique aussi que $\varphi'(\alpha_{k+1}) < 1$ et que la fonction $\varphi(x)$ doit avoir une pente strictement < 1 lorsqu'elle croise $f(x) = x$. La valeur de x où le croisement se produit est donc unique et doit se trouver dans l'intervalle $(0, \alpha_{k+1})$. C'est $x = \alpha_k$. Cela complète la preuve. \square

Option de type américaine pour une action dont le prix suit une marche aléatoire.

(Ross, "Introduction to Stochastic Dynamic Programming", 1983, Section 1.3.)

On suppose que le **prix de l'action** de la firme ABC est x_k au jour k , et évolue selon une marche aléatoire:

$$x_{k+1} = x_k + w_k = x_0 + \sum_{i=0}^k w_i$$

où les w_i sont des v.a. i.i.d. de moyenne μ .

On a une **option d'achat** pour une action au prix fixe K .

On peut l'exercer à l'un des N premiers jours.

Si on l'exerce au jour k , notre **profit** est $\max(0, x_k - K)$.

On veut maximiser notre profit espéré.

Soit $J_k(x_k)$ le profit espéré optimal si le prix est x_k au jour k et que l'on n'a pas encore exercé l'option. On a

$$J_k(x_k) = \begin{cases} \max(0, x_N - K) & \text{si } k = N; \\ \max(x_k - K, \mathbb{E}[J_{k+1}(x_k + w_k)]) & \text{si } k < N. \end{cases}$$

Proposition. $J_k(x)$ est continue et croissante en x et décroissante en k , tandis que $J_k(x) - x$ est décroissante en x .

La **politique optimale** consiste à exercer l'option au jour k si et seulement si $x_k \geq \alpha_k$, où α_k est la plus petite valeur qui satisfait $J_k(\alpha_k) = \alpha_k - K$. De plus, $\alpha_{k+1} \leq \alpha_k$ pour $k < N$.

Preuve. On montre la première partie par induction sur k .

$J_N(x) = \max(0, x - K)$ est continue et croissante en x , et
 $J_N(x) - x = \max(-x, -K)$ est décroissante en x .

De plus,

$$\begin{aligned} J_{N-1}(x) &= \max(x - K, \mathbb{E}[J_N(x + w_{N-1})]) \\ &\geq \max(x - K, 0, \mathbb{E}[x + w_{N-1} - K]) \geq J_N(x). \end{aligned}$$

Si on suppose que J_{k+1} a les propriétés désirées, pour $k < N$,

$$J_k(x) = \max(x - K, \mathbb{E}[J_{k+1}(x + w_k)])$$

est continue et croissante en x car J_{k+1} l'est,

$$J_k(x) \geq \max(x - K, \mathbb{E}[J_{k+2}(x + w_k)]) = J_{k+1}(x),$$

et

$$J_k(x) - x = \max(-K, \mathbb{E}[J_{k+1}(x + w_k) - (x + w_k)] + \mathbb{E}[w_k])$$

est décroissante en x grâce à l'hypothèse sur J_{k+1} .

Il est optimal d'exercer l'option à l'étape k si et seulement si
 $J_k(x_k) = x_k - K$. Posons

$$\alpha_k = \inf\{x : J_k(x) - x = -K\}.$$

($\alpha_k = \infty$ si $J_k(x) - x$ n'atteint jamais $-K$.)

Pour $x \geq \alpha_k$, on a $J_k(x) - x \leq J_k(\alpha_k) - \alpha_k = -K$ et il est donc optimal d'exercer. La politique optimale a donc la forme spécifiée.

Le fait que $\alpha_{k+1} \leq \alpha_k$ découle du fait que $J_{k+1}(x) \leq J_k(x)$.

