Décision dans l'incertain

École Nationale des Ponts et Chaussées
29 Janvier 2024

Jean-Philippe Chancelier   Bernard LAPEYRE

Table des matières

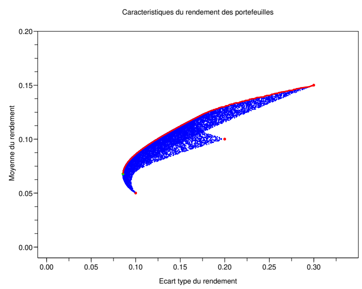
Chapitre 1

Probabilité, Espérance, Variance

Choix optimal de portefeuille

Probabilité, Espérance, Variance

Choix optimal de portefeuille



J.-P. Chancelier, B. Lapeyre

`http://cermics.enpc.fr/~jpc/decision-incertain`

Plan

Fonctionnement du cours

- Site du cours (Jean-Philippe Chancelier et Bernard Lapeyre)

  `http://cermics.enpc.fr/~jpc/decision-incertain`
- 6 séances : 1h de cours, 1h30 de TP informatique
  - Introduction aux chaînes de Markov et problèmes de décision associés
  - Résultats mathématiques en cours et concrétisés par des TP de simulation.
  - une conférence qui remplace le dernier cours.
- Enseignants (cours) :
  - Jean-Philippe Chancelier (`jean-philippe.chancelier@enpc.fr`)
- Enseignants (TPs informatiques) :
  - Roberta Flenghi (`roberta.flenghi@enpc.fr`)
  - Zoé Fornier (`zoe.fornier@enpc.fr`)
  - Raian Lefgoum (`rnl.persee@gmail.com`)
- Language de programmation: `Python` (`jupyter notebook`).

Organisation et règles de validation

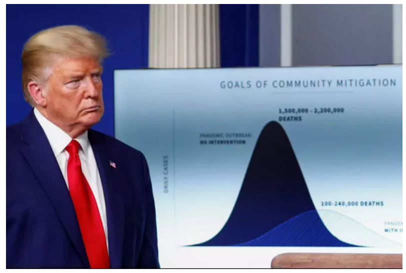
Organisation

- une semaine entre le cours et le TD correspondant ...
- ... pour: lire le cours, faire les exos, préparer le TP info
- une feuille d'exos par séance disponible sur le site

Évaluation

- contrôle sous forme de quizz: (portant sur cours, exercices, TD)
- envoi systématique à l'issu du TD (avant 12h00) du "Python notebook" à votre responsable de TD
- si plus d'une absence non justifiée examen oral final
- rendu (individuel) de l'un des 5 TD informatiques rédigé en détails (LaTeX suggéré).

Objet du cours



- Exposer des *situations* où la modélisation probabiliste est utile/indispensable pour prendre des décisions.
- Décrire un outil nouveau (les *chaînes de Markov*) et montrer comment cet outil peut être utilisé de façon effective.
- Implémenter les méthodes mathématiques proposées.

Aujourd'hui ...

- Des rappels de probabilité (espérance, variance, ...)
- Théorie du portefeuille de Markowitz.
  - Pourquoi des actifs (actions,...) ayant des rendements aux caractéristiques diverses peuvent ils coexister ?
  - Comment s'y prendre pour optimiser un portefeuille d'actif ?

Le rendement d'un actif

- Actif i (action, obligation, ...) de *rendement* R_i
- Rendement = sur une période de temps donnée T:

  1€ à l'instant 0 va rapporter $(1 + R_i)$€ en T.
- Plusieurs actifs (ex CAC40 $i = 1 \ldots 40$), vecteur de rendement $R = (R_1, \ldots, R_n)$.
- Rendements *déterministes* et un *marché* où l'on peut acheter et vendre ces actifs,

  tous les R_i doivent être égaux.
- Mais c'est un fait d'expérience que les rendements des actifs n'ont pas des caractéristiques identiques: il est *nécessaire* de les supposer *aléatoires*: "Risk is not an add-on"[1]...
- ce qui nous amène à quelques rappels de probabilité ...

[1] Voir R.C.Merton interviewed in [Bernstein(2007)].

Probabilité et espérance : rappels

- Description d'un expérience aléatoire : une *probabilité*, sur $(\Omega, \mathcal{A})$, $\mathbb{P}$ qui opère sur des ensembles de $\mathcal{A}$.
- $A \in \mathcal{A}$, $0 \leq \mathbb{P}(A) \leq 1$ et $\mathbb{P}(\Omega) = 1$. $\sum_{i \geq 1} A_i$ signifie une réunion *dénombrable* et *disjointe*.

$$\mathbb{P}\left(\sum_{i \geq 1} A_i\right) = \sum_{i \geq 1} \mathbb{P}(A_i)$$

- À une probabilité $\mathbb{P}$ est associée une espérance $\mathbb{E}$ (qui opère sur des variables aléatoires à valeurs dans $\mathbb{R}$)
  - $\mathbb{E}(1_A) = \mathbb{P}(A)$.
  - *linéarité* (X et Y positifs ou intégrables, X_i positifs).

$$\mathbb{E}(\lambda X + \mu Y) = \lambda \mathbb{E}(X) + \mu \mathbb{E}(Y)$$

$$\mathbb{E}\left(\sum_{i \geq 1} X_i\right) = \sum_{i \geq 1} \mathbb{E}(X_i)$$

Espérance et variance

- L'espérance est *linéaire*.
- Variance : $\mathrm{Var}(X) := \mathbb{E}\left\{(X - \mathbb{E}(X))^2\right\} = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.
- $\mathrm{Var}(X) = 0$ implique $X = Cte$ (p.s.).
- Covariance :
  $\mathrm{Cov}(X, Y) = \mathbb{E}\left\{(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right\} = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$.
- La linéarité de l'espérance permet de prouver

$$\mathrm{Var}\left(\sum_{i=1}^{n} \lambda_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \mathrm{Cov}(X_i, X_j)$$

- La variance fonctionne comme une forme *quadratique*, la covariance comme une *forme bilinéaire*.
- Cas particulier : si indépendance des X_i, $\mathrm{Cov}(X_i, X_j) = 0$, $i \neq j$

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Vecteur et matrice de variance-covariance

- $X = (X_1, \ldots, X_n)$ un vecteur de $\mathbb{R}^n$. Γ matrice de variance-covariance de X

$$\Gamma = (\mathrm{Cov}(X_i, X_j))_{1 \leq i,j \leq n}$$

- Permet de calculer $\mathrm{Var}\left(\sum_{i=1}^{n} \lambda_i X_i\right)$

$$\mathrm{Var}\left(\sum_{i=1}^{n} \lambda_i X_i\right) = \lambda^T \Gamma \lambda = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j \Gamma_{ij}.$$

- Γ est une matrice symétrique, positive ($\lambda^T \Gamma \lambda \geq 0$, pour tout λ).

Espérance et variance comme critères de décision

- Comment choisir entre 2 v.a. X et Y dont on ne connait que la loi ?
- On utilise (souvent) l'espérance et la variance comme critère de choix (plus l'espérance est grande ou plus la variance est petite, "meilleure" est la v.a.).
- C'est discutable[2], mais c'est l'approche la plus simple.
- Des "mesures de risque" mieux fondées (Voir [Föllmer and Schied(2008)]) existent et sont aussi utilisées : "value at risk", "expected shortfall", ...
- ... au prix toutefois d'une complexité accrue de la modélisation et/ou des calculs.

[2]On peut avoir $\mathbb{P}(X \leq Y) = 1$, sans que X et Y soient comparables pour l'ordre partiel suggéré (voir la feuille d'exercice pour un exemple).

Portefeuille de Markowitz : le modèle

- d actifs (actions, obligations, ...). R_i le rendement, aléatoires, du $i$ème actif.

$$1€ \text{ en actif } i \text{ en } 0 \rightarrow (1 + R_i)€ \text{ en } T.$$

- $R = (R_1, \dots, R_d)$ le vecteur, aléatoire, des rendements
- $r = (\mathbb{E}(R_1), \dots, \mathbb{E}(R_d))$, le vecteur des moyennes des rendements, par convention on suppose

$$r_1 \leq r_2 \leq \cdots \leq r_d.$$

- Γ la matrice de variance-covariance du vecteur R.

$$\sigma_i^2 = \mathrm{Var}(R_i) = \Gamma_{ii}.$$

Portefeuille de Markowitz : les portefeuilles

- On va considérer des portefeuilles, composés de quantités d'actifs λ_i, de valeur initiale V_0 égale (par convention) à 1€

$$V_0 = \sum_{i=1}^{d} \lambda_i = 1, \quad V_T = \sum_{i=1}^{d} \lambda_i (1 + \mathrm{R}_i) = 1 + \sum_{i=1}^{d} \lambda_i \mathrm{R}_i.$$

- Souvent $0 \leq \lambda_i \leq 1$. Mais dans certains cas on acceptera des quantités négatives pour λ_i (emprunt d'actif).
- Le gain du portefeuille est égal à $\mathrm{G} = V_T - V_0$

$$\mathrm{G} = \sum_{i=1}^{d} \lambda_i \mathrm{R}_i$$

- La moyenne et la variance de G se calcule facilement

$$\mathbb{E}(\mathrm{G}) = \lambda . r \qquad \mathrm{Var}(\mathrm{G}) = \lambda^T \Gamma \lambda$$

  en fonction de r et Γ. On va comparer les portefeuilles grâce à ces deux valeurs.

Portefeuille de Markowitz : les 2 critères

- L'hypothèse de base du modèle est que l'investisseur va classer les portefeuilles en utilisant la moyenne et l'écart-type (= la racine de la variance) du gain.
- Un portefeuille de moyenne inférieure et de variance supérieure ne sera jamais choisi par un investisseur rationnel (c'est réaliste, mais ça reste une hypothèse).
- On définit un ordre *partiel* sur les portefeuilles :

  G_1 préférable à G_2, ssi $\mathbb{E}(\mathrm{G}_1) \geq \mathbb{E}(\mathrm{G}_2)$ et $\mathrm{Var}(\mathrm{G}_1) \leq \mathrm{Var}(\mathrm{G}_2)$.
- On suppose que les actifs de base ne sont pas comparables pour cet ordre partiel, ce qui impose

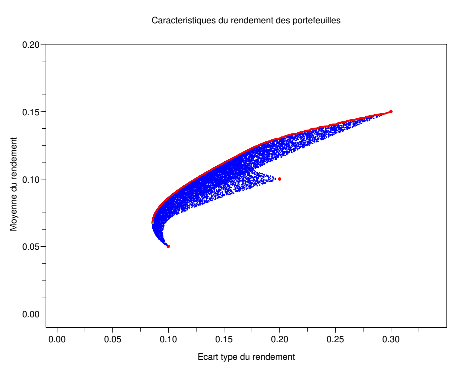
$$0 < \sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_d.$$

Portefeuille de Markowitz : frontière efficiente

- On s'intéresse aux portefeuilles "non dominés" (maximaux pour l'ordre partiel) : ceux sont les seuls susceptibles d'être utilisés par un intervenant rationnel.
- Ces portefeuilles "non dominés" forment une frontière de *Pareto* (ou d'indifférence).
- Markowitz l'appelle la *frontière efficiente*.
- On calculera cette frontière par exploration/simulation (en TD).
- Il existe d'autres méthodes plus efficaces (voir page 23) .

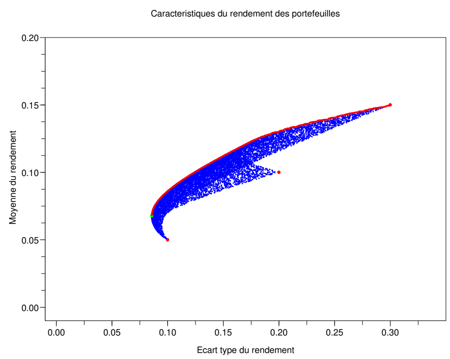
Exemple de frontière efficiente

- 3 actifs risqués, covariances nulles, frontière obtenue par simulation (cf TD), avec contraintes $0 \leq \lambda_i \leq 1$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$.
- frontière efficiente : points non dominés.



Frontière efficiente : remarques

- Le point de variance minimale.
- Effet de diversification : portefeuille de variance inférieure à l'actif de variance minimale, rendement meilleur.
- Sur la frontière efficiente on choisit un compromis espérance-variance.

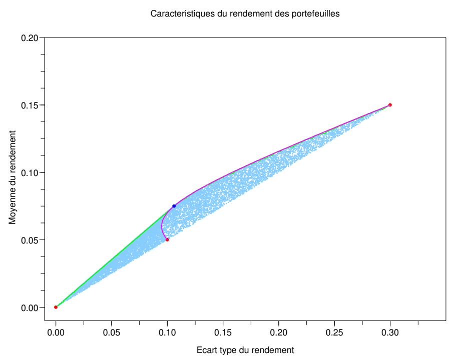


Actif sans risque et portefeuille de marché

- On suppose souvent l'existence d'un actif 0 sans risque (de variance nulle $\sigma_0 = 0$) et (donc) de rendement constant r_0.
- Cet actif correspond à un placement (quantité positive) ou un emprunt (quantité négative) dans une banque.
- Le rajout de cet actif modifie la forme de la frontière efficiente (cf TD) et fait apparaître un portefeuille remarquable: *le "portefeuille de marché"*.

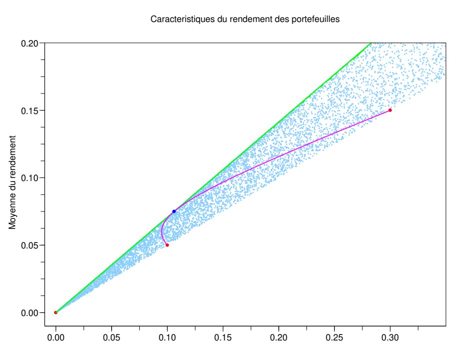
Le portefeuille de marché

- Cas 2 actifs risqués et un actif sans risque $r_0 = 0$, frontière obtenue par simulation (cf TD), avec contrainte $\lambda_i \geq 0$.
- Frontière efficiente pour les deux actifs risqués, les trois actifs.
- Portefeuille de marché : le point où les deux frontières se séparent.



Portefeuille de marché

- Son intérêt apparaît lorsque l'on autorise l'*emprunt* de l'actif non risqué.
- La frontière efficiente est modifiée (cf TD). La droite se prolonge au delà de P.
- Cette droite porte le nom de "droite de marché".



Intérêt du portefeuille de marché

- La droite de marché domine strictement la frontière efficiente "sans emprunt" (sauf pour le portefeuille de marché où elle coïncide).
- Le seul point "rationnel" de la frontière efficiente, qui ne contient pas d'actif sans risque, est le portefeuille de marché.
- Si l'on peut emprunter, les portefeuilles "rationnels" sont obtenus par combinaison de l'actif sans risque et du portefeuille de marché.
- En empruntant, on peut obtenir un rendement de même moyenne mais de variance inférieure à celle du meilleur actif individuel.
- En empruntant, on peut obtenir un rendement de même variance mais de moyenne supérieure à celle du meilleur actif individuel.
- Si l'on peut emprunter, il faut le faire !
- Le portefeuille de marché fait intervenir l'ensemble des actifs risqués (tous les actifs peuvent coexister).
- Variance nulle : il faut tout investir dans l'actif sans risque.

Conclusions

- finance et aléas
  - Sans hasard, pas d'activité financière.
  - ... ou dès qu'il y a du hasard, l'activité financière se développe.
  - cf développement des produits dérivées suite à l'abandon en 1971 de la convertibilité du dollars en or (accord de Bretton-Wood).
- Comment comparer des variables aléatoires ?
  - ici ordre partiel (Espérance,Variance)
  - Comme agréger des valeurs $X(\omega)$?

$$\min_{X\in\mathcal{X}} \mathbb{E}[f(X)] \;, \quad \min_{X\in\mathcal{X}} \sup_{\omega\in\Omega} f(X(\omega)) \;, \quad \min_{X\in\mathcal{X}} \sup_{P} \mathbb{E}_{P\in\mathcal{P}}[f(X)] \tag{1}$$

Le portefeuille de marché comme solution d'un problème d'optimisation

- Pour calculer P il faut résoudre le problème d'optimisation suivant : max en λ (vérifiant les contraintes) de la pente du point $(\sqrt{\mathrm{Var}(\mathsf{G})}, \mathbb{E}(\mathsf{G}))$ ou encore

$$\max_{\lambda} \frac{\mathbb{E}(\mathsf{G})^2}{\mathrm{Var}(\mathsf{G})} = \max_{\lambda} \frac{(r^T\lambda)^2}{\lambda^T\Gamma\lambda}$$

- Le quotient $\frac{\mathbb{E}(\mathsf{G})}{\sqrt{\mathrm{Var}(\mathsf{G})}}$ s'appelle le *ratio de Sharpe* du portefeuille[3].

[3]William Sharpe a été lui aussi prix Nobel d'économie en 1990 avec Harry Markowitz et Merton Miller.

Calcul du portefeuille de marché

$$\max_{\lambda} \frac{\mathbb{E}(\mathsf{G})^2}{\mathrm{Var}(\mathsf{G})} = \max_{\lambda} \frac{(r^T\lambda)^2}{\lambda^T\Gamma\lambda}$$

- Voir cours optimisation de 1A.
- Lorsque l'on ne tient pas compte des contraintes de positivité sur λ, mais seulement de $\sum_{i=1}^{n} \lambda_i = 1^T\lambda = 1$, on peut résoudre ce problème.
- On peut utiliser un algorithme générique d'optimisation de type gradient. Ce type de problème est accessible dans `Scicoslab` ou `Python`, via une fonction d'optimisation. Voir le corrigé du TD.
- Le problème évoqué peut être traité explicitement, en utilisant l'inégalité de Cauchy-Schwartz.

Calcul du portefeuille de marché

- En utilisant l'inégalité de Cauchy-Schwartz pour la norme $\sqrt{x^T \Gamma x}$, on obtient

$$r^T \lambda = (\Gamma^{-1} r)^T \Gamma \lambda \leq \sqrt{r^T \Gamma^{-1} r}\ \sqrt{\lambda^T \Gamma \lambda}.$$

  On en déduit que

$$\sup_{\lambda, \sum_{i=1}^d \lambda_i = 1} \frac{r^T \lambda}{\sqrt{\lambda^T \Gamma \lambda}} \leq \sqrt{r^T \Gamma^{-1} r}$$

  et que l'égalité est atteinte pour $\lambda = \Gamma^{-1} r / (1\ \Gamma^{-1} r)$ où $1 = (1, \ldots, 1)$

Paramètrisation de la frontière efficiente par des problèmes d'optimisation

- Une façon de calculer la frontière efficiente est de minimiser la variance à espérance égale à α, puis de faire varier α.

$$\min_{\lambda,\ 1^T \lambda = 1,\ r^T \lambda = \alpha} \lambda^T \Gamma \lambda.$$

- Toujours en ignorant les contraintes de positivité sur λ, on peut encore résoudre ce problème
- Soit en utilisant une routine d'optimisation en éliminant les 2 contraintes (voir TD).
- Soit une utilisant deux multiplicateurs de Lagrange, le problème se résolvant alors de façon explicite en utilisant l'inverse de la matrice Γ (exercice).

Conclusions : décider dans l'incertain

- Décider en environnement incertain suppose d'*identifier* un modèle et de *choisir des critères* de décision (ce qui n'est pas toujours simple).
- Un fois ceci fait, on utilise des algorithmes pour résoudre les problèmes d'*optimisation* (ce qui n'est pas toujours simple).
- Une modélisation aléatoire simple peut éclairer la pratique et conduire à des règles opératoires.