Supposons en effet par l'absurde que $\alpha_{k+1} > \alpha_k$. Comme $J_{k+1}(\alpha_k) \geq \alpha_k - K$, nous avons à présent (par définition de α_k et α_{k+1}) $J_{k+1}(\alpha_k) > \alpha_k - K = J_k(\alpha_k)$. \square

La règle nous donne des seuils, mais pas de règle précise de calcul.
Regardons par exemple l'étape $N - 1$. Nous exerçons si

$$\begin{aligned}x_{N-1} - K &> \mathbb{E}_w [J_N(x_{N-1} - K + w)] \\&= \mathbb{E}_w [\max\{0, x_{N-1} - K + w\}] \\&= \mathbb{E}_w [(x_{N-1} - K + w) \mathbb{I}[x_{N-1} - K + w > 0]] \\&= (x_{N-1} - K) P[x_{N-1} - K + w > 0] + \\&\quad \mathbb{E}_w [w \mathbb{I}[x_{N-1} - K + w > 0]],\end{aligned}$$

et donc si

$$(x_{N-1} - K) P[w \leq x_{N-1} - K] > \mathbb{E}_w [w \mathbb{I}[w > x_{N-1} - K]].$$

Notez que l'indicatrice a pour effet de "gommer" les valeurs négatives de w .

L'exemple n'a de sens que si $\mu < 0$. Comme aucun facteur d'actualisation n'est présent ici, seul l'évolution directe de x importe. Si $\mu \geq 0$, $E_w[x_{k+1}] \geq E_w[x_k]$, et il est préférable d'attendre. Si $\mu < 0$, même si en moyenne, la valeur de l'option baisse, les variations aléatoires peuvent conduire à une valeur élevée de l'option en certaines périodes. La stratégie d'exercice se base sur cette observation.

Ensemble d'arrêt absorbant et règle du “un coup à l'avance”

On considère un modèle **stationnaire** général où en plus (ou à l'intérieur) de la décision u_k à l'étape k , on peut arrêter le système et payer un coût terminal $t(x_k)$ fonction de l'état x_k , ou bien continuer. Si on se rend à l'étape N , on doit payer $t(x_N)$. On a $J_N(x_N) = t(x_N)$ et

$$J_k(x_k) = \min \left(t(x_k), \min_{u \in U(x_k)} \mathbb{E}_{w_k} [g(x_k, u, w_k) + J_{k+1}(f(x_k, u, w_k))] \right),$$

pour $k < N$, et il est **optimal de s'arrêter** au temps k si et seulement si x_k est dans

$$T_k = \left\{ x : t(x) \leq \min_{u \in U(x)} \mathbb{E} [g(x, u, w) + J_{k+1}(f(x, u, w))] \right\}.$$

On voit que $J_{N-1}(x) \leq J_N(x)$ et on montre facilement par induction que $J_k(x) \leq J_{k+1}(x)$, pour tout x et $k < N$. Il en découle que $T_k \subseteq T_{k+1}$ pour $0 \leq k < N$.

Proposition. Si l'ensemble T_{N-1} est absorbant tant que l'on ne s'arrête pas, i.e., $\{x \in T_{N-1}\} \Rightarrow \{f(x, u, w) \in T_{N-1}\}$, alors $T_k = T_{N-1}$ pour tout $k < N$.

Preuve. Si $x_{N-2} = x \in T_{N-1}$, alors $x_{N-1} = f(x, u, w) \in T_{N-1}$, de sorte que $J_{N-1}(x_{N-1}) = t(x_{N-1})$, et donc

$$\begin{aligned} t(x) &\leq \min_{u \in U(x)} \mathbb{E}[g(x, u, w) + J_N(f(x, u, w))] \\ &= \min_{u \in U(x)} \mathbb{E}[g(x, u, w) + t(f(x, u, w))] \\ &= \min_{u \in U(x)} \mathbb{E}[g(x, u, w) + J_{N-1}(f(x, u, w))], \end{aligned}$$

ce qui veut dire que $x \in T_{N-2}$. On a donc $T_{N-1} \subseteq T_{N-2}$, et donc $T_{N-2} = T_{N-1}$. On montre de la même manière que $T_k = T_{N-1}$ pour tout k . (Ou encore, par induction sur k .) \square

Ainsi, si T_{N-1} est absorbant, la politique optimale à chaque étape est de s'arrêter si et seulement s'il est préférable de s'arrêter maintenant plutôt que de continuer et de s'arrêter obligatoirement à la prochaine étape. Autrement dit, **il suffit de regarder un coup à l'avance** (“one-step look ahead policy”).

Cette proposition s'applique dans les deux exemples qui suivent.

Exemple: Problème du vendeur avec retention des offres.

Revenons au problème de vente d'un actif. On suppose maintenant que les offres qui ne sont pas acceptées immédiatement ne sont pas rejetées, elles peuvent être acceptées plus tard.

L'état x_k indique l'offre courante et on a

$$x_{k+1} = \max(x_k, w_k).$$

Question: Que fera-t-on si $r = 0$?