Harry Markowitz

- Né en 1927, prix Nobel d'Economie 1990.
- Article fondateur, à 25 ans, en 1952 ("Portfolio Selection", The Journal of Finance).

- " ... when I defended my dissertation as a student in the Economics Department of the University of Chicago, Professor Milton Friedman argued that portfolio theory was not Economics, and that they could not award me a Ph.D. degree in Economics for a dissertation which was not in Economics. ... at the time I defended my dissertation, portfolio theory was not part of Economics. But now it is." Nobel Lecture, December 7, 1990, Harry M. Markowitz.

Pour aller plus loin : cours de l'École des Ponts

- 1A : cours de "Probabilité", cours d'"Optimisation"
- 2A : cours "Processus Aléatoires", cours de "Recherche Opérationnelle", cours "Optimisation", cours "Méthodes Mathématiques pour la Finance", cours "Stratégie financière de l'entreprise"

Bibliographie

P.L. Bernstein.
Capital Ideas Evolving.
Jhon Wiley and Son, 2007.

Hans Föllmer and Alexander Schied.
Convex and coherent risk measures.
http://www.alexschied.de/Encyclopedia6.pdf, 2008.

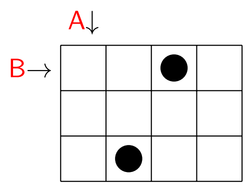
M. Haugh.
Asset allocation and risk management.
Available on ieor e4602 web page, Columbia University, 2009.

H.M. Markowitz.
Portfolio selection.
The Journal of Finance, 7(1):77–91, 1952.

H.M. Markowitz.
Portfolio Selection: Efficient Diversification of Investments.
John Wiley and Sons, 1959.

TP: Sur un puzzle (du site) de la NSA

- Deux œufs "au hasard" dans une matrice $p \times q$.
- Deux joueurs: le joueur A parcours la matrice en colonne, le joueur B en ligne.
- Le gagnant est celui qui atteint le premier l'un des deux œufs.



- Ce jeu est-il équitable ?
- La réponse dépend de p et q de façon bizarre:

  $\neq$ pour $(p = 3, q = 4)$, $(p = 4, q = 4)$ et $(p = 4, q = 5)$!

Exercice 1 1. Montrer que, si X est une variable réelle, $\mathbb{E}\left\{(X-\mathbb{E}(X))^2\right\} = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.

2. Montrer que si $X = (X_1, \dots, X_n)$ est un vecteur de $\mathbb{R}^n$ on a

$$\text{Var}\left(\sum_{i=1}^{n} \lambda_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j \text{Cov}(X_i, X_j)$$

   En déduire que la matrice symétrique $\Gamma = (\text{Cov}(X_i, X_j))_{1\leq i,j\leq n}$ est une matrice positive (i.e. $x.\Gamma x \geq 0$ pour tout $x \in \mathbb{R}^n$).

3. Montrer que la matrice de covariance est dégénérée (i.e. elle admet un noyau non réduit à $\{0\}$) si et seulement si le vecteur X prend ses valeurs dans un hyperplan affine strict de $\mathbb{R}^n$.

4. Soit ρ un nombre réel compris entre -1 et 1, à quelle condition sur ρ la matrice Γ, $n \times n$ suivante

$$\Gamma = \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \rho & \rho & \dots & \rho & 1 \end{pmatrix}$$

   peut elle être la matrice de covariance d'un vecteur aléatoire ?

5. On suppose $\rho \geq 0$ et on considère $(G_1, \dots, G_n, G)$, $n+1$ variables aléatoires indépendantes centrées de variance 1. Comment construire un vecteur aléatoire $(X_1, \dots, X_n)$ qui a pour matrice de variance covariance Γ ?

Exercice 2 Soit un couple de variables aléatoires (X, Y) tel que $\mathbb{P}(X \leq Y) = 1$ [1].

1. Si X et Y representent des rendements de portefeuille, laquelle de ces deux variables aléatoires vous paraît naturellement préférable ?
2. Donner un exemple de couple de variables aléatoires (X, Y) tel que $X \leq Y$ p.s., avec (bien sûr !) $\mathbb{E}(X) \leq \mathbb{E}(Y)$, mais telles que Var $(X) <$ Var (Y).

   Ces deux variables aléatoires *ne seront pas comparables pour l'ordre partiel sur les gains de portefeuille* défini dans le cours [2].
3. En quoi cela questionne t'il le choix de l'ordre partiel introduit dans la modèlisation du portefeuille proposée par Markowitz ?

Exercice 3 Soit Γ une matrice symétrique définie positive. Montrer que $\phi(x, y) = x^T\Gamma y$ est un produit scalaire. On note $\|x\|_\Gamma = \sqrt{x^T\Gamma x}$ la norme associée.

1. (Re)-démontrer l'inégalité de Cauchy-Schwartz :

$$|x^T\Gamma y| \leq \|x\|_\Gamma \|y\|_\Gamma$$

2. En déduire que :

$$r^T\lambda \leq \sqrt{r^T\Gamma^{-1}r} \times \sqrt{\lambda^T\Gamma\lambda}.$$

1. Remarquez que pour savoir si $\mathbb{P}(X \leq Y) = 1$, il faut connaître la loi du couple (X, Y).

2. Notez que cet ordre partiel ne suppose la connaissance que de la loi de X et de celle de Y (et non celle du couple (X, Y) comme précédemment).

Puis que

$$\sup_{\lambda, \sum_{i=1}^{d} \lambda_i = 1} \frac{r^T \lambda}{\sqrt{\lambda^T \Gamma \lambda}} \leq \sqrt{r^T \Gamma^{-1} r}$$

et que l'égalité est atteinte pour $\lambda = \Gamma^{-1} r / (\mathbf{1}\, \Gamma^{-1} r)$ où $\mathbf{1} = (1, \ldots, 1)$

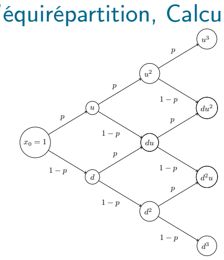
Chapitre 2

Loi et chaîne de Markov

Test d'équirépartition, Calcul de prix

Loi et chaîne de Markov

Test d'équirépartition, Calcul de prix



J.-P. Chancelier, B. Lapeyre

http://cermics.enpc.fr/~jpc/decision-incertain

Plan

Probabilité et espérance conditionnelle

- Une *probabilité* $\mathbb{P}$ sur $(\Omega, \mathcal{A})$.
- Une espérance $\mathbb{E}$ (qui opére *linéairement* sur des variables aléatoires à valeurs dans $\mathbb{R}$).
- Probabilité conditionnelle : A, $B \in \mathcal{A}$, $\mathbb{P}(B) > 0$ par définition

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

- Espérance conditionnelle : $B \in \mathcal{A}$, $\mathbb{P}(B) > 0$, X v.a., par définition

$$\mathbb{E}(X|B) = \frac{\mathbb{E}(X 1_B)}{\mathbb{P}(B)}$$

$\rightarrow$ On obtient $\mathbb{E}(1_A|B) = \mathbb{P}(A|B)$ et la linéarité de $X \rightarrow \mathbb{E}(X|B)$.

Loi d'une variable aléatoire

- Un espace d'état fini $\mathbb{X} = \{x_1, x_2, \ldots, x_n\}$,
  une variable aléatoire X à valeur dans $\mathbb{X}$,

$$\mathrm{X} : \Omega \rightarrow \mathbb{X}$$

- Loi de la variable aléatoire X, $p_\mathrm{X} : \mathbb{X} \rightarrow [0, 1]$

$$p_\mathrm{X}(x) = \mathbb{P}\big[\{X = x\}\big] \ , \ \ \forall x \in \mathbb{X}$$

  - On a $p_\mathrm{X}(x) \in [0, 1]$ et $\sum_{x \in \mathbb{X}} p_\mathrm{X}(x) = 1$.
- A quoi ca sert ?

$$\mathbb{E}\big[f(\mathrm{X})\big] = \sum_{x \in \mathbb{X}} p(x) f(x)$$

Loi d'un vecteur aléatoire

- $X = (X_1, \dots, X_n)$ un vecteur aléatoire à valeurs dans $\mathbb{X}^n$. Loi de X

$$p_X(x_1, \dots, x_n) = \mathbb{P}\left(X_1 = x_1, \dots, X_n = x_n\right), x_1, \dots, x_n \in \mathbb{X}$$

- $\sum_{x_1,\dots,x_n \in \mathbb{X}} p_X(x_1, \dots, x_n) = 1.$
- A quoi ca sert ?

$$\mathbb{E}\left[f(X_1, \dots, X_n)\right] = \sum_{x_1,\dots,x_n \in \mathbb{X}} p_X(x_1, \dots, x_n) f(x_1, \dots, x_n)$$

- Par contraction, la loi de X permet de calculer les lois marginales p_{X_i} des X_i

$$p_{X_i}(x) = \sum_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_n \in \mathbb{X}} p_X(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_n)$$

- Une remarque simple (mais utile) :

$$\mathbb{E}\left[f(X_i)\right] = \sum_{x_1,\dots,x_n \in \mathbb{X}} p_X(x_1, \dots, x_n) f(x_i) = \mathbb{E}_{p_{X_i}}\left[f(X_i)\right] = \sum_{x_i \in E} p_{X_i}(x_i) f(x_i)\ .$$

Loi d'un vecteur et indépendance

- $(X_1, \dots, X_n)$ est un vecteur de v.a. indépendantes si et seulement si (au choix), pour tout $x_1, \dots, x_n \in \mathbb{X}$, pour tout $f_1, \dots, f_n$ de $\mathbb{X}$ dans $\mathbb{R}$:
  - $\mathbb{P}\left(X_1 = x_1, \dots, X_n = x_n\right) = \mathbb{P}(X_1 = x_1) \times \dots \times \mathbb{P}(X_n = x_n)$
  - $\mathbb{E}\left[f_1(X_1) \times \cdots \times f_n(X_n)\right] = \mathbb{E}\left[f_1(X_1)\right] \times \cdots \times \mathbb{E}\left[(f_n(X_n)\right]$
- les lois marginales p^i des X_i ne caractérisent pas la loi du vecteur $(X_1, \dots, X_n)$
- ... sauf si on suppose que ces v.a. sont indépendantes

$$p_{(X_1,\cdots,X_n)} = p_{X_1} \otimes p_{X_2} \otimes \cdots p_{X_n}\ .$$

Système dynamique aléatoire

- $\mathrm{W} = (\mathrm{W}_n, n \geq 1)$ suite de variables aléatoires indépendantes de même loi sur un espace $\mathbb{W}$.
- On considère $F : \mathbb{X} \times \mathbb{W} \to \mathbb{X}$ et "On tire une application de $\mathbb{X}$ dans $\mathbb{X}$, au hasard" : $F(\cdot, \mathrm{W}_{n+1})$.
- Sytème dynamique aléatoire, $\mathrm{X}_0 = x_0 \in \mathbb{X}$ arbitraire, et par récurrence
$$\mathrm{X}_{n+1} = F(\mathrm{X}_n, \mathrm{W}_{n+1})$$
- La connaissance de $F : \mathbb{X} \times \mathbb{W} \to \mathbb{X}$ permet de simuler très simplement la suite $(\mathrm{X}_n, n \geq 0)$ à partir de la suite $(\mathrm{W}_n, n \geq 1)$.

Systèmes dynamiques aléatoires : exemples

$(W_n, n \geq 1)$ une suite de tirages à pile ou face indépendants, $0 \leq p \leq 1$
$$W_n = \begin{cases} P & \text{avec probabilité } p\ , \\ F & \text{avec probabilité } (1-p)\ . \end{cases}$$

- Marche aléatoire ($p = 1/2$, marche aléatoire symétrique)
$$X_0 = 0\ ,\ \ X_{n+1} = X_n + \left(1_{\{W_{n+1} = P\}} - 1_{\{W_{n+1} = F\}}\right).$$
- Processus en "dents de scie": (# pile consécutifs avant n)
$$X_0 = 0\ ,\ \ X_{n+1} = (1 + X_n)1_{\{W_{n+1} = P\}}.$$
- Processus de Cox-Ross, $(d, u) \in \mathbb{R}^2$ avec $d < 1 < u$
$$X_0 = 1\ ,\ \ X_{n+1} = X_n \left(u\, 1_{\{W_{n+1} = P\}} + d\, 1_{\{W_{n+1} = F\}}\right).$$

Système dynamique aléatoire et propriété de Markov

- Un système dynamique aléatoire vérifie la propriété de Markov

$$\mathbb{P}\left(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\right) = P(x_n, x_{n+1})$$

- où P une matrice de transition : $(P(x,y), x, y \in E)$ (parfois aussi noté $(P(x \to y))$ donnée par

$$P(x,y) = \mathbb{P}(F(x, W_1) = y).$$

- Matrice de transition : $P(x,y) \geq 0$ et $\sum_{y \in E} P(x,y) = 1$.
- On a forcement $P(x,y) = \mathbb{P}\left(X_{n+1} = y \,|\, X_n = x\right)$ et $\mathbb{P}\left(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\right) = \mathbb{P}\left(X_{n+1} = x_{n+1} | X_n = x_n\right)$

Preuve

$$\begin{aligned}
&\mathbb{P}\left(X_{n+1} = x_{n+1}, X_n = x_n, \ldots, X_0 = x_0\right) \\
&\quad = \mathbb{P}\left(F(X_n, W_{n+1}) = x_{n+1}, X_n = x_n, \ldots, X_0 = x_0\right) \\
&\quad = \mathbb{P}\left(F(x_n, W_{n+1}) = x_{n+1}, X_n = x_n, \ldots, X_0 = x_0\right).
\end{aligned}$$

- $F(x_n, W_{n+1})$ indépendante du vecteur $(W_1, \ldots, W_n)$
- $X_0, \ldots, X_n$ fonctions de $(W_1, \ldots, W_n)$.
- Donc $F(x_n, W_{n+1})$ indépendante de $X_0, \ldots, X_n$.

$$\begin{aligned}
&\mathbb{P}\left(X_{n+1} = x_{n+1}, X_n = x_n, \ldots, X_0 = x_0\right) \\
&\quad = \mathbb{P}\left(F(x_n, W_{n+1}) = x_{n+1}\right) \mathbb{P}\left(X_n = x_n, \ldots, X_0 = x_0\right) \\
&\quad = \underbrace{\mathbb{P}\left(F(x_n, W_1) = x_{n+1}\right)}_{=} \mathbb{P}\left(X_n = x_n, \ldots, X_0 = x_0\right) \\
&\quad = \qquad P(x_n, x_{n+1}) \quad \mathbb{P}\left(X_n = x_n, \ldots, X_0 = x_0\right)
\end{aligned}$$

- En sommant sur tous les $x_0, \ldots, x_{n-1}$ dans $\mathbb{X}$, on obtient aussi

$$\mathbb{P}\left(X_{n+1} = x_{n+1}, X_n = x_n\right) = P(x_n, x_{n+1}) \mathbb{P}\left(X_n = x_n\right).$$

Matrice de transition : exemples

- Marche aléatoire ($p = 1/2$, marche aléatoire symétrique)

$$X_0 = 0, X_{n+1} = X_n + \left(1_{\{W_{n+1} = P\}} - 1_{\{W_{n+1} = F\}}\right).$$

$x \in \mathbb{Z}$, $P(x, x+1) = p$, $P(x, x-1) = 1 - p$, 0 sinon

- Processus en "dents de scie": (# pile consécutifs avant n)

$$X_0 = 0, X_{n+1} = (1 + X_n)1_{\{W_{n+1} = P\}}.$$

$x \in \mathbb{N}$, $P(x, x+1) = p$, $P(x, 0) = 1 - p$, 0 sinon

- Processus de Cox-Ross (cf. finance), $d < 1 < u$

$$X_0 = 1, X_{n+1} = X_n \left(u\, 1_{\{W_{n+1} = P\}} + d\, 1_{\{W_{n+1} = F\}}\right).$$

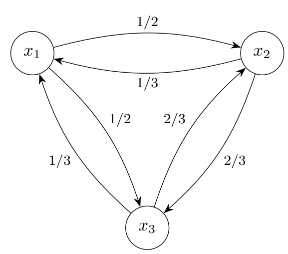
$P(x, xu) = p$, $P(x, xd) = 1 - p$, 0 sinon

Matrice de transition : exemples (suite)

- Matrice de transition P sur $\mathbb{X} = \{x_1, x_2, x_3\}$

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 0 & 2/3 \\ 1/3 & 2/3 & 0 \end{pmatrix}$$

- Graphe $G = (\mathbb{X}, \mathcal{E})$ avec $x\mathcal{E}y \iff P_{x,y} > 0$



- Que represente la matrice $P^2 = P \times P$?

Système dynamique aléatoire et chaîne de Markov

- une suite de variables aléatoires qui vérifie la propriété de Markov, s'appelle une chaîne de Markov.
- $P(x,y) = \mathbb{P}\left(X_{n+1} = y | X_n = x\right)$, $x, y \in E$, est la matrice de transition de la chaîne de Markov.
- Un système dynamique aléatoire est une chaîne de Markov et reciproquement on peut représenter (en loi) une chaîne de Markov comme un système dynamique aléatoire.
- La matrice de transition $P(x,y)$ se déduit de F et de la loi de W.
- P contient toute l'information utile pour mener à bien les calculs de loi concernant la chaîne.
- Pour simuler une chaîne de Markov, on utilise une représentation comme système dynamique aléatoire.

Chaîne de Markov : définition

Definition 1

Une suite de v.a. $(X_n, n \geq 0)$ à valeurs dans $\mathbb{X}$ est un chaîne de Markov de *matrice de transition* $(P(x,y), x \in \mathbb{X}, y \in \mathbb{X})$, si et seulement si, par définition, pour tout $x_0, x_1, \ldots, x_n, x_{n+1} \in \mathbb{X}$,

$$\begin{aligned}\mathbb{P}\left(X_{n+1} = x_{n+1} | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\right) \\ &= P(x_n, x_{n+1}) \\ &= \mathbb{P}\left(X_{n+1} = x_{n+1} | X_n = x_n\right)\end{aligned}$$

La loi de X_0, μ_0, est appelée la *loi initiale*.

- **Formellement, égalité vrai seulement si $\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n) > 0$, sinon l'interpréter comme**

$$\begin{aligned}\mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n, X_{n+1} = x_{n+1}) \\ = \mathbb{P}(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n) P(x_n, x_{n+1})\end{aligned}$$

- **Matrice de transition P_n dépendant de n : chaîne de Markov inhomogène**

Matrice de transition et calcul de lois

- $\mathbb{X}$ fini
- $(X_n, n \geq 0)$ une chaîne de Markov de matrice de transition P sur $\mathbb{X}$.
- μ_0 la loi de X_0 (loi initiale), $\mu_0(x) = \mathbb{P}(X_0 = x)$, $x \in \mathbb{X}$.
- La loi du vecteur $(X_0, \ldots, X_n)$ se calcule explicitement en fonction de P et μ_0

$$\mathbb{P}\left(X_0 = x_0, \ldots, X_n = x_n\right) = \mu_0(x_0)P(x_0, x_1) \times \cdots \times P(x_{n-1}, x_n).$$

- Preuve = propriété de Markov

$$\begin{aligned}\mathbb{P}\left(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = x_n\right) \\ = \mathbb{P}\left(X_0 = x_0, \ldots, X_{n-1} = x_{n-1}\right) P(x_{n-1}, x_n),\end{aligned}$$

puis on itère.

Matrice de transition : quelques notations

- $\mathbb{X}$ fini : μ est un vecteur *ligne*, P une *matrice* et f un vecteur *colonne*.
- $(Pf)(x) = \sum_{y \in \mathbb{X}} P(x, y)f(y)$, f une fonction de $\mathbb{X}$ dans $\mathbb{R}$. Pf est aussi une fonction de $\mathbb{X}$ dans $\mathbb{R}$.
- $(\mu P)(y) = \sum_{x \in \mathbb{X}} \mu(x)P(x, y)$, μ une loi de probabilité sur $\mathbb{X}$. μP est aussi une probabilité sur $\mathbb{X}$
- $(PQ)(x, y) = \sum_{z \in \mathbb{X}} P(x, z)Q(z, y)$, pour P et Q deux matrices de transition. PQ est aussi une matrice de transition.
- P^n est définie par récurrence par :
$$P^{n+1}(x, y) = \sum_{z \in \mathbb{X}} P(x, z)P^n(z, y).$$
- μ loi sur $\mathbb{X}$, f fonction de $\mathbb{X}$ dans $\mathbb{R}$: $\mu f = \sum_{x \in \mathbb{X}} \mu(x)f(x)$.
- Si X a pour loi μ, $\mathbb{E}[f(X)] = \mu f$.

Loi de X_n

- Par contraction de la loi du vecteur $(X_0, \ldots, X_n)$, on obtient

$$\mathbb{P}(X_n = x_n) = \sum_{x_0, x_1, \ldots, x_{n-1} \in \mathbb{X}} \mu_0(x_0) P(x_0, x_1) \times \cdots \times P(x_{n-1}, x_n).$$

- avec les notations précédentes :

$$\mathbb{P}(X_n = x_n) = (\mu_0 P^n)(x_n),$$

- $\text{loi}(X_n) = \mu_0 P^n$: un (simple) calcul matriciel à partir de μ_0 et P.

Un test d'équirépartition

- Question Les 100 tirages à pile ou face suivants sont ils "aléatoires" ?

P	F	F	F	P	F	P	P	F	F
F	P	P	P	F	F	P	F	P	F
P	F	F	F	P	P	F	F	P	P
P	F	F	P	F	P	P	F	F	P
F	P	P	F	P	P	F	F	F	P
P	P	F	P	F	P	P	F	P	P
F	F	P	F	P	P	F	F	P	F
P	F	F	F	P	F	P	P	F	F
P	F	F	F	P	P	P	F	P	F
P	F	F	P	P	F	P	F	P	P

- Ça n'a pas l'air trop mal : on trouve 50 piles et 50 faces ...

Un test d'équirépartition

- mais ... le nombre maximum de piles consécutifs est de 3, ce qui est (très) peu compatible avec l'hypothèse d'un tirage au hasard.
- On a du mal a imiter le hasard, ...
- Nous allons voir qu'il faut s'attendre à 4 piles consécutifs (avec proba 0.97) voire 5 (avec proba 0.81).
- On a plus de chance de voir 10 piles consécutifs (ou plus) que d'en voir 3 ou moins !
- On va calculer le loi du nombre maximum de piles consécutifs au bout de 100 tirages.

Chaîne de Markov arrêtée

- $(X_n, n \geq 0)$ un système dynamique aléatoire
$$X_{n+1} = F(X_n, W_{n+1}).$$
- $x \in \mathbb{X}$, soit $\tau_x = \inf \left\{ n \geq 0 \,\middle|\, X_n = x \right\}$: le premier temps d'atteinte de x (pas forcément fini, $+\infty$ si ensemble vide). C'est une variable aléatoire.
- $n \wedge \tau_x = \inf \{n, \tau\}$
- La suite $(Y_n, n \geq 0)$ définie par $Y_n = X_{n \wedge \tau}$ pour tout, $n \geq 0$ est aussi un système dynamique aléatoire puisque
$$Y_{n+1} = F(Y_n, W_{n+1}) 1_{\{Y_n \neq x\}} + x 1_{\{Y_n = x\}} = G(Y_n, W_{n+1}).$$
$$\forall (x', w) \in \mathbb{X} \times \mathbb{W}\,, \;\; G(x', w) = F(x', w) 1_{\{x' \neq x\}} + x 1_{\{x' = x\}} \,..$$
$(Y_n)_{n \in \mathbb{N}}$ est donc une chaîne de Markov
- On a $\{Y_n = x\} = \{n \geq \tau\}$.

Chaîne en dents de scie arrêtée

- $(X_n, n \geq 0)$ la chaîne en dents de scie
- On fixe $l > 0$ et on considère $\tau_l = \inf\{n \geq 0, X_n = l\}$
- $(Y_n^l, n \geq 0)$ avec $Y_n^l = X_{n \wedge \tau_l}$ pour tout $n \in \mathbb{N}$
  la chaîne de Markov en dents de scie arrêtée
- $Y_n^l \in \{0, \dots, l\}$ pour tout $n \in \mathbb{N}$.
- de matrice de transition P_l

$$P_l : \begin{cases} P_l(x, x+1) = p & \text{si } 0 \leq x < l, \\ P_l(x, 0) = 1 - p & \text{si } 0 \leq x < l, \\ P_l(l, l) = 1. & \\ P_l(x, y) = 0 & \text{sinon} \end{cases}$$

Loi du nombre de piles consécutifs après 100 tirages

- $(N_n^{\text{max}}, n \geq 0)$ nombre maximum de piles consécutifs après n tirages.

$$\{Y_n^l = l\} = \{N_n^{\text{max}} \geq l\}$$

- Le calcul de la loi de Y_n^l pour tout n donne $\mathbb{P}(Y_n^l = l)$
  - $\mathbb{P}(Y_0^l = 0) = 1$. Loi initiale $\mu_0(0) = 1$, $\mu_0(k) = 0$ sinon.
    $\mu_0 = (1, 0, \dots, 0)$.
  - Loi de $Y_n^l = \mu_0(P_l)^n = (P_l^n(0, k), 0 \leq k \leq l)$ (un vecteur ligne).
  - $\mathbb{P}(Y_n^l = l) = (P_l)^n(0, l)$.
- On obtient la *loi* du nombre de piles consécutifs maximum après 100 tirages avec N_{100}^{max}.

$$\begin{aligned} \mathbb{P}(N_{100}^{\text{max}} = l) &= \mathbb{P}(Y_{100}^l = l) - \mathbb{P}(Y_{100}^{l+1} = l + 1) \\ &= P_l^{100}(0, l) - P_{l+1}^{100}(0, l + 1). \end{aligned}$$

Loi du nombre maximum de piles consécutifs

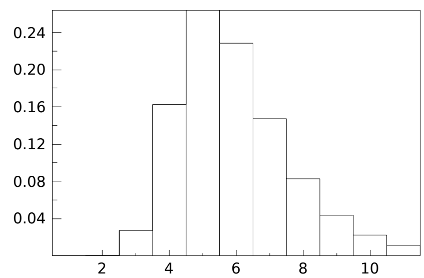


Figure: Loi du nombre de piles consécutifs.

- $\mathbb{P}$(nombre de piles consécutifs $\leq 3) = 0.0273$
- $\mathbb{P}$(nombre de piles consécutifs $\geq 5) = 0.8101$

Une régle de décision

- Si l'on voit au moins 4 piles consécutifs, la suite sera déclarée aléatoire, sinon, la suite sera déclarée non aléatoire.
- Si la suite a été tirée aléatoirement, je vais me tromper avec moins de 3% d'erreur.
- Si la suite n'est pas aléatoire, je suis aidé par la tendance naturelle à ne pas faire de trop longues series.
- "Mon" exemple ne passe pas le test. Mais si on le lit par colonne ou lieu de ligne, il le passe aisément!
- À vérifier avec l'un de vos camarades, si vous êtes sceptique.

Loi de X_n : une autre façon d'écrire les choses

- $(X_n, n \geq 0)$ chaîne de Markov de matrice de transition P sur $\mathbb{X}$.
- $\mathbb{P}(X_0 = x_0) = 1 \quad$ (i.e. $\mu_0(x_0) = 1$, $\mu_0(x) = 0$ si $x \neq x_0$.)
- On sait exprimer $\mathbb{E}\left[f(X_N)\right]$

$$\mathbb{E}\left[f(X_N)\right] = \mu_N f = \mu_0 P^N f = \sum_{x \in \mathbb{X}} \mu_0(x)(P^N f)(x) = (P^N f)(x_0).$$

Une formulation alternative plus algorithmique.

Theorem 2

Soit $(u(n,x), n = 0, \ldots, N, x \in \mathbb{X})$ *la solution unique de*

$$(1) \qquad \begin{cases} u(n,x) = \sum_{y \in \mathbb{X}} P(x,y) u(n+1,y), & n < N \\ u(N,x) = f(x), \end{cases}$$

Alors $\mathbb{E}[f(X_N)] = u(0, x_0)$.

Remarques

- La première équation de (1) peut aussi s'écrire

$$u(n,x) = P[u(n+1,.)](x) = \sum_{y \in \mathbb{X}} P(x,y) u(n+1,y)$$

- $u(n,x)$ peut s'interpréter comme

$$u(n,x) = \overbrace{P \times \cdots \times P}^{N-n \text{ fois}} f(x) = P^{N-n} f(x),$$

- Lorsque $\mathbb{P}(X_n = x) > 0$, on a

$$u(n,x) = \mathbb{E}\left[f(X_N) \,|\, X_n = x\right] = P^{N-n} f(x).$$

Preuve

$$\begin{aligned}&\mathbb{P}(X_0 = x_0, \dots, X_n = x_n, X_{n+1} = x_{n+1}, \dots, X_N = x_N)\\ &\quad = \mathbb{P}(X_0 = x_0)P(x_0, x_1) \times \cdots \times P(x_n, x_{n+1}) \times P(x_{n+1}, x_{n+2}) \times \cdots \times P(x_{N-1}, x_N)\end{aligned}$$

en sommant sur toutes les valeurs de $x_0, \cdots, x_{n-1}$ et $x_{n+1}, \cdots, x_{N-1}$ on obtient la loi de (X_n, X_N)

$$\mathbb{P}(X_n = x_n, X_N = x_N) = \mathbb{P}(X_n = x_n)P^{N-n}(x_n, x_N)$$

$$\begin{aligned}\mathbb{E}\left(f(X_N)1_{\{X_n = x_n\}}\right) &= \mathbb{P}(X_n = x_n)\sum_{x_N \in \mathbb{X}} P^{N-n}(x_n, x_N)f(x_N)\\ &= \mathbb{P}(X_n = x_n)(P^{N-n}f)(x_n)\end{aligned}$$

Remarques et notations

- (1) est *une équation de programmation dynamique*
- $\mathbb{X}$ est fini, (1) est un *algorithme* qui termine.

Une notation commode

- $x_{0:n}$ plutôt que $(x_0, \dots, x_n)$
- $X_{0:n}$ plutôt que $(X_0, \dots, X_n)$
- $X_{0:n} = x_{0:n}$ plutôt que $(X_0 = x_0, \dots, X_n = x_n)$

Preuve formelle

On le sait déjà ... mais voici une autre méthode de preuve. On va montrer que

$$\mathbb{E}\left[u(n+1, X_{n+1})\right] = \mathbb{E}\left[u(n, X_n)\right].$$

Si cela est vrai :

$$u(0, x_0) = \mathbb{E}\left[u(0, X_0)\right] = \cdots = \mathbb{E}\left[u(N, X_N)\right] = \mathbb{E}\left[f(X_N)\right].$$

La loi de $X_{0:n+1} = (X_{0:n}, X_{n+1})$ est donnée par (c'est un façon d'exprimer la propriété de Markov)

$$\mathbb{P}(X_{0:n+1} = x_{0:n+1}) = \mathbb{P}(X_{0:n} = x_{0:n}, X_{n+1} = x_{n+1}) = \mathbb{P}(X_{0:n} = x_{0:n}) P(x_n, x_{n+1}).$$

$$\begin{aligned}
\mathbb{E}\left[u(n+1, X_{n+1})\right] &= \sum_{x_{0:n+1} \in \mathbb{X}^{n+2}} u(n+1, x_{n+1}) \mathbb{P}(X_{0:n+1} = x_{0:n+1}), \\
\text{[propriété de Markov]} &= \sum_{x_{0:n} \in \mathbb{X}^{n+1}, x_{n+1} \in \mathbb{X}} u(n+1, x_{n+1}) \mathbb{P}(X_{0:n} = x_{0:n}) P(x_n, x_{n+1}), \\
&= \sum_{x_{0:n} \in \mathbb{X}^{n+1}} \mathbb{P}(X_{0:n} = x_{0:n}) \sum_{x_{n+1} \in \mathbb{X}} P(x_n, x_{n+1}) u(n+1, x_{n+1}), \\
\text{[définition de } u(n,.)\text{]} &= \sum_{x_{0:n} \in \mathbb{X}^{n+1}} \mathbb{P}(X_{0:n} = x_{0:n}) u(n, x_n) = \mathbb{E}\left[u(n, X_n)\right].
\end{aligned}$$

Calcul du prix d'une option

- Un modèle : le processus de Cox-Ross

$$X_0 = x_0, X_{n+1} = X_n \left(u\, 1_{\{W_{n+1} = P\}} + d\, 1_{\{W_{n+1} = F\}} \right).$$

- Un profil d'option $f(X_N)$ (ce que je vais recevoir en N)

$$f(x) = (x - K)_+, \text{ option d'achat de prix d'exercice } K.$$

- Prix : $\mathbb{E}\left[f(X_N)\right]$ ("intuitif", voir cours de maths financière 2A).
- Calcul du prix : $(u(n, x), n = 0, \ldots, N, x \in \mathbb{X})$

$$(2) \qquad \begin{cases} u(n, x) = p u(n+1, xu) + (1-p) u(n+1, xd), \\ u(N, x) = f(x). \end{cases}$$

- $\mathbb{E}\left[f(X_N)\right] = u(0, x_0)$.

Un programme

Un programme en Python, catastrophique (complexité 2^N), mais qui a le bon goût de terminer (lentement)

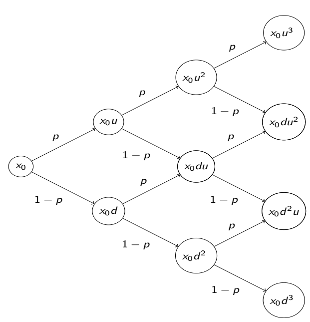
```
N=10;p=1/2;
down=1/2;up=2;
K=1;

def f(x): return max(x-K,0)

def u(n,x):
   if n==N:
      return f(x);
   else:
      return p * u(n+1,x*up)+ (1-p) * u(n+1,x*down);

def prix_rec(x): return u(0,x)
```

Pour aller plus vite, un dessin pour $N = 3$



Un algorithme qui fait la même chose, plus vite ...

- Si $X_0 = x_0$ les seules valeurs que peut prendre X_n sont données par (k = nombre de fois où l'on tire u entre 0 et $n-1$)
$$(x_k^{(n)} = x_0 u^k d^{n-k}, 0 \leq k \leq n).$$
- Il suffit de calculer les vecteurs $U^{(n)} = (u(n, x_k^{(n)}), 0 \leq k \leq n)$, qui vérifient
$$(3) \quad \begin{cases} U^{(N)}(k) = f(x_k^{(N)}), \quad k = 0, \ldots, N \\ U^{(n)}(k) = pU^{(n+1)}(k+1) + (1-p)U^{(n+1)}(k), \quad k = 0, \ldots, n. \end{cases}$$
- Ce qui donne un agorithme (de complexité N^2).

Un algorithme qui fait la même chose, plus vite ...

```
def prix_iter(x_0):
  U=np.zeros((N+1,N+1))
  for k in range(N+1): # = intervalle [0,1,...,N]
    U[N,k] = f(x_0 * up**k * down**(N-k))

  for n in range(N-1,-1,-1):  # = intervalle [N-1,N-2,...,0]
    for k in range(n+1):  # intervalle [0,1,...,n]
        U[n,k] = p*U[n+1,k+1]+(1-p)*U[n+1,k]
  return U[0,0]
```

Andry Markov

- Mathematicien russe 1856 – 1922, élève de Chebyshev, "known for his work in number theory, analysis, and probability theory."
- ... "He introduced a new sequence of dependent variable, called a chain, as well as a few basic concepts of chains such as transition probabilities, irreductibility and stationarity. His ideas were taken up and developed further by scientists around the world and now the theory of Markov Chains is one of the most powerful theories for analyzing various phenomena of the world"

А. А. Марков (1886).

Pour des détail sur la vie et l'œuvre de Markov voir [Basharin et al.(2004)Basharin, Langville, and Naumov].

Pour aller plus loin : cours de l'Ecole des Ponts

- 2A : cours "Processus Aléatoires"
- cours "Méthodes Mathématiques pour la Finance"

Bibliographie

Gely P. Basharin, Amy N. Langville, and Valeriy A. Naumov.
The life and work of A. A. Markov.
Linear Algebra Appl., 386:3–26, 2004.

Nathanaël Berestycki.
Notes on card shuffling.
Technical report, http://www.statslab.cam.ac.uk/~beresty/Articles/shuffle.pdf, 2007.

P.L. Bernstein.
Capital Ideas: The Improbable Origins of Modern Wall Street.
The Free Press, 1992.

P.L. Bernstein.
Capital Ideas Evolving.
Jhon Wiley and Son, 2007.

F. Thomas Bruss.
The art of a right decision.
from Oxfo, page 15, 2006.

Jean-François Delmas and Benjamin Jourdain.
Modèles aléatoires.
Mathématiques & Applications. Springer-Verlag, Berlin, 2006.

Mark R. Dennis, Paul Glendinning, Paul A. Martin, Fadil Santosa, and Jared Tanner, editors.
The Princeton companion to applied mathematics.
Princeton University Press, Princeton, NJ, 2015.
p. 648–658.

Stuart Dreyfus.
Richard bellman on the birth of dynamic programming.
Operations Resarch, 50(1):48–51, 2002.

Hans Föllmer and Alexander Schied.
Convex and coherent risk measures.
http://www.alexschied.de/Encyclopedia6.pdf, 2008.

Exercice 4 On considère une suite de variables aléatoires indépendantes de même loi $(U_n, n \geq 1)$ prenant ces valeurs dans G (noter qu'il n'est pas nécessaire de supposer G fini). Soit E un ensemble fini et F une application de $E \times G$ dans E.

On construit, par récurrence, une suite $(X_n, n \geq 0)$ en posant, $X_0 = x \in E$ et

$$X_{n+1} = F(X_n, U_{n+1}), n \geq 0.$$

1. Montrer que $(X_n, n \geq 0)$ est une chaîne de Markov et exprimer sa matrice de transition P en fonction de F.
2. On définit τ par
$$\tau = \inf \{n \geq 0, X_n = x_0\}.$$
Montrer que $Y_n = X_{n \wedge \tau}$ est aussi une chaîne de Markov. Calculer sa matrice de transition en fonction de celle de X.

Exercice 5
1. Soit $(X_n, n \geq 0)$ une chaîne de Markov de matrice de transition $(P(x,y), x \in \mathbb{E}, y \in E)$ et de loi initiale (i.e. la loi de X_0) $(\mu(x), x \in \mathbb{E})$. Calculer la loi du vecteur $(X_0, X_1, \ldots, X_n)$ et en déduire la loi de X_n.
2. Montrez que $P^n(x,y) = \mathbb{P}(X_n = y | X_0 = x)$. En particulier, montrez que si la chaîne part d'un point déterministe x_0 [i.e. $X_0 = x_0$ avec probabilité 1], alors
$$\mathbb{P}(X_n = y) = P^n(x_0, y).$$
3. Calculer la loi du couple (X_n, X_N), pour $n < N$, et en déduire que
$$\mathbb{E}\left(f(X_N) | X_n = x\right) = P^{N-n} f(x),$$
et, en particulier, que si $X_0 = x_0$ avec probabilité 1, $\mathbb{E}(f(X_N)) = P^N f(x_0)$.