Les récurrences pour V_k deviennent

$$V_N(x_N) = x_N,$$

$$V_k(x_k) = \max(x_k, \mathbb{E}[V_{k+1}(\max(x_k, w_k))]/(1+r)), \quad k < N.$$

L'ensemble d'arrêt optimal à l'étape $N - 1$ est

$$T_{N-1} = \{x : x \geq \mathbb{E}[\max(x, w)]/(1 + r)\} = \{x : x \geq \bar{\alpha}\},$$

où $\bar{\alpha}$ est la solution de

$$(1 + r)\bar{\alpha} = \mathbb{E}[\max(\bar{\alpha}, w)] = \bar{\alpha}\mathbb{P}[w \leq \bar{\alpha}] + \mathbb{E}[w\mathbb{I}[w > \bar{\alpha}]]$$

si la solution existe, $\bar{\alpha} = \infty$ sinon.

Cet ensemble T_{N-1} est absorbant, car x_k ne peut jamais diminuer lorsque k augmente. On a donc $T_k = T_{N-1}$ pour tout k , i.e., on accepte la première offre qui atteint $\bar{\alpha}$.

Exemple: Un voleur qui sait calculer.

à chaque période k (e.g., chaque nuit), un voleur peut tenter un nouveau vol ou prendre sa retraite avec son profit déjà accumulé, x_k . S'il tente un vol, avec probabilité p il se fait prendre et perd tout, et avec probabilité $1 - p$ il fait un gain w_k (aléatoire).

Après N périodes, il doit nécessairement se retirer avec son profit x_N , s'il ne l'a pas fait avant et s'il n'a pas été pris.

Il veut maximiser son profit espéré total, $\mathbb{E}[x_N]$.

Les équations de la PD:

$$J_N(x_N) = x_N,$$

$$J_k(x_k) = \max(x_k, (1 - p)\mathbb{E}[J_{k+1}(x_k + w_k)]), \quad k < N.$$

On a ici

$$\begin{aligned} T_{N-1} &= \{x : x \geq (1-p)(x + \mathbb{E}[w])\} \cup \{\Delta\} \\ &= \{x : x \geq \mathbb{E}[w](1-p)/p\} \cup \{\Delta\} \end{aligned}$$

où Δ est l'état d'arrestation. Cet ensemble est **absorbant** au sens de la proposition. La politique optimale est donc de se retirer dès que le profit accumulé x_k atteint $\bar{\alpha} = \mathbb{E}[w](1-p)/p$, peu importe la valeur de k .

Ce modèle s'applique aussi à d'autres types de situations.

Problèmes d'ordonnancement et argument d'échange de voisins.

On a un ensemble de N tâches à accomplir et on doit les ordonner.

Par exemple, N pièces à fabriquer dans un atelier, où N voitures à réparer, ou N dossiers à étudier, ou N articles à arbitrer, etc. On veut minimiser un critère de performance qui s'exprime comme l'espérance de la somme des coûts pour les différentes tâches.

Algorithme de PD standard: l'état à l'étape k est l'ensemble des tâches restantes.

Les modèles considérés ici sont stochastiques, mais l'information obtenue au cours des premières étapes n'est pas utile pour améliorer les décisions futures, de sorte que la politique optimale sera une politique en **boucle ouverte** (on peut trouver l'ordonnancement optimal dès le départ).

L'argument d'échange de voisins dit que si

$$L = \{i_0, i_1, \dots, i_{k-1}, \textcolor{red}{i}, \textcolor{red}{j}, i_{k+2}, \dots, i_{N-1}\}$$

est un ordonnancement optimal et

$$L' = \{i_0, i_1, \dots, i_{k-1}, \textcolor{teal}{j}, \textcolor{teal}{i}, i_{k+2}, \dots, i_{N-1}\}$$

alors le coût espéré total pour L ne doit pas dépasser celui pour L' .
En général, cela ne donne que des conditions nécessaires
d'optimalité, mais dans certains cas il devient évident que ces
conditions sont aussi suffisantes.

Exemple: Ordonnancement des questions d'un quiz.

Il y a N questions, auxquelles on peut répondre dans l'ordre que l'on veut. On répondra correctement à la question i avec probabilité p_i , et si on le fait on gagne R_i . Dès que l'on échoue à une question, c'est terminé. On veut maximiser notre gain total espéré.