Exercice 6 Soit $(X_n, n \geq 0)$ un processus de Markov de matrice de transition P sur un espace fini E. Pour une fonction f de E dans $\mathbb{R}$, on note $P^0 f(x) = f(x)$, puis récurence

$$P^{k+1} f(x) = \sum_{y \in E} P(x,y) P^k f(y).$$

1. Vérifier que P^k est bien la puissance k-ième (au sens du produit des matrices) de P.

On définit $u(n,x)$ par

$$u(n,x) = P^{N-n} f(x), x \in E, n \leq N,$$

2. Vérifier que l'on a
$$\left\{ \begin{array}{lll} u(N,x) & = & f(x) \quad x \in E \\ u(n,x) & = & \sum_{y \in E} P(x,y) u(n+1,y) \quad x \in E, n < N. \end{array} \right.$$
3. Montrer que $\mathbb{E}(u(n, X_n)) = \mathbb{E}(u(n+1, X_{n+1}))$ pour $0 \leq n < N$.
4. Redémontrer à partir de cette propriété, le résultat de l'exercice précédent : si la chaîne part de x_0 à l'instant 0, $\mathbb{E}(f(X_n)) = u(0, x_0) = P^N f(x_0)$.

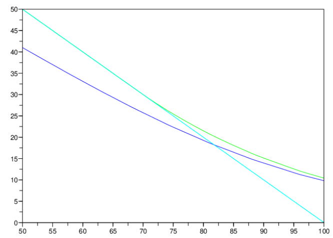
Chapitre 3

Temps d'arrêt, Arrêt optimal

Recrutement optimal, Options américaines

Temps d'arrêt, Arrêt optimal

Recrutement optimal, Options américaines



J.-P. Chancelier, B. Lapeyre

http://cermics.enpc.fr/~jpc/decision-incertain

Plan

Processus de Markov

- Processus de Markov et propriété de Markov

$$P\left(X_{n+1}=x_{n+1}|X_n=x_n, X_{n-1}=x_{n-1},\dots,X_0=x_0\right)$$
$$=\mathbb{P}\left(X_{n+1}=x_{n+1}|X_n=x_n\right)$$

- la matrice de transition : $(P(x,y), x, y \in \mathbb{X})$

$$P(x,y)=\mathbb{P}\left(X_{n+1}=y|X_n=x\right)$$

- Loi de $(X_0,\dots,X_n)$ pour $n\in\mathbb{N}$ déterminée par P et la loi de X_0 (notée μ_0).

$$\mathbb{P}\left(X_0=x_0,\dots,X_n=x_n\right)=\mu_0(x_0)P(x_0,x_1)\times\cdots\times P(x_{n-1},x_n)$$

- Loi de X_n pour $n\in\mathbb{N}$, μ_N donnée par

$$\mu_n=\mu_0 P^n=\mu_{n-1}P$$

Loi de X_n : une autre façon d'écrire les choses

- $(X_n, n\geq 0)$ chaîne de Markov de matrice de transition P sur $\mathbb{X}$.
- $\mathbb{P}(X_0=x_0)=1$ (i.e. $\mu_0(x_0)=1$, $\mu_0(x)=0$ si $x\neq x_0$.)
- $N\geq 0$ fixé. On sait exprimer $\mathbb{E}[f(X_N)]$:

$$\mathbb{E}[f(X_N)]=\mu_N f=\mu_0 P^N f=\sum_{x\in\mathbb{X}}\mu_0(x)(P^N f)(x)=(P^N f)(x_0)$$

Une formulation alternative plus algorithmique.

Theorem 1

Soit $(u_n(x), n\in[\![0,N]\!], x\in\mathbb{X})$ la solution unique de

$$(1)\qquad \begin{cases} u_n(x)=\sum_{y\in\mathbb{X}}P(x,y)u_{n+1}(y)\ ,\ \ \forall x\in\mathbb{X}\ ,\ \ n<N \\ u_N(x)=f(x)\ ,\ \ \forall x\in\mathbb{X}\end{cases}$$

Alors

$$\mathbb{E}[f(X_N)]=u_0(x_0)$$

.

Remarques

- La première équation de (1) peut aussi s'écrire

$$u_n = P u_{n+1} \quad \text{ou} \quad u_n = P^{N-n} f = \overbrace{P \times \cdots \times P}^{N-n \text{ fois}} f$$

- Pour $x \in \mathbb{X}$, lorsque $\mathbb{P}(X_n = x) > 0$, on a

$$u_n(x) = \mathbb{E}[f(X_N) \mid X_n = x] = P^{N-n} f(x)$$

Preuve $\mathbb{P}(X_0 = x_0, \ldots, X_n = x, X_{n+1} = x_{n+1}, \ldots, X_N = x_N)$
$= \mathbb{P}(X_0 = x_0) P(x_0, x_1) \times \cdots \times P(x, x_{n+1}) \times \cdots \times P(x_{N-1}, x_N)$

en sommant sur toutes les valeurs de $x_0, \cdots, x_{n-1}, x_{n+1}, \cdots, x_{N-1}$ on obtient $\mathcal{L}_{(X_n, X_N)} : \mathbb{P}(X_n = x_n, X_N = x_N) = \mathbb{P}(X_n = x_n) P^{N-n}(x_n, x_N)$

$$\begin{aligned}\mathbb{E}\left[f(X_N) 1_{\{X_n = x_n\}}\right] &= \mathbb{P}(X_n = x_n) \sum_{x_N \in \mathbb{X}} P^{N-n}(x_n, x_N) f(x_N) \\ &= \mathbb{P}(X_n = x_n)(P^{N-n} f)(x_n)\end{aligned}$$

Remarques et notations

- (1) est une équation de programmation dynamique [1]
- $\mathbb{X}$ est fini, (1) est un algorithme qui termine.

Notations pour la suite

- $x_{0:n}$ plutôt que $(x_0, \ldots, x_n)$
- $X_{0:n}$ plutôt que $(X_0, \ldots, X_n)$
- $X_{0:n} = x_{0:n}$ plutôt que $(X_0 = x_0, \ldots, X_n = x_n)$

[1] Sur l'origine du terme programmation dynamique voir [Dreyfus(2002)]. "It was something not even a congressman could object to" selon Bellman.

Une autre preuve

On va montrer que pour tout $n \in \{0, \dots, N-1\}$

$$\mathbb{E}\left[u_{n+1}(X_{n+1})\right] = \mathbb{E}\left[u_n(X_n)\right]$$

Et ainsi, on aura :

$$u_0(x_0) = \mathbb{E}\left[u_0(X_0)\right] = \dots = \mathbb{E}\left[u_N(X_N)\right] = \mathbb{E}\left[f(X_N)\right].$$

Par la propriété de Markov, on obtient

$$\mathbb{P}\left(X_{0:n+1} = x_{0:n+1}\right) = \mathbb{P}\left(X_{0:n} = x_{0:n}, X_{n+1} = x_{n+1}\right) = \mathbb{P}\left(X_{0:n} = x_{0:n}\right) P(x_n, x_{n+1}).$$

$$\begin{aligned}
\mathbb{E}\left[u_{n+1}(X_{n+1})\right] &= \sum_{x_{0:n+1} \in \mathbb{X}^{n+2}} u_{n+1}(x_{n+1}) \mathbb{P}\left(X_{0:n+1} = x_{0:n+1}\right) \\
&= \sum_{x_{0:n} \in \mathbb{X}^{n+1}, x_{n+1} \in \mathbb{X}} u_{n+1}(x_{n+1}) \mathbb{P}\left(X_{0:n} = x_{0:n}\right) P(x_n, x_{n+1}) \\
&= \sum_{x_{0:n} \in \mathbb{X}^{n+1}} \mathbb{P}\left(X_{0:n} = x_{0:n}\right) \sum_{x_{n+1} \in \mathbb{X}} P(x_n, x_{n+1}) u_{n+1}(x_{n+1}) \\
&= \sum_{x_{0:n} \in \mathbb{X}^{n+1}} \mathbb{P}\left(X_{0:n} = x_{0:n}\right) u_n(x_n) = \mathbb{E}\left[u(n, X_n)\right]
\end{aligned}$$

La question du jour

- $(X_n, n \geq 0)$ une chaîne de Markov de matrice de transition P sur $\mathbb{X}$. $\mathbb{P}(X_0 = x_0) = 1$.
- On cherche à calculer

$$\sup_{\tau \leq N} \mathbb{E}\left[f(X_\tau)\right]$$

- τ appartenant à une famille de temps aléatoires,
  - plus grande (au sens de l'inclusion) que les temps déterministes,
  - mais plus petite que tous les temps aléatoires
- plus grande que les temps déterministes : parce que l'on souhaite pouvoir tenir compte des valeurs de X_n observées au cours du temps.
- plus petite que tous les temps aléatoires : parce que à l'instant n on ne connait pas la trajectoire future $X_{n+1}, \dots, X_N$

$$\tau = \text{Argmax}\{f(X_n), 0 \leq n \leq N\}$$

  est optimum, mais réclame de connaitre le futur !
- Voir Pour la Sciences [Hill(2009)] pour une introduction à ce genre de problème.

Temps d'arrêt

La notion adéquate est celle de temps d'arrêt : on souhaite prendre la décision de s'arrêter au temps n avec l'information que l'on a au temps n.

Definition 2

τ est un temps d'arrêt, si, pour tout n, il existe $A_n \subset \mathbb{X}^{n+1}$ tel que

$$\{\tau = n\} = \{(X_0, X_1, \dots, X_n) \in A_n\}$$

- "Je peux déterminer si $\tau = n$ en ne considérant que la portion de trajectoire avant n".
- Argmax$\{f(X_n), 0 \leq n \leq N\}$ ne peut pas être un temps d'arrêt (sauf cas particulier).
- Le temps d'atteinte d'un point z de $\mathbb{X}$ est un temps d'arrêt
$$\tau = \inf \{n \geq 0, X_n = z\}.$$
En effet $\{\tau = n\} = \{X_0 \neq z, X_1 \neq z, \dots, X_{n-1} \neq z, X_n = z\}$.

Formulation d'un problème d'arrêt optimal

- On se donne une chaîne de Markov $(X_n, n \geq 0)$ sur $\mathbb{X}$, de matrice de transition P, issue de x_0 en 0 ($\mathbb{P}(X_0 = x_0) = 1$).
- $f : \mathbb{N} \times \mathbb{X}$ donné, on cherche à calculer le sup suivant
$$\sup_{\tau \text{ t.a.} \leq N} \mathbb{E}\left[f(\tau, X_\tau)\right]$$
- et aussi à trouver un temps d'arrêt τ qui réalise ce sup.
- ... on sait répondre complétement à ces deux questions.

Les données : P, x_0, f.

Solution du problème d'arrêt optimal

Theorem 3

Si $(u_n(x), n = 0, \dots, N, x \in \mathbb{X})$ est la solution unique de

$$(2) \quad \begin{cases} u_n(x) = \max\left\{ \sum_{y\in\mathbb{X}} P(x,y)u_{n+1}(y), f(n,x)\right\}, & n < N, \ \forall x \in \mathbb{X} \\ u_N(x) = f(N,x), \ \forall x \in \mathbb{X}. \end{cases}$$

Alors

- $\sup_{\tau\leq N} \mathbb{E}\left[f(\tau, X_\tau)\right] = u_0(x_0)$
- $\tau_0 = \inf\{n \geq 0, u_n(X_n) = f(n, X_n)\}$ *est un temps d'arrêt optimal.*

- Lorsque la matrice de transition dépend de n il faut remplacer P par P_n (la matrice de transition entre les instants n et $n+1$) dans l'équation.
- τ_0 est bien un temps d'arrêt $\leq N$.

$$\{\tau_0 = n\} = \{u(0, X_0) \neq f(0, X_0), \dots, u(n-1, X_{n-1}) \neq f(n-1, X_{n-1}), u(n, X_n) = f(n, X_n)\}$$

- Preuve formelle du théorème : transparent 16.

Commentaires

- (2) est une équation de programmation dynamique.
- Interprétation de u_n : $u_n(x) = \sup_{n\leq\tau\leq N} \mathbb{E}\left[f(\tau, X_\tau) \,\middle|\, X_n = x\right]$.
- Preuve informelle ("Principe d'optimalité") :

  Au temps n, si $X_n = x$

  - *soit je m'arrête et je gagne $f(n,x)$*
  - *soit j'attends le temps suivant n+1 où je peux gagner $u_{n+1}(X_{n+1})$.*
  - *La moyenne de mes gains futurs si je passe à $n+1$ est donnée par*

  $$\mathbb{E}\left[u_{n+1}(X_n+1) \,\middle|\, X_n = x\right] = \sum_{y\in\mathbb{X}} P(x,y)u_{n+1}(y)$$

  - *je m'arrête ou je continue en regardant*

  $$\max\left(f(n,x), \sum_{y\in\mathbb{X}} P(x,y)u_{n+1}(y)\right)$$

- On l'utilise (2) pour écrire un algorithme

Exemple 1: Calcul du prix d'options américaines

- $(X_n, 0 \leq n \leq N)$ une chaîne de Markov de matrice de transition P décrivant l'évolution des prix des actifs (e.g. modèle de Cox-Ross).
- J'ai la possibilité si j'"exerce" mon option en $n \leq N$ de gagner $f(n, X_n)$.
- Que vaut ce droit ?. À quel moment dois-je exercer ce droit pour maximiser mon gain ?
- Pour calculer le prix (i.e. la valeur de ce droit), il est naturel de chercher à maximiser l'espérance du flux $\mathbb{E}\left[f(\tau, X_\tau)\right]$ parmi tous les temps d'arrêt de X.
- Justification compléte : cours de Mathématiques Financière (2A) ou [Lamberton and Lapeyre(1997)].

Le "put américain" dans le modèle de Cox-Ross

- On cherche à calculer $\sup_{\tau \leq N} \mathbb{E}\left[f(\tau, X_\tau)\right]$, le prix de l'option et à determiner le moment d'exercice optimum.
- $(X_n, 0 \leq n \leq N)$ est le processus de Cox-Ross

$$X_0 = 1, X_{n+1} = X_n \left(u 1_{\{U_{n+1} = P\}} + d 1_{\{U_{n+1} = F\}}\right).$$

  $(U_n, n \geq 1)$ une suite de tirage à pile ou face indépendant,

$$0 \leq p \leq 1, \quad \mathbb{P}(U_n = P) = p = 1 - \mathbb{P}(U_n = F).$$

- r le taux d'intérêt sur une période, K le strike, $d < 1 + r < u$.

$$f(n, x) = \frac{1}{(1+r)^n}(K - x)_+$$

Solution

- On commence par calculer (cf TD) $(u_n(x), n = 0, \ldots, N, x \in \mathbb{X})$ la solution de l'équation (2) du théorème 3, ici

$$(3) \qquad \begin{cases} u_n(x) = \max\left(p\ u_{n+1}(xu) + (1-p)\ u_{n+1}(xd), \dfrac{(K-x)_+}{(1+r)^n}\right) \\ u_N(x) = \dfrac{(K-x)_+}{(1+r)^N}. \end{cases}$$

- Souvent on calcule $v_n(x) = (1+r)^n u_n(x)$ plutot que $u_n(x)$.

$$\begin{cases} v_n(x) = \max\left(\dfrac{p\ v_{n+1}(xu) + (1-p)\ v_{n+1}(xd)}{1+r}, (K-x)_+\right) \\ v_N(x) = (K-x)_+. \end{cases}$$

- $u_0(x) = v_0(x) = \sup_{\tau \leq N} \mathbb{E}\left[f(\tau, X_\tau)\right]$
- τ_0 le temps d'arrêt optimal

$$\begin{aligned} \tau_0 &= \inf\left\{n \geq 0, u_n(X_n) = (K - X_n)_+/(1+r)^n\right\} \\ &= \inf\left\{n \geq 0, v_n(X_n) = (K - X_n)_+\right\}. \end{aligned}$$

Preuve du théorème 3 : τ temps d'arrêt quelconque

- u solution (2), on va voir que, pour tout τ t.a., $\mathbb{E}\left[u_{n\wedge\tau}(X_{n\wedge\tau})\right]$ décroît en n.
- En admettant la décroissance en n, on obtient pour tout τ t.a.

$$\begin{aligned} u_0(x_0) = \mathbb{E}\left[u_0(X_0)\right] &= \mathbb{E}\left[u_{0\wedge\tau}(X_{0\wedge\tau})\right] \\ &\geq \mathbb{E}\left[u_{N\wedge\tau}(X_{N\wedge\tau})\right] = \mathbb{E}\left[u_\tau(X_\tau)\right] \geq \mathbb{E}\left[f(\tau, X_\tau)\right]. \end{aligned}$$

$$\Longrightarrow \sup_{0 \leq \tau \leq N, \text{t.a.}} \mathbb{E}\left[f(\tau, X_\tau)\right] \leq u_0(x_0)$$

- Pour montrer la décroissance annoncée, on remarque que

$$\begin{aligned} \Delta_{n+1} &= \mathbb{E}\left[u_{(n+1)\wedge\tau}(X_{(n+1)\wedge\tau})\right] - \mathbb{E}\left[u_{n\wedge\tau}(X_{n\wedge\tau})\right] \\ &= \mathbb{E}\left[(u_{n+1}(X_{n+1}) - u_n(X_n))1_{\{\tau\, \geq\, n+1\}}\right] \end{aligned}$$

Et on montre que $\Delta_{n+1} \leq 0$

Preuve du théorème 3 : τ temps d'arrêt quelconque

- τ est un temps d'arrêt, $\{\tau \geq n+1\} = \{\tau \leq n\}^c$, s'écrit sous la forme
$$\{\tau \geq n+1\} = \{X_{0:n} \in \bar{A}_n\}$$
- La loi de $(X_{0:n}, X_{n+1})$ est donnée par (c'est la propriété de Markov)
$$\mathbb{P}\left(X_{0:n} = x_{0:n}, X_{n+1} = x_{n+1}\right) = \mathbb{P}\left(X_{0:n} = x_{0:n}\right) P(x_n, x_{n+1}).$$

$$\begin{aligned}
\Delta_{n+1} &= \mathbb{E}\left[(u_{n+1}(X_{n+1}) - u_n(X_n))1_{\{X_{0:n} \in \bar{A}_n\}}\right] \\
&= \sum_{x_{0:n} \in \bar{A}_n, x_{n+1} \in \mathbb{X}} (u_{n+1}(x_{n+1}) - u_n(x_n))\mathbb{P}(X_{0:n} = x_{0:n})P(x_n, x_{n+1}) \\
&= \sum_{x_{0:n} \in \bar{A}_n} \mathbb{P}(X_{0:n} = x_{0:n})\Big(\sum_{x_{n+1} \in \mathbb{X}} u_{n+1}(x_{n+1})P(x_n, x_{n+1}) - u_n(x_n)\Big)
\end{aligned}$$

u est sol. de (2), $\sum_{x_{n+1} \in \mathbb{X}} u_{n+1}(x_{n+1})P(x_n, x_{n+1}) \leq u_n(x_n)$

$$\leq 0$$

Preuve du théorème 3 : τ_0 temps d'arrêt optimal

- $\tau_0 = \inf\{n \geq 0, u_n(X_n) = f(n, X_n)\}$.
- On va montrer que $\mathbb{E}\left[u\left(n \wedge \tau_0, X_{n \wedge \tau_0}\right)\right]$ est constant en n (i.e. $\Delta_{n+1} = 0$).
- En admettant ceci, il est facile de conclure
$$\begin{aligned}
u_0(x_0) = \mathbb{E}\left[u_0(X_0)\right] &= \mathbb{E}\left[u_{0 \wedge \tau_0}(X_{0 \wedge \tau_0})\right] \\
&= \mathbb{E}\left[u_{N \wedge \tau_0}(X_{N \wedge \tau_0})\right] = \mathbb{E}\left[u_{\tau_0}(X_{\tau_0})\right]
\end{aligned}$$
Mais, par définition de τ_0, $u_{\tau_0}(X_{\tau_0}) = f(\tau_0, X_{\tau_0})$, donc
$$u_0(x_0) = \mathbb{E}\left[f(\tau_0, X_{\tau_0})\right]$$
- Ce qui finit la démonstration puisque τ_0 réalise alors le sup.
- Il nous reste à montrer que, pour ce temps d'arrêt $\Delta_{n+1} = 0$

Preuve du théorème 3 : τ_0 temps d'arrêt optimal

- τ_0 est un temps d'arrêt, $\{\tau_0 \geq n+1\} = \{X_{0:n} \in \bar{A}_n^0\}$ où
$$\bar{A}_n^0 = \{u_0(x_0) \neq f(0, x_0), \dots, u_n(x_n) \neq f(n, x_n)\}$$
- Sur l'événement $\{\tau_0 \geq n+1\}$, on a $u_n(X_n) \neq f(n, X_n)$ et comme u sol. de (2)
$$\sum_{x_{n+1} \in \mathbb{X}} u_{n+1}(x_{n+1}) P(X_n, x_{n+1}) = u_n(X_n)$$
- On termine alors comme tout à l'heure, mais en tenant compte de cette égalité :
$$\begin{aligned}
\Delta_{n+1} &= \mathbb{E}\Big[(u_{n+1}(X_{n+1}) - u_n(X_n)) 1_{\left\{X_{0:n} \in \bar{A}^0{}_n\right\}}\Big] \\
&= \sum_{x_{0:n} \in \bar{A}_n^0, x_{n+1} \in \mathbb{X}} (u_{n+1}(x_{n+1}) - u_n(x_n)) \mathbb{P}(X_{0:n} = x_{0:n}) P(x_n, x_{n+1}), \\
&= \mathbb{E}\Big[1_{\left\{X_{0:n} \in \bar{A}_n^0\right\}} \big(\sum_{x_{n+1} \in \mathbb{X}} u_{n+1}(x_{n+1}) P(X_n, x_{n+1}) - u_n(X_n)\big)\Big] \\
&= 0.
\end{aligned}$$

Exemple 2: un problème de recrutement

- Il y a N candidats à un poste
- Je reçois les candidats consécutivement
- Les circonstances m'imposent de décider tout de suite du recrutement (soit je recrute la personne que je viens de recevoir, soit je la refuse définitivement).
- La seule information que j'ai sur les candidats est leur classement (je peux comparer un candidat à ceux que j'ai vu précédemment)
- Je souhaite maximiser la probabilité de recruter le meilleur candidat
- Quelle est la façon de s'y prendre ?

Le résultat

- Il faut recevoir (environ) 37% des candidats sans les recruter
- Puis, choisir le premier candidat qui suit qui est meilleur que tous les précédents (le dernier si cela n'arrive jamais).
- On obtient ainsi un temps d'arrêt optimal.
- La probabilité d'obtenir le meilleur candidat est (environ) de 37%.
- Les transparents qui suivent expliquent comment arriver à ces résultats à l'aide de la théorie précédente.

Le modèle

- $\omega = (\omega_1, \dots, \omega_N)$ une permutation de $(1, \dots, N)$.
- ω_k le classement du $\#k$-ième individu dans la permutation.

$$\begin{array}{llllllll} \textit{Indice} & (\#1 & \#2 & \dots & \#k & \dots & \#N) \\ \textit{Rang} & (\omega_1 & \omega_2 & \dots & \omega_k & \dots & \omega_N) \end{array}$$

- Ω_N l'ensemble des permutations de $(1, \dots, N)$ muni de la probabilité uniforme.
- B_n l'évenement "le n-ième candidat est le meilleur".
- On cherche un temps d'arrêt τ qui maximise $\mathbb{P}(B_\tau)$.

Où est la chaîne de Markov ?

- Un temps d'arrêt mais pour quel processus de Markov ?
- $R_k(\omega)$ le rang du $\#k$-ième individu parmi les k premiers individus.
- $B_n = \{R_n = 1, R_{n+1} > 1, \dots, R_N > 1\}$.
- Quelle est la loi de $(R_1, \dots, R_N)$?

Un exemple de calcul de $R = (R_1, \dots, R_N)$

- $\omega = (2\ 3\ 1\ 4)$ donne $R = (1\ 2\ 1\ 4)$.
- $R = (1\ 2\ 1\ 4)$ donne $\#3 \leq \#1 \leq \#2 \leq \#4$ et donc ω

$(1) \to (\#1)$	$\#1$ est le premier puisqu'il est tout seul!
$(1\ 2) \to (\#1 \leq \#2)$	$\#2$ est le deuxième parmi les 2 premiers
$(1\ 2\ 1) \to (\#3 \leq \#1 \leq \#2)$	$\#3$ est le premier parmi les 3 premiers
$(1\ 2\ 1\ 4) \to (\#3 \leq \#1 \leq \#2 \leq \#4)$	$\#4$ est le quatrième parmi les 4 premiers

On obtient la permutation de départ $\omega = (2\ 3\ 1\ 4)$

$$\begin{array}{lcccc} \textit{Indice} & (\#1 & \#2 & \#3 & \#4) \\ \textit{Rang} & (2 & 3 & 1 & 4) \end{array}$$

En Python

Si vous avez un doute voici, deux fonctions Python qui font le job.

```
import numpy as np

def Omega2R(omega):
# Calcule les rang d'insertion pour un omega donne
  R=np.zeros(np.size(omega), dtype=int)
  # dtype=int, pour specifier que l'on veut un tableau d'entier
  for n in range(np.size(omega)):
    # classe le vecteur omega[1,...,n] en croissant
    y=np.sort(omega[0:n+1]);
    # R[n] = le classement de omega[n] parmi les n premiers
    R[n]=np.where(y==omega[n])[0][0]+1
  return R

def R2Omega(R):
    # Calcule omega connaissant les rangs d'insertion
    iomega=np.zeros(0, dtype=int);
    for n in range(np.size(R)):
        # J'insére n à l'indice R[n]
        iomega=np.concatenate([iomega[0:int(R[n]-1)],[n+1],iomega[int(R[n]-1):n+1]])
    # On inverse la permutation
    omega=np.zeros(np.size(R), dtype=int)
    for n in range(np.size(omega)):
        omega[iomega[n]-1]=n+1
    return omega

Taille=6
omega=np.random.permutation(Taille)+1 # permutation aléatoire
print(omega)
print(R2Omega(Omega2R(omega))) # devrait être égal à omega
```

Calcul de lois

- À un R correspond un et un seul ω, **donc**, pour $\alpha_k \in \{1, \ldots, k\}$

$$P(R_1 = \alpha_1, \ldots, R_N = \alpha_N) = \mathbb{P}(\{\omega\}) = \frac{1}{N!}.$$

- (Par contraction) les R_k suivent des lois uniformes sur $\{1, \ldots, k\}$. Elles sont indépendantes
- $S_k = 1_{\{R_k = 1\}}$ sont aussi des variables aléatoires indépendantes. Elles suivent des lois de Bernoulli de paramètre $p_k = \mathbb{P}(S_k = 1) = 1/k$
- La suite de variable $(S_1, \ldots, S_N)$ suffira pour traiter le problème.

Le critère

- B_n l'évenement "le n-ième candidat est le meilleur".

$$B_n = \{R_n = 1, R_{n+1} > 1, \dots, R_N > 1\} = \{S_n = 1, S_{n+1} = 0, \dots, S_N = 0\}.$$

- τ un temps d'arrêt par rapport au processus $R = (R_1, \dots, R_N)$. À l'instant n, $(R_1, \dots, R_n)$ "contient toute l'information disponible".
- Pour un temps d'arrêt[2] τ, on va voir que

$$\mathbb{P}(B_\tau) = \mathbb{E}\left[\frac{\tau}{N} S_\tau\right].$$

ce qui permettra de mettre le problème "sous forme markovienne".

[2]Ce n'est plus vrai sinon ... Exercice: trouver un contre-exemple.

Preuve

$$\{\tau = n\} = \{(R_1, \dots, R_{n-1}, R_n) \in A_n\} = \{R_{1:n} \in A_n\}.$$

Notation : $R_{1:n} = (R_1, \dots, R_{n-1}, R_n)$.

$$\begin{aligned}
\mathbb{P}(\tau = n, B_\tau) = \mathbb{P}(\tau = n, B_n).& \\
&= \mathbb{P}(\tau = n, R_n = 1, R_{n+1} > 1, \dots, R_N > 1) \\
&= \mathbb{P}(R_{1:n} \in A_n, R_n = 1, R_{n+1} > 1, \dots, R_N > 1) \\
&= \mathbb{P}(R_{1:n} \in A_n, R_n = 1)\mathbb{P}(R_{n+1} > 1, \dots, R_N > 1)
\end{aligned}$$

car indépendance des vecteurs $R_{1:n}$ et $R_{n+1:N}$, puis par indépendance des R_n entre eux

$$\mathbb{P}(R_{n+1} > 1, \dots, R_N > 1) = \frac{n}{n+1}\frac{n+1}{n+2} \times \dots \times \frac{N-1}{N} = \frac{n}{N}$$

Preuve

$$\begin{aligned}\mathbb{P}(\tau = n, B_\tau) &= \frac{n}{N}\mathbb{P}(R_{1:n} \in A_n, R_n = 1) = \frac{n}{N}\mathbb{P}(\tau = n, R_n = 1) \\ &= \frac{n}{N}\mathbb{E}\big[1_{\{\tau = n,\, R_n = 1\}}\big] = \mathbb{E}\big[1_{\{\tau = n\}}\frac{n}{N}S_n\big] \\ &= \mathbb{E}\big[1_{\{\tau = n\}}\frac{\tau}{N}S_\tau\big].\end{aligned}$$

En sommant pour n variant de 1 à N

$$\mathbb{P}(B_\tau) = \mathbb{E}\big[\frac{\tau}{N}S_\tau\big].$$

Résumons nous !

- $(S_1, \dots, S_N)$ est une suite de variables aléatoires indépendantes de Bernoulli $1/k$,
  donc une chaîne de Markov sur l'espace $\mathbb{X} = \{0, 1\}$, non homogène, de matrice de transition, dépendant du temps, P_n

$$\begin{aligned}P_n(0,0) = P_n(1,0) &= 1 - \frac{1}{n+1} = \frac{n}{n+1} \\ P_n(0,1) = P_n(1,1) &= \frac{1}{n+1}\end{aligned}$$

- On chercher à maximiser $\mathbb{P}(B_\tau) = \mathbb{E}\big[\frac{\tau}{N}S_\tau\big] = \mathbb{E}\big[f(\tau, S_\tau)\big]$ parmi tous les temps d'arrêt plus petits que N.

Résolution

- On peut résoudre le problème grâce à (2)
- On calcule ($u_n(0)$ et $u_n(1)$ pour $1 \leq n \leq N$)
  la solution du problème est donnée par $u_1(1)$ (noter que $S_1 = 1$)

$$\begin{cases} u_n(x) = \max\left\{\frac{n}{n+1}u_{n+1}(0) + \frac{1}{n+1}u_{n+1}(1), \frac{n}{N}x\right\}, n < N\ , \quad \forall x \in \{0,1\} \\ u_N(x) = x\ , \quad \forall x \in \{0,1\} \end{cases}$$

- Deux suites à calculer $u_n(0)$ et $u_n(1)$ pour $1 \leq n \leq N$)
- Une fois ceci fait, un temps d'arrêt optimal est obtenu par

$$\tau = \inf\left\{n \geq 0, u_n(S_n) = \frac{n}{N}S_n\right\}.$$

- Dans ce cas, on peut mener des calculs explicites (voir [Delmas and Jourdain(2006)] et TD) pour obtenir les résultats annoncés.

- En posant, $f_n(x) = \frac{n}{N}x$ on obtient pour $x \in \{0,1\}$

$$u_n(x) = \max\left(\frac{n}{n+1}u_{n+1}(0) + \frac{1}{n+1}u_{n+1}(1), f_n(x)\right)\ , \quad u_N(x) = f_N(x)\ .$$

- Comme $u_n(x)$ est positive et $f_n(0) = 0$ on obtient

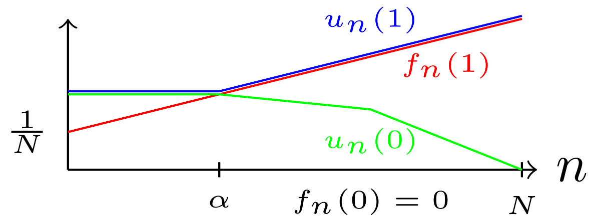
$$u_n(0) = \frac{1}{n+1}u_{n+1}(1) + \frac{n}{n+1}u_{n+1}(0), \quad u_n(1) = \max\left(u_n(0), \frac{n}{N}\right).$$

- Pour $N > 1$, $\exists\alpha \in [1, N)$ tel que $\max\left(u_\alpha(0), \frac{\alpha}{N}\right) = u_\alpha(0)$. **preuve par contradiction: on autait sinon $u_n(1) = n/N > u_n(0)$ pour tout $n \in [1, N]$ ce qui donne**

$$u_n(0) = \frac{n}{N}\sum_{i=n}^{N-1}\frac{1}{i}$$

  On aurait alors $u_1(0) \geq 1/N$ ce qui donne une contradiction.
- $\forall n \in [1, N)$, $u_n(1) = u_n(0) = \beta \geq n/N$
  $\Longrightarrow\ \forall p \in [1, n-1]u_p(1) = u_p(0) = \beta > p/N$ **La preuve se fait par recurence**
- On appelle α^* le plus grand α qui vérifie la propriété au dessus.

- Le coût optimal



  - $u_n(1) = u_n(0) = u_{\alpha^*}(1) > n/N\ ,\ \ \forall n \in [1, \alpha^*[$
  - $u_{\alpha^*}(1) = u_{\alpha^*}(0) \geq \alpha^*/N$
  - $u_n(1) = f_n(1) > u_n(0) = \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i}\ ,\ \ \forall n \in]\alpha^*, N]$

- La stratégie optimale

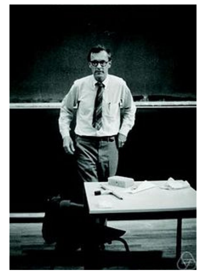
$$\tau = \inf\{n \geq 1 \,|\, u_n(S_n) = f_n(S_n)\}$$

  - Si $S_n = O$ comme $u_n(0) > f_n(0)$ pour $n \neq N$ et $u_N(0) > f_N(0)$ $\Longrightarrow$ On ne s'arrête jamais quand $S_n = 0$ sauf si on est a la fin
  - Si $S_n = 1$ et $u_n(1) = f_n(1)$ (n'arrive que si $n \geq \alpha^*$) $\Longrightarrow$ On s'arrête.
  - α^* est caractérisé par :

$$\alpha^* = \max\left\{n \geq 1 : \frac{n}{N} \sum_{i=n}^{N-1} \frac{1}{i} \geq \frac{n}{N}\right\}$$

    et $\lim_{N \mapsto \infty} \alpha_N^*/N = e^{-1}$.
  - On s'arrête après α^*, la première fois que $S_n = 1$

Snell, James Laurie



- La théorie de l'arrêt optimal porte le nom d'enveloppe de Snell (voir [Snell(1952)]), "so named by the Russian mathematician, Kolmogorov".
- Snell (1925-2011), mathématicien américain, élève de Doob.
- Il utilise la théorie des martingales pour traiter le problème d'arrêt optimal.

Bibliographie

Gely P. Basharin, Amy N. Langville, and Valeriy A. Naumov.
The life and work of A. A. Markov.
Linear Algebra Appl., 386:3–26, 2004.

Nathanaël Berestycki.
Notes on card shuffling.
Technical report, http://www.statslab.cam.ac.uk/~beresty/Articles/shuffle.pdf, 2007.

P.L. Bernstein.
Capital Ideas: The Improbable Origins of Modern Wall Street.
The Free Press, 1992.

P.L. Bernstein.
Capital Ideas Evolving.
Jhon Wiley and Son, 2007.

F. Thomas Bruss.
The art of a right decision.
from Oxfo, page 15, 2006.

Jean-François Delmas and Benjamin Jourdain.
Modèles aléatoires.
Mathématiques & Applications. Springer-Verlag, Berlin, 2006.

Mark R. Dennis, Paul Glendinning, Paul A. Martin, Fadil Santosa, and Jared Tanner, editors.
The Princeton companion to applied mathematics.
Princeton University Press, Princeton, NJ, 2015.
p. 648–658.

Stuart Dreyfus.
Richard bellman on the birth of dynamic programming.
Operations Resarch, 50(1):48–51, 2002.

Hans Föllmer and Alexander Schied.
Convex and coherent risk measures.
http://www.alexschied.de/Encyclopedia6.pdf, 2008.

Exercice 7 On considère un dé à 6 faces équiprobables. Un joueur a le choix de lancer ce dé 1, 2 ou 3 fois, il obtient alors pour gain la valeur du dernier lancé effectué.

En supposant les 3 tirages indépendants, l'objectif de cet exercice est déterminer la stratégie optimale en moyenne de ce joueur.

1. Calculer l'espérance du gain si l'on effectue un seul tirage.
2. Si le joueur doit choisir un nombre de tirages fixe (déterministe), *avant le début du jeu*, a t'il intérêt à choisir entre jouer 1, 2 ou 3 fois ?
3. Quel est le gain optimal moyen du joueur, s'il connait les résultats des 3 tirages (le gain est alors le maximum des 3 tirages indépendants) ?

On suppose maintenant que le joueur peut tenir compte des tirages déjà effectués.

4. Montrer que la stratégie optimale (en moyenne), si l'on se restreint à 2 tirages maximum, est la suivante : si le premier tirage vaut 4, 5 ou 6 s'arrêter, sinon rejouer. Calculer la moyenne du gain de cette stratégie optimale.
5. En utilisant la question précédente identifier la stratégie optimale lorsque l'on joue 3 fois. Calculer la moyenne du gain de cette stratégie optimale.

Exercice 8 On considère une chaîne de Markov $(X_n, n \geq 0)$, sur un espace fini E, de matrice de transition P, issue d'un point $x_0 \in E$ (i.e. $\mathbb{P}(X_0 = x_0) = 1$).

On cherche $(u(n,x), n = 0, \ldots, N, x \in E)$ une solution à l'équation

$$\begin{cases} u(n,x) = \max \left\{ \sum_{y \in E} P(x,y) u(n+1,y), f(n,x) \right\}, n < N, x \in E \\ u(N,x) = f(N,x), x \in E. \end{cases} \tag{3.1}$$

Montrer qu'il existe une solution à cette équation (3.1) et que cette solution est unique.

Exercice 9 Avec les notation de l'exercice précèdent, on note u la solution unique de (3.1).