A noter que la politique optimale est en boucle ouverte, car une fois que l'on aura répondu aux k premières questions, on n'aura pas davantage d'information qu'au départ qui puisse justifier de changer notre ordonnancement des $N - k$ questions restantes.

Soit $J(S)$ le revenu espéré pour une suite ordonnée de questions S , et soient L et L' comme à la page précédente, où L est optimal:

$$\begin{aligned}L &= \{i_0, i_1, \dots, i_{k-1}, i, j, i_{k+2}, \dots, i_{N-1}\}, \\L' &= \{i_0, i_1, \dots, i_{k-1}, j, i, i_{k+2}, \dots, i_{N-1}\}, \\J(L) &= J(\{i_0, \dots, i_{k-1}\}) + p_{i_0} \cdots p_{i_{k-1}} (\textcolor{teal}{p_i R_i} + \textcolor{teal}{p_i p_j R_j}) \\&\quad + p_{i_0} \cdots p_{i_{k-1}} p_i p_j J(\{i_{k+2}, \dots, i_{N-1}\}); \\J(L') &= J(\{i_0, \dots, i_{k-1}\}) + p_{i_0} \cdots p_{i_{k-1}} (\textcolor{teal}{p_j R_j} + \textcolor{teal}{p_j p_i R_i}) \\&\quad + p_{i_0} \cdots p_{i_{k-1}} p_j p_i J(\{i_{k+2}, \dots, i_{N-1}\}).\end{aligned}$$

Puisque L est optimale, on doit avoir $J(L) \geq J(L')$, i.e.,

$$\begin{aligned}\textcolor{teal}{p_i R_i} + \textcolor{teal}{p_i p_j R_j} &\geq \textcolor{teal}{p_j R_j} + \textcolor{teal}{p_j p_i R_i}, \quad \text{i.e.,} \\\textcolor{teal}{p_i R_i / (1 - p_i)} &\geq \textcolor{teal}{p_j R_j / (1 - p_j)}.\end{aligned}$$

Il faut ordonner les tâches par ordre décroissant de leur valeur de $\textcolor{teal}{p_i R_i / (1 - p_i)}$.

Ordonnancement sur un processeur, avec revenu actualisé.

N tâches. La tâche i requiert un temps T_i aléatoire de loi connue et rapporte un revenu $\alpha^t R_i$ si on la complète au temps t , où $\alpha < 1$ est un facteur d'actualisation. Les v.a. T_i sont supposées indépendantes.

Les valeurs des T_i des tâches déjà réalisées n'affectent les revenus futurs que via un facteur d'actualisation commun, donc n'affectent pas l'optimisation de l'ordonnancement des tâches qui restent. On peut ainsi déterminer l'ordonnancement optimal dès le départ.

Soient L et L' comme dans l'exemple précédent, et soit t_k l'instant où on termine la tâche $k - 1$ et débute la suivante.

Puisque les T_i sont indépendants, on a

$$\begin{aligned} J(L) &= J(\{i_0, \dots, i_{k-1}\}) + \mathbb{E}[\alpha^{t_k+T_i} R_i + \alpha^{t_k+T_i+T_j} R_j] \\ &\quad + \mathbb{E}[\alpha^{t_k+T_i+T_j} J(\{i_{k+2}, \dots, i_{N-1}\})] \\ &= J(\{i_0, \dots, i_{k-1}\}) + \mathbb{E}[\alpha^{t_k}] \mathbb{E}[\alpha^{T_i} R_i + \alpha^{T_i+T_j} R_j] \\ &\quad + \mathbb{E}[\alpha^{t_k+T_i+T_j}] J(\{i_{k+2}, \dots, i_{N-1}\}). \end{aligned}$$

Tout comme dans l'exemple précédent, $J(L) \geq J(L')$ implique que

$$\begin{aligned} \mathbb{E}[\alpha^{T_i} R_i + \alpha^{T_i+T_j} R_j] &\geq \mathbb{E}[\alpha^{T_j} R_j + \alpha^{T_j+T_i} R_i], \quad \text{i.e.,} \\ R_i \mathbb{E}[\alpha^{T_i}] + R_j \mathbb{E}[\alpha^{T_i}] \mathbb{E}[\alpha^{T_j}] &\geq R_j \mathbb{E}[\alpha^{T_j}] + R_i \mathbb{E}[\alpha^{T_j}] \mathbb{E}[\alpha^{T_i}] \quad \text{i.e.,} \\ \frac{p_i R_i}{1 - p_i} &\geq \frac{p_j R_j}{1 - p_j} \end{aligned}$$

où $p_k = \mathbb{E}[\alpha^{T_k}]$. On doit ordonner les tâches par ordre décroissant de leur valeur de $p_i R_i / (1 - p_i)$, exactement comme dans l'exemple précédent.