1. Montrer que pour tout temps d'arrêt τ, et pour tout n, il existe, un ensemble $\bar{A}_n \subset E^{n+1}$ tel que
$$\{\tau \geq n+1\} = \{\tau \leq n\}^c = \{X_{0:n} \in \bar{A}_n\}.$$
où $X_{0:n} = (X_0, \ldots, X_n)$.
2. Montrer que, quel que soit le temps d'arrêt τ plus petit que N, $\mathbb{E}\left(u\left(n \wedge \tau, X_{n \wedge \tau}\right)\right)$ est décroissant en n pour $0 \leq n \leq N$ (s'aider des transparents du cours au besoin).
3. En déduire que pour tout temps d'arrêt τ plus petit que N, $u(0,x_0) \geq \mathbb{E}\left(f\left(\tau, X_\tau\right)\right)$.
4. On note $\tau_0 = \inf\{n \geq 0, u(n,X_n) = f(n,X_n)\}$. Montrer que τ_0 est un temps d'arrêt plus petit que N.
5. Montrer que $\mathbb{E}\left(u\left(n \wedge \tau_0, X_{n \wedge \tau_0}\right)\right)$ ne dépend pas de n, pour $0 \leq n \leq N$ (s'aider des transparents du cours au besoin).
6. En déduire que $u(0,x_0) = \mathbb{E}\left(f\left(\tau_0, X_{\tau_0}\right)\right) = \sup_{0 \leq \tau \leq N, \text{t.a.}} \mathbb{E}\left(f\left(\tau, X_\tau\right)\right)$.

Exercice 10 On considère une chaîne de Markov $(X_n, n \geq 0)$ sur l'espace d'état $\{1,2\}$ de matrice de transition P $(0 < p < 1)$

$$P = \begin{pmatrix} p & 1-p \\ p & 1-p \end{pmatrix},$$

issue de 1 à l'instant 0 (i.e. $\mathbb{P}(X_0 = 1) = 1$).

1. Montrer que, si $n \geq 1$, la loi de X_n est donnée par $\mathbb{P}(X_n = 1) = p$ et $\mathbb{P}(X_n = 2) = 1 - p$.
2. En déduire que $(X_n, n \geq 1)$ est une suite de variables aléatoires indépendantes de même loi.
3. On s'intéresse au problème d'arrêt optimal pour cette chaîne de Markov avec comme gain à l'instant n, $f(n,x) = \rho^n \mathbf{1}_{\{x=1\}}$. On suppose que $\rho > 0$ et $p\rho > 1$. Montrer que la solution u de l'équation (3.1) est donnée, pour $n < N$, par $u(n, 1 \text{ ou } 2) = p\rho^N$ et pour $n = N$ par $u(N,x) = \rho^N \mathbf{1}_{\{x=1\}}$.
4. En déduire que le temps d'arrêt τ_{opt} optimal (unique dans ce cas) est donné par $\tau_{opt} = N$. Comment interpréter ce résultat ?
5. On pose $\bar{\tau}_{opt} = \sup\{n \geq 0, X_n = 1\}$. Vérifier que $\bar{\tau}_{opt}$ est l'unique temps aléatoire tel que
$$f(\bar{\tau}_{opt}, X_{\bar{\tau}_{opt}}) = \arg\max\{0 \leq n \leq N, f(n, X_n)\}.$$
6. Vérifier que $\mathbb{E}\left(f(\bar{\tau}_{opt}, X_{\bar{\tau}_{opt}})\right) > \mathbb{E}\left(f(\tau_{opt}, X_{\tau_{opt}})\right)$. $\bar{\tau}_{opt}$ peut-il être un temps d'arrêt ?

Chapitre 4

Le problème du vendeur de journaux

Le problème du vendeur de journaux

J.-P. Chancelier, B. Lapeyre

http://cermics.enpc.fr/~jpc/decision-incertain

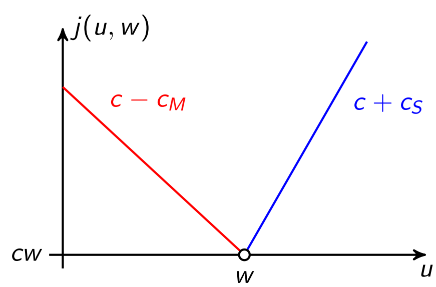
Plan

Le problème du vendeur de journaux (une période)

- Chaque matin, le vendeur doit décider d'un nombre de journaux à commander $u \in \mathbb{U} = \{0, 1, \dots\}$ au prix unitaire $c > 0$.
- La demande du jour est incertaine $w \in \mathbb{W} = \{0, 1, \dots\}$
  - Si, à la fin de la journée, il lui reste des invendus : coût unitaire $c_S \in \mathbb{R}$.
  $$c_S (u - w)_+ = c_s \underbrace{\max(u - w, 0)}_{\text{invendus}}, \qquad c > -c_S$$
  - Si, à la fin de la journée, il n'a pas pu faire face à la demande on associe un coût unitaire c_M. Le coût lié à la non satisfaction de la demande est
  $$c_M (w - u)_+ = c_M \underbrace{\max(w - u, 0)}_{\text{manquants}}$$

Face à une demande aléatoire, le coût est aléatoire

$$j(u, w) = \underbrace{cu}_{\text{achats}} + \underbrace{c_s (u - w)_+}_{\text{invendus}} + \underbrace{c_M (w - u)_+}_{\text{manquants}}$$



$\arg\min_{u \in \mathbb{U}} j(u, w) = \{w\}$: Quantité inconnue w !

Choix d'un critère en espérance

Le choix du critère : attitude face au risque du décideur. Comme le vendeur de journaux va vendre ses journaux chaque jour, on choisit ici un critère en espérance :

$$\min_{u \in \mathbb{U}} J(u) \quad \text{avec } J(u) = \underbrace{\mathbb{E}_w[j(u,w)]}_{\text{Espérance}}$$

Ce n'est pas le seul critère possible : Un vendeur de journaux pessimiste pourrait minimiser le pire des cas (utilise le support $\overline{\mathbb{W}}$ de la loi de la demande) :

$$\min_{u \in \mathbb{U}} J(u) \quad \text{avec } J(u) = \underbrace{\max_{w \in \overline{\mathbb{W}}} j(u,w)}_{\text{Pire des cas}}$$

Minimisation du coût en espérance

- On suppose que la demande, $\mathbf{W}$, est une variable aléatoire et le vendeur de journaux connaît, $\mathbb{P}_{\mathbf{W}}$, la loi de $\mathbf{W}$.
- Le coût à minimiser devient

$$J(u) = \mathbb{E}_{\mathbf{W}}[cu + c_s(u - \mathbf{W})_+ + c_M(\mathbf{W} - u)_+]$$

  fonction convexe pour $u \in \mathbb{R}_+$
- Le problème du vendeur de journaux devient

$$u^* \overset{?}{\in} \arg\min_{u \in \mathbb{U}} J(u)\,.$$

  Le vendeur de journaux prend une décision déterministe face à un aléa dont il connaît la loi mais pas la réalisation (Décision-Hasard).

Minimisation du coût en espérance

Commande optimale u^* :

- Si $c_M \leq c$ alors J est une fonction croissante et ne rien commander est optimal.
- Si $c_M > c$ alors J atteint son minimum en u^* :

  $$u^* = \inf\{z \in \mathbb{R} \mid F(z) \geq (c_M - c)/(c_M + c_S)\} \ ,$$

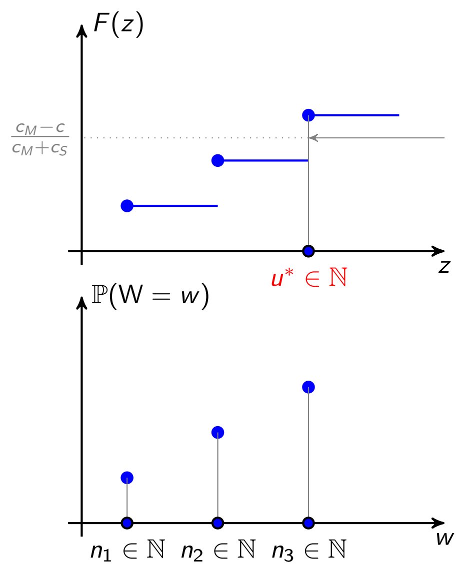
  où F est la fonction de répartition de W, $F(z) = \mathbb{P}(\mathrm{W} \leq z)$.
- Si W est une variable aléatoire entière alors $u^* \in NN$.

En effet :

$$\begin{aligned} J(u) &= c_M \mathbb{E}_\mathrm{W}[\mathrm{W}] + (c - c_M)u + (c_M + c_S)\mathbb{E}_\mathrm{W}[(u - \mathrm{W})_+] \\ &= c_M \mathbb{E}_\mathrm{W}[\mathrm{W}] + (c - c_M)u + (c_M + c_S)\int_0^u F(z)dz \end{aligned}$$

J est continue et coercive pour $u \in \mathbb{R}_+$. Pour $c_M > c$ elle décroît puis croît avec un minimum en u^*.

Minimisation du coût en espérance

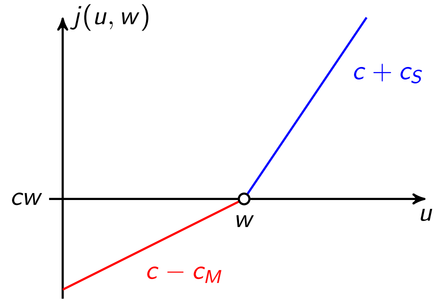


La commande optimale prend des valeurs entières

$$u^* = \inf\left\{z \in \mathbb{R} \,|\, F(z) \geq \frac{c_M - c}{c_M + c_S}\right\} \ ,$$

si la demande est une variable aléatoire entière.

Minimisation du coût en espérance $c_M \leq c$

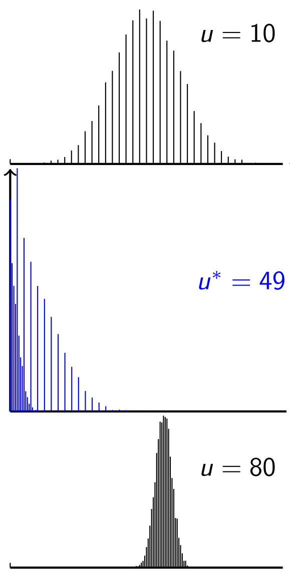


On remarque que

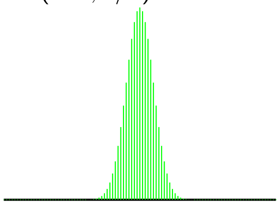
$$\min_{u\in\mathbb{U}} J(u) \geq \min_{u=\phi(\mathrm{W})} J(u) = \mathbb{E}_{\mathrm{W}}[\min_{u\in\mathbb{U}} j(u,\mathrm{W})]$$

Or $\min_{u\in\mathbb{U}} j(u,w)$ est atteint pour $u=0$ pour tout w et donne une commande admissible pour le problème $\min_{u\in\mathbb{U}} j(u,w)$.

Distribution des coûts $j(u, \mathrm{W})$



Distribution de la demande $\mathrm{bin}(100, 1/2)$.



Stock initial et coût fixe d'approvisionnement

On suppose ici que l'on peut avoir un stock initial $x \in \mathbb{Z}$ et que toute commande a un coût fixe c_F. Le coût à minimiser devient maintenant :

$$\begin{aligned}\widetilde{J}(u) &= \mathbb{E}_{\mathrm{W}}[c_F \mathbb{I}_{\{u>0\}} + cu + c_s(x+u-\mathrm{W})_+ + c_M(\mathrm{W}-x-u)_+] \\ &= c_F \mathbb{I}_{\{u>0\}} - cx + \mathbb{E}_{\mathrm{W}}[j(u+x,\mathrm{W})] \\ &= c_F \mathbb{I}_{\{u>0\}} + J(u+x) - cx\end{aligned}$$

Commande optimale $u^*(x)$ pour $c_M > c$:

La commande optimale est une stratégie (s, S) :

$$u^*(x) = (S-x)\mathbb{I}_{\{x \leq s\}}$$

Quand le stock devient en dessous de s, on commande des journaux de façon à ramener le stock à S. La valeur de S est donnée par $\arg\min_{u \in \mathbb{U}} J(u)$.

Stock initial et coût fixe d'approvisionnement

Soit S le stock donné par $\{S\} = \arg\min_{u \in \mathbb{U}} J(u)$.

- Si $x \geq S$, il est optimal de ne rien faire ($J(\cdot) \nearrow$ et $c_F \geq 0$) :

$$\tilde{J}(0) = J(x) - cx \leq J(x+u) - cx + c_F = \tilde{J}(u)$$

- Si $x \leq S$, comme $J(\cdot)$ est minimale en S :
  - Si on commande, il faut commander $u = S - x$ de coût

  $$\tilde{J}(u) = J(S) - cx + c_F$$

  - Si on ne commande rien le coût est $J(x) - cx$

  On doit choisir la solution la moins coûteuse, Il faut donc compléter son stock jusqu'à S si $J(x) \geq c_F + J(S)$ et ne rien faire sinon. De plus, on remarque que :

$$\{x \leq S \,|\, J(x) \geq c_F + J(S)\} = \{x \,|\, x \leq s\}$$

où s est donné par

$$\text{où} \quad s = \sup\{z \in (-\infty, S) \,|\, J(z) \geq c_F + J(S)\}$$

Stock initial et coût fixe d'approvisionnement

Posons $\mathrm{X}_0 = x$ et $f(x, u, w) = x + u - w$ et
$\tilde{j}(u, x_1) = c_F \mathbb{I}_{\{u>0\}} + cu + c_s(x_1)_+ + c_M(-x_1)_+$

Le vendeur de journaux : problème dynamique

$$\min_{\mathrm{u} \in \mathcal{U}, \mathrm{X_1}, \mathrm{X_0}} \mathbb{E}_{\mathrm{W}}[\tilde{j}(\mathrm{u}, \mathrm{X}_1)]$$

$$\text{sous contraintes}$$

$$\mathrm{X}_0 = x \quad \mathrm{X}_1 = f(\mathrm{X}_0, \mathrm{u}, \mathrm{W})$$

Avec une contrainte de non-anticipativité

$$\mathcal{U} = \{\mathrm{u} : \mathbb{X}_0 \to \mathbb{N}\}$$

La commande du nombre de journaux est contrainte à n'utiliser que l'information disponible $\mathrm{X}_0 = x$ et pas la demande future W.

Une "chaîne" de Markov

Les deux stocks de journaux X_0 et X_1 peuvent être vus comme les états d'une chaîne de Markov controlée. On peut autoriser des stocks négatifs et on a alors a priori $X_0 \in \mathbb{Z}$ et $X_1 \in \mathbb{Z}$.

- Supposant $u \in \mathbb{N}$ fixé, la matrice de transition de la chaîne de Markov est
$$P^u_{x_0,x_1} = \mathbb{P}(\mathrm{X}_1 = x_1 | \mathrm{X}_0 = x_0)$$
$$P^u_{x_0,x_1} = \begin{cases} \mathbb{P}(\mathrm{W} = w_0) & \text{si } x_1 = x_0 + u - w_0 \\ 0 & \text{sinon.} \end{cases}$$
- Supposant que u puisse être choisi comme une fonction de X_0, $u = \phi(\mathrm{X}_0)$, alors
$$P^\phi_{x_0,x_1} = \begin{cases} \mathbb{P}(\mathrm{W} = w_0) & \text{si } x_1 = x_0 + \phi(x_0) - w_0 \\ 0 & \text{sinon.} \end{cases}$$
- Étudier des problèmes sur N pas de temps sera l'objet du prochain cours.

Temps d'arrêt et problème de contrôle

Soit P la matrice de transition d'une chaine de Markov d'espace d'état $\mathbb{X}$. On considère $\mathbb{Y} = \mathbb{X} \cup \Delta$ et deux matrices de transition

$$\forall (x,y) \in \mathbb{X} \quad P^{C}_{x,y} = P_{x,y} \quad \text{et} \quad P^{C}_{\Delta,\Delta} = 1 \quad \text{(continue)}$$

et

$$\forall (x,y) \in \mathbb{X} \quad P^{S}_{x,y} = 0 \quad \text{et} \quad \forall x \in \mathbb{Y} \quad P^{S}_{x,\Delta} = 1 \quad \text{(stop)}$$

On cherche à évaluer un coût sur un pas de temps pour la chaîne partant de $X_0 = x \in \mathbb{X}$

$$v_0(x) = \sup_{u \in \{C,S\}, X_0 = x} \mathbb{E}(c(u, X_0) + f_1(X_1))$$

où la matrice de transition de la chaîne est P^u si le contrôle utilisé est u et avec $c(C, \cdot) = 0$ et $c(S,x) = f_0(x)$ et $f_1(\Delta) = 0$. On obtient :

$$v_0(x) = \sup \left(\sum_{y \in \mathbb{X}} P_{x,y} v_1(y), f_0(x) \right) \quad \text{et} \quad v_1(x) = f_1(x) \;, \quad \forall x \in \mathbb{X}$$

Similaire à l'Equation du temps d'arrêt optimal

Un programme linéaire

Le problème du vendeur de journaux

$$\min_{u \in \mathbb{U}} J(u) = \mathbb{E}_{\mathrm{W}}[j(u, \mathrm{W})]$$

avec

$$j(u,w) = c_M w + (c - c_M) u + (c_M + c_S)(u - w)_+$$

peut s'écrire comme un problème linéaire

Explicitons l'espérance

La demande W prend un nombre fini S de valeurs w_s, $s \in S$, avec probabilité π_s

$$\mathbb{P}(\mathrm{W} = s) = \pi_s \ , \ \ \forall s \in S$$

$$\sum_{s \in S} \pi_s = 1 \quad \text{et} \quad \pi_s \geq 0 \ , \ \ \forall s \in S$$

Le critère à minimiser s'écrit alors

$$J(u) = \sum_{s \in S} \pi_s \big(c_M w_s + (c - c_M) u + (c_M + c_S)(u - w_s)_+ \big)$$

Explicitons le max $(\cdot)_+$ avec des variables supplémentaires (Optim S1)

$$(u - w_s)_+ = \max\{0, u - w_s\} = \min \{v_s | vs \geq 0 \text{ et } v_s \geq u - w_s\}$$

$$\begin{aligned} J(u) &= \sum_{s \in S} \pi_s \big(c_M w_s + (c - c_M) u + (c_M + c_S)(u - w_s)_+ \big) \\ &= \sum_{s \in S} \pi_s \min_{\substack{v_s \geq \mathbf{0} \ , \ \ \forall s \in S \\ v_s \geq u - w_s \ , \ \ \forall s \in S}} c_M w_s + (c - c_M) u + (c_M + c_S) v_s \\ &= \min_{\substack{v_s \geq \mathbf{0} \ , \ \ \forall s \in S \\ v_s \geq u - w_s \ , \ \ \forall s \in S}} \sum_{s \in S} \pi_s \big(c_M w_s + (c - c_M) u + (c_M + c_S) v_s \big) \end{aligned}$$

Le problème du vendeur de journaux comme un programme linéaire

$$\min_{u \in \mathbb{U}, (v_s)_{s \in S} \in \mathbb{V}^S} \sum_{s \in S} \pi_s \big(c_M w_s + (c - c_M) u + (c_M + c_S) v_s \big)$$

$$\text{sous les contraintes}$$

$$v_s \geq u - w_s \quad \forall s \in S$$

$$v_s \geq 0 \quad \forall s \in S$$

$$u \geq 0$$

On obtient un programme linéaire en nombres entiers ($\mathbb{U} = \mathbb{N}$ et $\mathbb{V} = \mathbb{N}$) et on peut étudier la version relâchée ($\mathbb{U} = \mathbb{R}$ et $\mathbb{V} = \mathbb{R}$).

Question :

Est-ce que la version relâchée donne une solution entière ?