Ordonnancement sur 2 processeurs en série.

N tâches. La tâche i requiert un temps a_i sur la machine A , puis un temps b_i sur la machine B (déterministe). On veut minimiser le temps total de traitement.

Soit X_k l'ensemble des tâches qui restent à traiter sur A et τ_k la quantité de travail en attente ("workload") sur la machine B , à l'instant où l'on termine la k -ième tâche sur la machine A (étape k). L'état à l'étape k est (X_k, τ_k) .

Si i est la prochaine tâche sur A , on aura à l'étape $k + 1$:

$$X_{k+1} = X_k - \{i\}, \quad \tau_{k+1} = \max(0, \tau_k - a_i) + b_i.$$

Lorsqu'on termine la dernière tâche sur A , il reste un temps τ_N avant de tout terminer. La récurrence de la PD s'écrit donc

$$J_N(\phi, \tau_N) = \tau_N,$$

$$J_k(X_k, \tau_k) = \min_{i \in X_k} [a_i + J_{k+1}(X_k - \{i\}, \max(0, \tau_k - a_i) + b_i)]$$

pour $k < N$.

Soient L et L' comme dans les exemples précédents. L'argument d'échange des voisins i et j nous donne la **condition d'optimalité**

$$J_{k+2}(X_k - \{i, j\}, \tau_{i,j}) \leq J_{k+2}(X_k - \{i, j\}, \tau_{j,i})$$

où τ_{ij} et τ_{ji} sont les valeurs de τ_{k+2} pour L et L' , respectivement.
Ceci implique que $\tau_{ij} \leq \tau_{ji}$ car $J_k(X_k, \tau_k)$ est croissante en τ_k .

Pour L , on a

$$\begin{aligned}\tau_{ij} &= \tau_{k+2} = \max(0, \tau_{k+1} - a_j) + b_j \\&= \max(0, \max(0, \tau_k - a_i) + b_i - a_j) + b_j \\&= \max(0, \max(a_i, \tau_k) + b_i - a_i - a_j) + b_j \\&= \max(-b_i + a_i + a_j, \max(a_i, \tau_k)) + b_i - a_i - a_j + b_j \\&= \max(\tau_k, a_i, a_i + a_j - b_i) - a_i - a_j + b_i + b_j.\end{aligned}$$

De même,

$$\tau_{ji} = \max(\tau_k, a_j, a_i + a_j - b_j) - a_i - a_j + b_i + b_j.$$

Dès lors,

$$\tau_{ij} \leq \tau_{ji} \Leftrightarrow \max(\tau_k, a_i, a_i + a_j - b_i) \leq \max(\tau_k, a_j, a_i + a_j - b_j)$$

Ainsi, $\tau_k \geq \max(a_j, a_i + a_j - b_j)$, implique

$\tau_k \geq \max(a_i, a_i + a_j - b_i)$, et on a $\tau_{ij} = \tau_{ji}$. Autrement dit, L et L' ont le même coût.

Si $\tau_k < \max(a_j, a_i + a_j - b_j)$, alors on doit avoir $\max(a_i, a_i + a_j - b_i) \leq \max(a_j, a_i + a_j - b_j)$, et on peut montrer que ceci est équivalent à

$$\min(a_i, b_j) \leq \min(a_j, b_i).$$

Cela détermine un ordre unique pour chaque paire (i, j) , sauf s'il y a égalité ci-haut. Dans le cas où il y a égalité, l'ordre n'a pas d'importance.

Mis à part ces cas d'égalité, l'ordre unique pour chaque paire détermine un ordre unique pour l'ensemble des tâches.

Intuition: Il faut éviter de laisser la machine B innocupée, car c'est là que le temps est perdu. Pour cela, on donne priorité aux tâches qui demandent très peu de temps sur la machine A et/ou beaucoup de temps sur la machine B.