La contrainte de mesurabilité

Le fait que la commande $u \in \mathbb{U}$ doit être la même pour toutes les réalisations de w_s peut s'exprimer en introduisant de nouvelles commandes $u_s \in \mathbb{U}^S$ (duplication de variables) et en les contraignant a être toutes égales, $u_s = u \quad \forall s \in S$.

$$\min_{u \in \mathbb{U}, u_s \in \mathbb{U}^S, (v_s)_{s \in S} \in \mathbb{V}^S} \sum_{s \in S} \pi_s \big(c_M w_s + (c - c_M) u_s + (c_M + c_S) v_s \big)$$

$$\text{sous les contraintes}$$

$$v_s \geq u - w_s \quad \forall s \in S$$

$$v_s \geq 0 \quad \forall s \in S$$

$$u \geq 0$$

$$u_s = u \quad \forall s \in S$$

Si on dualise les contraintes, $u_s = u \quad \forall s \in S$, on découple les problèmes s par s. Des méthodes numériques sont basées sur cette idée.

POUR ALLER PLUS LOIN : COURS DE L'ÉCOLE DES PONTS

- 1A : cours d'« Optimisation » (matrices unimodulaires)
- 2A :
  - cours de « Recherche Opérationnelle »
  - cours d'« Optimisation et contrôle »
  - cours « Modéliser l'Aléa »

Exercice 11 1. On note F, la fonction de répartition d'une variable aléatoire réelle W, i.e. $F(z) = \mathbb{P}(W \leq z)$. Montrer que F est (toujours !) continue à droite. A quelle condition F est elle continue en un point z ?

2. Montrer que si u et w sont deux nombres réels positifs, on a $\int_0^u \mathbf{1}_{\{w \leq z\}} dz = (u - w)_+$. En déduire que, si W est une variable aléatoire positive :

$$\mathbb{E}\left((u - W)_+\right) = \int_0^u F(z) dz,$$

où $F(z) = \mathbb{P}(W \leq z)$ est la fonction de répartition de W.

Exercice 12 On suppose que c_M, c et c_S sont des nombres réels positifs et W une variable aléatoire positive. On définit, comme dans le cours, la perte moyenne résultant d'une commande en quantité u positive

$$J(u) = \mathbb{E}\left[c_F \mathbf{1}_{\{u>0\}} + cu + c_S(u - W)_+ + c_M(W - u)_+\right].$$

1. Vérifier que $J(u) = c_F \mathbf{1}_{\{u>0\}} + c_M \mathbb{E}(W) + (c - c_M)u + (c_M + c_S)\mathbb{E}\left((u - W)_+\right)$.
2. Montrer que $J(u) = c_M \mathbb{E}(W) + (c - c_M)u + (c_M + c_S) \int_0^u F(z) dz$ et que J est continue et admet une dérivée à droite en tout point.
3. On pose

$$u^* = \inf \left\{ z \in \mathbb{R}^+, F(z) \geq \frac{c_M - c}{c_M + c_S} \right\}.$$

   Montrez que, si W ne prend que des valeurs entières strictement positives, u^* prend lui aussi une valeur entière strictement positive.
4. Montrez que, si $c_M > c$, alors $u^* > 0$ et J est décroissante lorsque $u \in [0, u^*]$ et croissante si $u \geq u^*$ (elle atteint donc son minimum en u^*).
5. Montrer que si $c_M \leq c$ alors J est une fonction croissante sur $\mathbb{R}^+$ (elle atteint donc son minimum en 0).

Exercice 13 On s'intéresse maintenant à la fonction (dépendant de x) $\tilde{J}^x$, représentant la perte moyenne lorsque le stock initial est égal à x, définie par

$$\tilde{J}^x(u) = \mathbb{E}\left[c_F \mathbf{1}_{\{u>0\}} + cu + c_S(x + u - W)_+ + c_M(W - x - u)_+\right].$$

1. Vérifier que $\tilde{J}^x(u) = c_F \mathbf{1}_{\{u>0\}} + J(u + x) - cx$.

On suppose dans la suite que $c_M > c$. On note S pour l'unique point qui réalise $\mathrm{argmin}_{u \geq 0} J(u)$.

2. Lorsque $x \geq S$, vérifier que pour tout $u > 0$ $\tilde{J}^x(0) \leq \tilde{J}^x(u)$. *Il est optimal de ne rien commander.*
3. Lorsque $x \leq S$ et $J(x) < c_F + J(S)$, vérifier que pour tout $u > 0$ $\tilde{J}^x(0) \leq \tilde{J}^x(u)$. *Il est, là aussi, optimal de ne rien commander.*
4. Lorsque $x \leq S$ et $J(x) > c_F + J(S)$, vérifier que pour tout $u > 0$ $\tilde{J}^x(u) \geq \tilde{J}^x(S - x)$. *Il est optimal de commander la quantité $S - x$.*
5. On définit s par (notez que s est forcément inférieur à S par définition)

$$s = \sup \left\{ z \leq S, J(z) \geq c_F + J(S) \right\}.$$

   Vérifier que, pour $x \leq S$, $J(x) \geq c_F + J(S)$ est équivalent à $x \leq s$. En déduire que la commande optimale $u^*(x)$ partant d'un stock initial x est donnée par

$$u^*(x) = (S - x)\mathbf{1}_{\{x \leq s\}}.$$

Chapitre 5

Le problème du vendeur de journaux

Le problème du vendeur de journaux en horizon fini

J.-P. Chancelier, B. Lapeyre

`http://cermics.enpc.fr/~jpc/decision-incertain`

Plan

Programmation dynamique

Preuve (cas $L \equiv 0$)

Preuve $L \not\equiv 0$

Graphes

Rappel sur le problème à un pas de temps

Le vendeur de journaux : problème dynamique

$$\min_{u \in \mathcal{U}, X_1, X_0} \mathbb{E}_W \left[\tilde{\jmath}(u, X_1) \right]$$

sous contraintes

$$X_0 = x \quad X_1 = f(X_0, u, W)$$

$$f(x, u, w) = x + u - w$$
$$\tilde{\jmath}(u, x) = c_F \mathbb{I}_{\{u>0\}} + cu + \alpha \left(c_S(x)_+ + c_M(-x)_+ \right)$$

- Le « stock » $X_t \in \mathbb{Z}$ peut être positif (c'est alors un stock physique) ou négatif (c'est alors moins la quantité de manquants)
- La loi de la demande W, est supposée connue (de moyenne finie)
- Taux d'actualisation $\alpha \in (0, 1]$ (valoriser des gains futurs à l'instant présent)

Le problème à horizon fini T

Le problème sur un horizon fini

$$\min_{U \in \mathcal{U}, X} \mathbb{E}_W \left[\sum_{t=0}^{T-1} \alpha^t \tilde{\jmath}(U_t, X_{t+1}) \right]$$

sous contraintes

$$X_0 = x \quad \text{et} \quad \forall t \in [0, T-1], X_{t+1} = f(X_t, U_t, W_{t+1})$$

- La loi de la demande W_t, est supposée connue (de moyenne finie) pour tout t.
- On sera amené à supposer que les demandes $(W_1, W_2, \ldots, W_T)$ sont indépendantes.
- $\alpha \in (0, 1]$ est un taux d'actualisation (valoriser des gains futurs à l'instant présent).

Le problème à horizon T sous une forme canonique

Le problème sur un horizon fini

$$
\begin{aligned}
\min_{\mathsf{U}\in\mathcal{U},\mathsf{X}} \mathbb{E}_{\mathsf{W}} \Big[\sum_{t=0}^{T-1} \alpha^t c_t(\mathsf{U}_t,\mathsf{X}_t) + \alpha^T c_T(\mathsf{X}_T) \Big] \\
\text{sous contraintes} \\
\mathsf{X}_0 = x \quad \text{et} \quad \forall t \in [0, T-1]\,, \mathsf{X}_{t+1} = f(\mathsf{X}_t,\mathsf{U}_t,\mathsf{W}_{t+1})
\end{aligned}
$$

avec

$$
\begin{aligned}
c_t(u,x) &= c_F \mathbb{I}_{\{u>0\}} + cu + c_S(x)_+ + c_M(-x)_+ \\
c_0(u,x) &= c_F \mathbb{I}_{\{u>0\}} + cu \\
c_T(x) &= c_S(x)_+ + c_M(-x)_+
\end{aligned}
$$

Ré-écriture du coût

$$
\begin{aligned}
\tilde{\jmath}(u_0,x_1) + \alpha\tilde{\jmath}(u_1,x_2) &= \overbrace{c_F\mathbb{I}_{\{u_0>0\}} + cu_0 + \alpha\left(c_S(x_1)_+ + c_M(-x_1)_+\right)}^{\tilde{\jmath}(u_0,x_1)} \\
&+ \alpha\,\overbrace{\left(c_F\mathbb{I}_{\{u_1>0\}} + cu_1 + \alpha\left(c_S(x_2)_+ + c_M(-x_2)_+\right)\right)}^{\tilde{\jmath}(u_1,x_2)} \\
&= \overbrace{c_F\mathbb{I}_{\{u_0>0\}} + cu_0}^{c_0(u_0,x_0)} \\
&+ \alpha\,\overbrace{\left(c_F\mathbb{I}_{\{u_1>0\}} + cu_1 + c_S(x_1)_+ + c_M(-x_1)_+\right)}^{c_1(u_1,x_1)} \\
&+ \alpha^2\,\overbrace{\left(c_S(x_2)_+ + c_M(-x_2)_+\right)}^{c_T(x_2)} \\
&= c_0(u_0,x_0) + \alpha c_1(u_1,x_1) + \alpha^2 c_T(x_2)
\end{aligned}
$$

La contrainte de non anticipativité

- On suppose que le vendeur de journaux note chaque jour la valeur de la demande qui s'est réalisée et conserve cette valeur.
  - À l'instant t, il connaît $(\mathsf{W}_1, \cdots, \mathsf{W}_t)$ et X_0 et peut utiliser cette information pour calculer sa commande de journaux U_t
  - On pourrait donc choisir pour $\mathcal{U}$ les commandes de journaux qui à l'instant t sont fonction de $(\mathsf{X}_0, \mathsf{W}_1, \cdots, \mathsf{W}_t)$

$$\mathcal{U}_t = \{\phi : \mathbb{X} \times \mathbb{W}_1 \times \cdots \times \mathbb{W}_t \to \mathbb{N}\}$$

- Si les bruits sont indépendants et indépendants de X_0, on peut montrer que la stratégie optimale pour des fonctions de $(\mathsf{X}_0, \mathsf{W}_1, \cdots, \mathsf{W}_t)$ ne dépend à l'instant t que du stock X_t. La forme de la fonction coût à son importance dans ce résultat.

$$\mathcal{U}_t = \{\phi : \mathbb{X}_t \to \mathbb{N}\}$$

- Noter que, dans le cas des chaînes de Markov données par leur matrices de transition le, « bruit » n'est pas observé.
- On parle d'observation complète si on observe l'état de la chaîne de Markov.

Equation de Programmation Dynamique (Bellman)

On pose pour simplifier $L_t(x, u, w) = \alpha^t c_t(x, u)$ et $K = \alpha^T c_T$

Définition de $V_0(x)$

$$\begin{aligned} V_0(x) = \min_{\mathsf{x}, \mathsf{u} \in \mathcal{U}} \quad & \mathbb{E}\big(\sum_{s=0}^{T-1} L_s\left(\mathsf{x}_s, \mathsf{u}_s, \mathsf{w}_{s+1}\right) + K\left(\mathsf{x}_T\right)\big), \\ \text{s.c.} \quad & \mathsf{x}_0 = x \text{ fixé}, \\ & \mathsf{x}_{s+1} = f_s\left(\mathsf{x}_s, \mathsf{U}_s, \mathsf{w}_{s+1}\right), \qquad \forall s = 0, \ldots, T-1, \end{aligned}$$

Pour obtenir $V_0(x)$ on calcule la famille de fonctions valeurs $\{V_t\}_{t \in [0,T]}$

$$\begin{aligned} V_T &= K \\ V_t(x) &= \min_{u \in \mathbb{U}} \mathbb{E}\Big[L_t(x, u, \mathsf{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \mathsf{W}_{t+1})\big)\Big] \end{aligned}$$

Commande optimale

- Pour $t \in [0, T-1]$

$$V_t(x) = \min_{u \in \mathbb{U}} \mathbb{E}\Big[L_t(x, u, \mathrm{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \mathrm{W}_{t+1})\big)\Big]$$

$$\phi_t^\sharp(x) \in \arg\min_{u \in \mathbb{U}} \mathbb{E}\Big[L_t(x, u, \mathrm{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \mathrm{W}_{t+1})\big)\Big]$$

- A t fixé, autant de problème d'optimisation qu'il y a de x possibles
- Pour chaque x possible, on trouve une commande $u^\sharp$ optimale
- On construit ainsi une stratégie $\phi_t^\sharp : \mathbb{X} \to \mathbb{U}$
- La commande optimale pour le problème demarrant en $t = 0$ avec $X_0 = x$ est pour $t \in [0, T-1]$, $\mathrm{U}_t^\sharp = \phi_t^\sharp(\mathrm{X}_t^\sharp)$ avec

$$\begin{aligned} &\mathrm{x}_0^\sharp = x_0 \\ &\mathrm{x}_{s+1}^\sharp = f_s\left(\mathrm{x}_s^\sharp, \phi_t^\sharp(\mathrm{X}_t^\sharp), \mathrm{w}_{s+1}\right), \qquad \forall s = 0, \dots, T-1, \end{aligned}$$

Résumé

Sous une hypothèse d'indépendance des bruits $\{\mathrm{W}_t\}_{t \in [1,T]}$
l'équation de Programmation Dynamique $V_K = K$ et

$$V_t(x) = \min_{u \in \mathbb{U}} \mathbb{E}\Big[L_t(x, u, \mathrm{W}_{t+1}) + V_{t+1}\big(f_t(x, u, \mathrm{W}_{t+1})\big)\Big] \qquad \forall t \in [0, T-1]$$

donne la valeur du problème d'optimisation

$$\begin{aligned} V_0(x) = \min_{\mathrm{x}, \mathrm{u} \in \mathcal{U}} \quad & \mathbb{E}\Big(\sum_{s=0}^{T-1} L_s\left(\mathrm{x}_s, \mathrm{u}_s, \mathrm{w}_{s+1}\right) + K\left(\mathrm{x}_T\right)\Big), \\ \text{s.c.} \quad & \mathrm{x}_0 = x \text{ fixé}, \\ & \mathrm{x}_{s+1} = f_s\left(\mathrm{x}_s, \mathrm{U}_s, \mathrm{w}_{s+1}\right), \qquad \forall s = 0, \dots, T-1, \end{aligned}$$

et la commande optimale à chaque instant est une fonction de l'état

$$\phi_t^\sharp \in \mathcal{U}_t = \{\phi : \mathbb{X}_t \to \mathbb{N}\}$$

Difficultés de la programmation dynamique

- Difficultés mathématiques
  - Existence d'un minimum pour le problème initial
  - Existence de fonctions de Bellman au moins mesurables solution de l'Equation de Bellman et eventuellement plu srégulières.
  - Existence de stratégies optimales mesurable, problèmes de selections mesurables

    D. P. Bertsekas and S. E. Shreve
    Stochastic Optimal Control : The Discrete-Time Case (1996)
  - Le cas convexe s.c.i : Algorithme particulier SDDP qui utilise la convexité.
- Difficultés numériques, malédiction de la dimension. Si le cardinal de $\mathbb{X}$ est trop grand on ne sait plus résoudre
  - Tétris : espace d'état énorme 10^{61} pour hauteur 20 largeur 10 et 7 pièces.
  - Vallée hydraulique : problème quand on à plus de quatre stocks (barrages).
  - Gestion de portefeuille : problème quand on à plus de quatre actifs.

Preuve

On introduit $(\mathcal{P}_t)$ démarrant en x à l'instant t :

$$
\begin{aligned}
V_t(x) = \min_{\mathrm{x},\mathrm{u}\in\mathcal{U}} \quad & \mathbb{E}\Big(\sum_{s=t}^{T-1} L_s\left(\mathrm{x}_s, \mathrm{u}_s, \mathrm{w}_{s+1}\right) + \overline{K}\left(\mathrm{x}_T\right)\Big), \\
\text{s.c.} \quad & \mathrm{x}_t = x \text{ fixé}, \\
& \mathrm{x}_{s+1} = f_s\left(\mathrm{x}_s, \mathrm{U}_s, \mathrm{w}_{s+1}\right), \qquad \forall s = t, \dots, T-1,
\end{aligned}
$$

- V_t est appelé fonction valeur à l'instant t.
- V_T est connu $V_T = \overline{K}$
- On cherche V_0 la valeur du problème du vendeur de journaux $(\mathcal{P}_0)$
- On cherche la stratégie optimale du vendeur de journaux

On commence par le cas où $L_s \equiv 0$

On commence par un problème avec juste un coût final

$(\mathcal{P}_t)$ démarrant en x à l'instant t :

$$\begin{aligned} V_t(x) = \min_{\mathrm{x},\mathrm{u}\in\mathcal{U}} \quad & \mathbb{E}(K(\mathrm{x}_T)), \\ \text{s.c.} \quad & \mathrm{x}_t = x \text{ fixé}, \\ & \mathrm{x}_{s+1} = f_s(\mathrm{x}_s, \mathrm{U}_s, \mathrm{w}_{s+1}), \qquad \forall s = t, \dots, T-1, \end{aligned}$$

- Les bruits $\mathrm{w} = (\mathrm{w}_t)_{t=1,\dots,T}$.
- Les contrôles $\mathrm{u} = (\mathrm{u}_t)_{t=0,\dots,T-1}$.
- Les états $(\mathrm{x}_t)_{t=0,\dots,T-1}$.

On se limite aussi à

$$\mathcal{U}_t = \{\phi : \mathbb{X}_t \to \mathbb{N}\}$$

Dynamique Markovienne

Hypothèse Markovienne : bruits $\mathrm{x}_0, \mathrm{w}_1, \dots, \mathrm{w}_T$ indépendants et les w_t ont même loi (homogénéité).

- Matrice de transition : dynamique non contrôlée $f : \mathbb{X} \times \mathbb{W} \to \mathbb{X}$

$$\mathrm{X}_{t+1} = f(\mathrm{X}_t, \mathrm{W}_{t+1}), \quad P(x,y) = \mathbb{P}(f(x, \mathrm{w}_1) = y)$$

- Matrice de transition : dynamique contrôlée $f : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \to \mathbb{X}$ avec politique Markovienne $(\phi_s)_{s\in[0,T-1]}$, $\mathrm{U}_t = \phi_t(\mathrm{X}_t)$

$$\mathrm{X}_{t+1} = f(\mathrm{X}_t, \phi_t(\mathrm{X}_t), \mathrm{W}_{t+1}), \quad P_t(x,y) = \mathbb{P}(f(x, \phi_t(x), \mathrm{w}_1) = y)$$

  Pour $u \in \mathbb{U}$, soit $P^u = \mathbb{P}(f(x,u,\mathrm{w}_1) = y)$. Pour une politique $(\phi_s)_{s\in[0,T-1]}$ Markovienne, P_t^ϕ définie par $P_t^\phi(x,y) = P^{\phi_t(x)}(x,y)$

$$\mathrm{X}_{t+1} = f(\mathrm{X}_t, \phi_t(\mathrm{X}_t), \mathrm{W}_{t+1}), \quad P_t^\phi(x,y) = P^{\phi_t(x)}(x,y)$$

Deux familles de problèmes

$(\mathcal{P}_t)$ démarrant en x à l'instant t

$$V_t(x) = \min_{\mathrm{x},\phi} \quad V_t^{\phi}(x)$$
$$V_t^{\phi}(x) = \mathbb{E}(K(\mathrm{x}_T)),$$
$$\text{s.c. } \mathrm{x}_t = x \text{ fixé},$$
$$\mathrm{x}_{s+1} = f_s\left(\mathrm{x}_s, \phi_s(\mathrm{x}_s), \mathrm{w}_{s+1}\right), \ \forall s \in [t, T-1]$$

$(\mathcal{P}'_t)$ démarrant en μ (loi) à l'instant t

$$\mathcal{V}_t(\mu) = \min_{\mu_t,\dots,\mu_T,\phi(\cdot)} \quad \mathcal{V}_t^{\phi}(\mu)$$
$$\mathcal{V}_t^{\phi}(\mu) = \sum_x \mu_T(x)K(x)$$
$$\text{avec} \quad \mu_t = \mu, \quad \mu_{s+1} = \mu_s P_s^{\phi}\ , \ \forall s \in [t, T-1]$$

Liens entre les fonctions valeurs $\mathcal{V}_t^{\phi}(\cdot)$ et $V_t^{\phi}(\cdot)$

$\mathcal{V}_t^{\phi}(\mu)$ est la moyenne de la fonction V_t^{ϕ} pour une loi μ sur $\mathbb{X}$

$$\mathcal{V}_t^{\phi}(\mu) = \left\langle \mu,\ V_t^{\phi} \right\rangle \text{ avec } \langle \mu,\ F \rangle = \sum_{x \in \mathbb{X}} \mu(x)F(x)$$

$V_t^{\phi}(x)$ est la valeur de $\mathcal{V}_t^{\phi}(\mu)$ quand $\mu = \delta_x(\cdot)$

$$V_t^{\phi}(x) = \mathcal{V}_t^{\phi}(\delta_x(\cdot))$$

$\rightarrow$ Les mêmes relations sont vraies entre $\mathcal{V}_t$ et v_t.