Algorithme d'ordonnancement:

$T \leftarrow \{1, \dots, N\}; \quad L1 \leftarrow \phi; \quad L2 \leftarrow \phi;$

TANTQUE $T \neq \phi$ FAIRE

$k \leftarrow \arg \min_{i \in T} \min(a_i, b_i);$

SI $a_k < b_k$, mettre k à la fin de $L1$

SINON mettre k au début de $L2$;

L'ordonnancement choisi est la liste $L1$ suivie de la liste $L2$.

Proposition. Cet algorithme retourne un ordonnancement optimal.

Preuve: Exercice.

PD pour évaluer une option américano-asiatique

Le prix d'un actif évolue selon un processus stochastique $\{S(t), t \geq 0\}$ à valeurs dans $[0, \infty)$. Ce processus est observé aux instants (fixes) $0 = t_0 < t_1 < \dots < t_n = T$. Un contrat financier (option) de type américain-bermudien procure un revenu immédiat $g_j(S(t_1), \dots, S(t_j))$ si on exerce l'option au temps t_j , pour $m^* \leq j \leq n$.

En supposant que le processus évolue sous la mesure de risque neutre et que le taux d'intérêt est r , la valeur du contrat au temps t_j si $(S(t_0), \dots, S(t_j)) = (s_0, \dots, s_j)$ est $v_j(s_0, \dots, s_j)$, où les fonctions v_j obéissent aux équations de récurrence:

$$v_j(s_0, \dots, s_j) = \begin{cases} \max(0, g_n(s_1, \dots, s_n)) & \text{si } j = n; \\ \max(g_j(s_1, \dots, s_j), v_j^h(s_0, \dots, s_j)) & \text{si } m^* \leq j < n; \\ v_j^h(s_0, \dots, s_j) & \text{si } j < m^*. \end{cases}$$

avec la valeur de rétention

$$v_j^h(s_0, \dots, s_j) = \mathbb{E} \left[e^{-r(t_{j+1}-t_j)} v_{j+1}(s_0, \dots, s_j, S(t_{j+1})) \mid s_0, \dots, s_j \right]$$

Si g est très simple, on peut parfois trouver une formule analytique pour v_j .

Si v_j peut s'écrire comme une fonction de peu de variables, alors on peut résoudre numériquement.

Exemple: Option de type asiatique, pour laquelle

$$g_j(s_1, \dots, s_j) = \max(0, \bar{s}_j - K)$$

pour une constante $K > 0$, où $\bar{s}_j = (s_1 + \dots + s_j)/j$.

Dans ce cas, v_j ne dépend que de (s_j, \bar{s}_j) (état à 2 dimensions):

$$v_j(s_j, \bar{s}_j) = \begin{cases} \max(0, \bar{s}_n - K) & \text{pour } j = n; \\ \max(\bar{s}_j - K, v_j^h(s_j, \bar{s}_j)) & \text{pour } m^* \leq j < n; \\ v_j^h(s_j, \bar{s}_j) & \text{pour } j < m^*. \end{cases}$$

où

$$v_j^h(s_j, \bar{s}_j) = e^{-r(t_{j+1} - t_j)} \mathbb{E}[v_{j+1}(S(t_{j+1}), (j\bar{s}_j + S(t_{j+1}))/j) | s_j].$$

Pour $m^* \leq j < n$, on exercera l'option ssi $\bar{s}_j - K \geq v_j^h(s_j, \bar{s}_j)$.

Quelle est la forme de la région d'exercice? Et des fonctions v_j^h et v_j ?

Pour répondre à ces questions, il faut des hypothèses sur le processus S .

Mouvement Brownien géométrique.

$$S(t) = S(0)e^{(r-\sigma^2/2)t+\sigma B(t)}$$

où r est le taux d'intérêt sans risque, σ est la volatilité, et $B(\cdot)$ est un mouvement Brownien standard.

Pour $t_2 > t_1 \geq 0$, $B(t_2) - B(t_1)$ est une v.a. normale de moyenne 0 et variance $t_2 - t_1$, indépendante des accroissements de $B(\cdot)$ en dehors de $[t_1, t_2]$.

On peut par exemple supposer $t_{j+1} - t_j = h$, $\forall j$, et poser $\rho = e^{-rh}$.
Technique... voir Ben-Ameur, Breton et L'Ecuyer, Management Science, 2002.