En effet on obtient

- Problème $\mathcal{P}_t$ démarrant en x à t :

$$V_t^\phi(x) = (P_t^\phi \cdots P_{T-1}^\phi K)(x)$$
$$\text{(Par exemple } V_{T-1}^\phi(x) = \sum_{y\in\mathbb{X}} P_{T-1}^\phi(x,y)K(y))$$

- Problème $\mathcal{P}'_t$ démarrant en μ à t :

$$\mathcal{V}_t^\phi(\mu) = \mu P_t^\phi \cdots P_{T-1}^\phi K$$
$$\text{(Par exemple } \mathcal{V}_{T-1}^\phi(\mu) = \sum_{x,y\in\mathbb{X}} \mu(x) P_{T-1}^\phi(x,y)K(y))$$

- Les deux formules s'en déduisent

Lien entre V et $\mathcal{V}$ sur un pb d'horizon $T = 1$

$V(x)$

$$V(x) = \mathbb{E}_{X_0=x}\left[K(x_1)\right] = \sum_y P_{x,y}K(y)$$

$\mathcal{V}(\mu)$

$$\mathcal{V}(\mu) = \mathbb{E}_\mu\left[K(x_1)\right] = \sum_x \mu(x) \sum_y P_{x,y}K(y)$$

On obtient

$$\mathcal{V}(\mu) = \sum_y \mu(x)V(x) = \langle \mu,\ V\rangle\ ,$$

$$\mathcal{V}(\delta_{x'}) = \sum_x \delta_{x'}(x) \sum_y P_{x,y}K(y) = \sum_y P_{x',y}K(y) = V(x')\,.$$

Formule Récursive pour $\mathcal{V}_t$

On a la relation suivante :

$$\mathcal{V}_t(\mu) = \min_{\phi_t} \mathcal{V}_{t+1}(\mu P_t^{\phi})$$

Preuve : Le problème $\mathcal{P}_t'$ démarrant en μ à t :

$$\mathcal{V}_t(\mu) = \min_{\phi(\cdot)} \quad \mu P_t^{\phi} \cdots P_{T-1}^{\phi} K = \min_{\phi(\cdot)} \left\langle \mu P_t^{\phi},\ P_{t+1}^{\phi} \cdots P_{T-1}^{\phi} K \right\rangle$$

- À l'instant t, P_t^{ϕ} ne dépend que de la fonction ϕ_t.
- Le vecteur ligne μP_t^{ϕ} est à éléments positifs (loi de probabilité)

$$\begin{aligned}\mathcal{V}_t(\mu) &= \min_{\phi_t} \left\langle \mu P_t^{\phi},\ \min_{(\phi_s)_{s>t}} P_{t+1}^{\phi} \cdots P_{T-1}^{\phi} K \right\rangle \\ &= \min_{\phi_t} \left\langle \mu P_t^{\phi},\ V_{t+1}(\cdot) \right\rangle = \min_{\phi_t} \mathcal{V}_{t+1}(\mu P_t^{\phi})\end{aligned}$$

Formule Récursive pour V_t

Équation de Bellman

$$\begin{aligned} V_t(x) &= \min_{u \in \mathbb{U}} \mathbb{E}\left[V_{t+1}\left(f_t(x, u, \mathrm{W}_{t+1})\right)\right] \\ & \qquad u^{\sharp}(x) \in \arg\min_{u \in \mathbb{U}} \mathbb{E}\left[V_{t+1}\left(f_t(x, u, \mathrm{W}_{t+1})\right)\right] \\ V_T(x) &= K(x) \end{aligned}$$

Preuve : On a obtenu pour $\mathcal{V}_t$

$$\mathcal{V}_t(\mu) = \min_{\phi_t} \mathcal{V}_{t+1}(\mu P_t^{\phi})$$

- $V_t(x) = \mathcal{V}_t(\delta_x(\cdot))$
- $V_{t+1}\left(\delta_x(\cdot) P_t^{\phi}\right) = \mathbb{E}\left[V_{t+1}(f_1(x, \phi_t, \mathrm{W}_{t+1}))\right]$

Fin de la preuve

Écriture avec la matrice de transition de la chaîne de Markov

Équation de Bellman

$$
\begin{aligned}
V_t(x) &= \min_{u \in \mathbb{U}} \sum_y P^u_{x,y} V_{t+1}(y) \\
V_T(x) &= K(x) \\
u^\sharp(x) &\in \arg\min_{u \in \mathbb{U}} \sum_y P^u_{x,y} V_{t+1}(y)
\end{aligned}
$$

Preuve :

$$
\begin{aligned}
\sum_y P^u_{x,y} V_{t+1}(y) &= \sum_y \mathbb{P}(f(x, u, \mathrm{w}_{t+1}) = y) V_{t+1}(y) \\
&= \mathbb{E}\big[V_{t+1}(f_t(x, u, \mathrm{W}_{t+1}))\big]
\end{aligned}
$$

Horizon fini avec coût instantané

Le problème $(\mathcal{P}_0)$ démarrant en x à l'instant initial $t = 0$:

$$
\begin{aligned}
V_0(x) = \min_{\mathrm{x}, \mathrm{u} \in \mathcal{U}} \quad & \mathbb{E}\Big(\sum_{t=0}^{T-1} L_t\left(\mathrm{x}_t, \mathrm{u}_t, \mathrm{w}_{t+1}\right) + K\left(\mathrm{x}_T\right)\Big), \\
\text{s.c.} \quad & \mathrm{x}_0 = x \text{ fixé}, \\
& \mathrm{x}_{t+1} = f_t\left(\mathrm{x}_t, \mathrm{U}_t, \mathrm{w}_{t+1}\right), \qquad \forall t = 0, \ldots, T-1
\end{aligned}
$$

Équation de programmation dynamique

$$
V_t(x) = \min_{u \in \mathbb{U}} \mathbb{E}\left[L_t(x, u, \mathrm{W}_{t+1}) + V_{t+1}(f_t(x, u, \mathrm{W}_{t+1}))\right]
$$

Preuve

Le problème $(\mathcal{P}_0)$ démarrant en x à $t=0$ à un coût $\tilde{V}(0,x)$ où :

$$
\begin{aligned}
\tilde{V}_0(z,x) = \min_{\mathrm{x},\mathrm{u}\in\mathcal{U}} \quad & \mathbb{E}(Z_T + K(\mathrm{x}_T)),\\
\text{s.c.} \quad & \mathrm{x}_0 = x, \mathrm{z}_0 = z, \text{ fixées},\\
& \mathrm{x}_{t+1} = f_t(\mathrm{x}_t, \mathrm{U}_t, \mathrm{w}_{t+1}),\\
& \mathrm{z}_{t+1} = \mathrm{z}_t + L_t(\mathrm{x}_t, \mathrm{U}_t, \mathrm{w}_{t+1}), \qquad \forall t = 0,\dots,T-1,
\end{aligned}
$$

Le couple (Z,X) est Markovien pour des feedbacks $\phi_t(z,x)$.

Équation de Bellman

$$
\begin{aligned}
\tilde{V}_t(z,x) &= \min_{u\in\mathbb{U}} \mathbb{E}\left[\tilde{V}_{t+1}(z + L_t(x,u,\mathrm{W}_{t+1}), f_t(x,u,\mathrm{W}_{t+1}))\right]\\
\tilde{V}_T(z,x) &= z + K(x)
\end{aligned}
$$

Preuve

- On montre que $\tilde{V}_t(z,x) = z + V_t(x)$
- C'est vrai pour $t = T$, en effet $\tilde{V}_T(z,x) = z + K(x)$. Puis :

$$
\begin{aligned}
\tilde{V}_t(z,x) &= \min_{u\in\mathbb{U}} \mathbb{E}\left[z + L_t(x,u,\mathrm{W}_{t+1}) + V_{t+1}(f_t(x,u,\mathrm{W}_{t+1}))\right]\\
\tilde{V}_t(z,x) &= z + \underbrace{\min_{u\in\mathbb{U}} \mathbb{E}\left[L_t(x,u,\mathrm{W}_{t+1}) + V_{t+1}(f_t(x,u,\mathrm{W}_{t+1}))\right]}_{V_t(x)}
\end{aligned}
$$

- On remarque alors que le min en $u \in \mathbb{U}$ ne dépend que de x. Le contrôle optimal est en feedback seulement sur x.
- On note aussi que $\tilde{V}_t(0,x) = V_t(x)$, ce qui donne l'équation de Bellman pour le problème avec coût instantané (V).

Plus court chemin dans un graphe

Si le chemin $A \to B \to C \to D \to E$ est optimal alors $C \to D \to E$ est optimal.

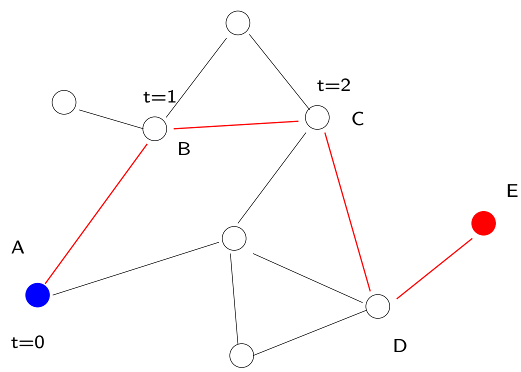


Figure – Plus court chemin

Principe de programmation dynamique

- Plus court chemin pour toutes les paires (i, j) du graphes ($w_{i,j}$ poids de l'arc (i, j)).
- $d_{ij}^{(m)}$: valeur du plus court chemin allant de i à j avec au plus m arcs.
- $d_{i,j}^{(0)} = 0$ si $i = j$ et $+\infty$ sinon.
- Principe de programmation dynamique

$$d_{ij}^{(m)} = \min\left(d_{ij}^{(m-1)}, \min_{1\leq k\leq n, k\neq j}(d_{ik}^{(m-1)} + w_{kj})\right) \\ = \min_{1\leq k\leq n}(d_{ik}^{(m-1)} + w_{kj})$$

- Si pas de cycles de poids négatif alors $d^{(n-1)}$ ou n est le nombre de nœuds du graphe donne la solution du problème.

Quelques éléments toujours présents

- Un problème de départ : trouver $d^{(n-1)}$ est remplacé par une famille de problèmes ($d^{(m)}$, $m = 0, \ldots, n-1$).
- Le problème de départ est bien sûr l'un des problèmes.
- On établit une relation de récurrence entre les $d^{(m)}$.
- Dans la récurrence apparaît un problème d'optimisation local.
- On a remplacé le calcul du plus court chemin par le calcul de la valeur du plus court chemin.
- Le plus court chemin se retrouve en gardant en mémoire l'argmin des problèmes récursifs.

Retour sur le plus court chemin

- Fonction valeur :

$$d_{ij}^{(m)} = \min_{1 \le k \le n} (d_{ik}^{(m-1)} + w_{kj})$$

- Comment aller de i à j par le plus court chemin. On regarde l'argmin du problème

$$d_{ij}^{(n-2)} = \min_{1 \le k \le n} (d_{ik}^{(n-1)} + w_{kj})$$

  Si cet argmin vaut $k^\sharp$ cela veut dire que le plus court chemin $i \rightsquigarrow j = i \rightsquigarrow k^\sharp \to j$
- On regarde alors comment aller de i à $k^\sharp$ en au plus $n-2$ arcs.
- Noter que $d^{(n-1)} = (W)^{(n-1)}$ (dans l'algèbre $(\max, +)$) : carré itérés.

Récursion

- $Fib(n) = (n <= 1)?1 : Fib(n-1) + Fib(n-2)$.

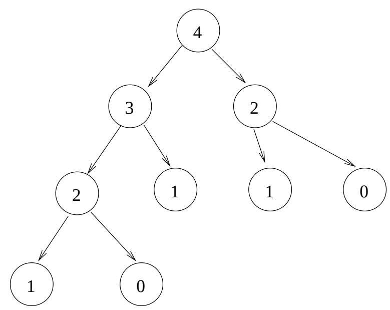


Figure – Fibonacci

Récursion

- Complexité exponentielle si l'on y prend garde !
- Solution 1 : partir de $(Fib(0), Fib(1))$ et calculer itérativement (plus de récursion).
- Solution 2 : utiliser la récursion mais garder en mémoire les valeurs déjà calculées (fonction à mémoire ou memoization (en anglais)).

POUR ALLER PLUS LOIN : COURS DE L'ÉCOLE DES PONTS

- cours de « Recherche Opérationnelle »
- cours d'« Optimisation et contrôle »
- cours « Modéliser l'Aléa »
- cours « Finance : aspects mathématiques et numériques ».

Cours de Décision dans l'incertain
Exercices : Jeudi 16 mars 2023

Exercice 14 On se donne une suite de fonctions (ou contrôles) $\tilde{u} = (\tilde{u}(n, x_{0:n}), n \geq 0)$ de E^n à valeurs dans un ensemble A ("on choisit la contrôle $\tilde{u}$ en tenant compte de toute la trajectoire avant l'instant n").

$(W_n, n \geq 1)$ une suite i.i.d. à valeurs dans F et x_0 un point de E. Les espaces E, F et A sont supposés finis.

On pose $X_0 = x_0$, puis itérativement on définit $X_{0:n} = (X_0, \ldots, X_n)$, $U_n = \tilde{u}(n, X_{0:n})$ puis X_{n+1} par

$$X_{n+1} = f(X_n, U_n, W_{n+1}). \tag{5.1}$$

A ce système dynamique contrôlé (5.1) on associe une famille $(P_u(x,y), x, y \in E, u \in A)$ de matrice de transition, en posant

$$P_u(x,y) = \mathbb{P}(f(x, W_1, u) = y)$$

1. Vérifier que pour un u fixé, $P_u(x,y)$ est une matrice de transition d'un chaine de Markov.
2. Montrer que, pour toute fonction v de E dans $\mathbb{R}$, $\sum_{y \in E} P_u(x,y) v(y) = \mathbb{E}[v(f(x,u,W_1)]$.
3. Montrer que $\mathbb{P}(X_{0:n+1} = x_{0:n+1}) = \mathbb{P}(X_{0:n} = x_{0:n}) P_{\tilde{u}(n, x_{0:n})}(x_n, x_{n+1})$.

Exercice 15 Soit $c(n,x,u)$ une fonction de $\mathbb{N} \times E \times A$ dans $\mathbb{R}$ et $K(x)$ un fonction de E dans $\mathbb{R}$. On considère l'équation de (Hamilton-Jacobi-)Bellman suivante

$$\begin{cases} v(n,x) = \min_{u \in A} \left\{ \sum_{y \in E} P_u(x,y) v(n+1,y) + c(n,x,u) \right\},\ n < N, \\ v(N,x) = K(x). \end{cases} \tag{HJB}$$

Montrer que l'équation précédente définit une *unique* fonction v. Ecrire un algorithme calculant cette fonction v.

Exercice 16 On spécifie le coût d'une stratégie $\tilde{u}$, en suppose que

— le choix de $U_j = \tilde{u}(j, X_{0:j})$ à l'instant j, coûte $c(j, X_j, U_j)$,

— le coût à l'instant final N est donné par $K(X_N)$.

Le coût moyen d'une stratégie $\tilde{u}$ est alors défini par

$$J(\tilde{u}) = \mathbb{E}\left(\sum_{j=0}^{N-1} c(j, X_j, U_j) + K(X_N) \right).$$

On cherche à identifier une stratégie qui minimise $J(\tilde{u})$, parmi tous les contrôles possibles $\tilde{u}$.

Pour cela on considère $\hat{u}(n,x)$ une fonction (pas forcément unique) à valeurs dans A qui réalise le minimum en u présent dans l'équation (HJB) à n et x fixés. On note $\hat{X}$ le processus définit par réccurence à partir de $\hat{u}$: $X_0 = x_0$, puis

$$\hat{X}_{n+1} = f(\hat{X}_n, \hat{U}_n, W_{n+1}),\ \text{où } \hat{U}_n = \hat{u}(n, \hat{X}_n).$$

Le but est de montrer que $\hat{u}$ réalise le minimum de de J.

1. Vérifier que l'on peut définir une fonction $\hat{u}$ qui réalise le minimum en u présent dans l'équation (HJB) mais que cette fonction n'est pas forcément unique.
2. Montrer, en utilisant la question 3 du premier exercice, que

$$\mathbb{E}\left(v(n+1, X_{n+1})\right) = \sum_{x_{0:n} \in E^{n+1}} \mathbb{P}(X_{0:n} = x_{0:n}) \sum_{x_{n+1} \in E} P_{\tilde{u}(n, x_{0:n})}(x_n, x_{n+1}) v(n+1, x_{n+1}),$$

puis, en utilisant l'équation (HJB), en déduire que

$$\mathbb{E}(v(n, X_n) - c(n, X_n, U_n)) \leq \mathbb{E}(v(n+1, X_{n+1})). \tag{5.2}$$

3. En se souvenant que $\hat{u}(n, x)$ réalise le minimum présent dans l'équation (HJB), montrer que

$$\mathbb{E}(v(n, \hat{X}_n) - c(n, \hat{X}_n, \hat{u}(n, \hat{X}_n))) = \mathbb{E}(v(n+1, \hat{X}_{n+1})). \tag{5.3}$$

4. Montrer par récurrence, en utilisant (5.2) et $v(N, x) = K(x)$, que :

$$v(0, x_0) \leq \mathbb{E}\left(\sum_{j=0}^{N-1} c(j, X_j, U_j) + K(X_N)\right) = J_{\tilde{u}},$$

5. Montrer à partir de (5.3) que $v(0, x_0) = \mathbb{E}\left(\sum_{j=0}^{N-1} c(j, \hat{X}_j, \hat{u}(j, \hat{X}_j)) + K(\hat{X}_N)\right) = J_{\hat{u}}$.
6. En déduire l'optimalité de $\hat{u}$: $\min_{\tilde{u}} J(\tilde{u}) = J(\hat{u})$.
   On constate que le contrôle $\hat{u}$ est optimal parmi tous les contrôles mais est de type "feedback" (il ne depend de la trajectoire avant n que par sa dernière valeur X_n).