Décisions dans l'incertain

Le problème du vendeur de journaux en horizon fini

Jean-Philippe Chancelier

Mai 2020

Rappel sur le problème à un pas de temps

Le vendeur de journaux : problème dynamique

$$\min_{u \in \mathcal{U}, \boldsymbol{X}_1, \boldsymbol{X}_0} \mathbb{E}_{\boldsymbol{W}}\left[\widetilde{\jmath}(u, \boldsymbol{X}_1)\right]$$

sous contraintes

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

avec

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

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

LE PROBLÈME À HORIZON T

Le problème sur un horizon fini

$$\min_{\boldsymbol{U}\in\mathcal{U},\boldsymbol{X}} \mathbb{E}_{\boldsymbol{W}}\Big[\sum_{t=0}^{T-1} \alpha^t \tilde{\jmath}(\boldsymbol{U}_t, \boldsymbol{X}_{t+1})\Big]$$

sous contraintes

$$\boldsymbol{X}_0 = x \quad \boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1})$$

- La loi de la demande $\boldsymbol{W}_t$, est supposée connue (de moyenne finie) pour tout t.
- On sera amené à supposer que les demandes $(\boldsymbol{W}_1, \boldsymbol{W}_2, \dots, \boldsymbol{W}_T)$ 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

$$\min_{\boldsymbol{U}\in\mathcal{U},\boldsymbol{X}} \mathbb{E}_W\Big[\sum_{t=0}^{T-1} \alpha^t c_t(\boldsymbol{U}_t, \boldsymbol{X}_t) + \alpha^T K(\boldsymbol{X}_T)\Big]$$

sous contraintes

$$\boldsymbol{X}_0 = x \quad \boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1})$$

avec

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

CALCUL DU COÛT

$$\begin{aligned}
\widetilde{J}(u_0, x_1) + \alpha \widetilde{J}(u_1, x_2) &= \overbrace{c_F \mathbb{I}_{\{u_0>0\}} + c u_0 + \alpha \left(c_s (x_1)_+ + c_M(-x_1)_+\right)}^{\widetilde{J}(u_0,x_1)} \\
&+ \alpha \overbrace{\left(c_F \mathbb{I}_{\{u_1>0\}} + c u_1 + \alpha \left(c_s (x_2)_+ + c_M(-x_2)_+\right)\right)}^{\widetilde{J}(u_1,x_2)} \\
&= \overbrace{c_F \mathbb{I}_{\{u_0>0\}} + c u_0}^{c_0(u_0,x_0)} \\
&+ \alpha \overbrace{\left(c_F \mathbb{I}_{\{u_1>0\}} + c u_1 + c_s (x_1)_+ + c_M(-x_1)_+\right)}^{c_1(u_1,x_1)} \\
&+ \alpha^2 \overbrace{\left(c_s (x_2)_+ + c_M(-x_2)_+\right)}^{K(x_2)} \\
&= c_0(u_0, x_0) + \alpha c_1(u_1, x_1) + \alpha^2 K(x_2)
\end{aligned}$$

La contrainte de non anticipativité

- On suppose que le vendeur de journaux chaque jours note la valeur de la demande qui s'est réalisée et conserve cette valeur.
  - À l'instant t, il connaît $(\boldsymbol{W}_1, \cdots, \boldsymbol{W}_t)$ et $\boldsymbol{X}_0$ et peut utiliser cette information pour calculer sa commande de journaux $\boldsymbol{U}_t$.
  - On pourrait donc choisir pour $\mathcal{U}$ les commandes de journaux qui à l'instant t sont fonction de $(\boldsymbol{X}_0, \boldsymbol{W}_1, \cdots, \boldsymbol{W}_t)$.
- Si les bruits sont indépendants et indépendants de $\boldsymbol{X}_0$, on peut montrer que la stratégie optimale pour des fonctions de $(\boldsymbol{X}_0, \boldsymbol{W}_1, \cdots, \boldsymbol{W}_t)$ ne dépend à l'instant t que du stock $\boldsymbol{X}_t$. La forme de la fonction coût à son importance dans ce résultat.
- 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.

On commence par un problème avec juste un coût final

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

$$
\begin{aligned}
V_0(x) = \min_{\boldsymbol{X},\boldsymbol{U}\in\mathcal{U}} \quad & \mathbb{E}\left[K\left(\boldsymbol{X}_T\right)\right], \\
\text{s.c.} \quad & \boldsymbol{X}_0 = x \text{ fixé}, \\
& \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \qquad \forall t = 0, \ldots, T-1,
\end{aligned}
$$

- Les bruits $\boldsymbol{W} = (\boldsymbol{W}_t)_{t=1,\ldots,T}$.
- Les contrôles $\boldsymbol{U} = (\boldsymbol{U}_t)_{t=0,\ldots,T-1}$.
- Les états $(\boldsymbol{X}_t)_{t=0,\ldots,T-1}$.

Dynamique Markovienne

Hypothèses Markoviennes : bruits $\boldsymbol{X}_0, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_T$ indépendants et les $\boldsymbol{W}_t$ ont même loi.

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

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

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

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

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

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

Une famille de problèmes

- Problème $(\mathcal{P}_{t_0})$ démarrant en x à t_0 :

$$
\begin{aligned}
V_{t_0}(x) = \min_{\boldsymbol{X},\phi(\cdot)} \quad & V^{\phi}_{t_0}(x) \\
& V^{\phi}_{t_0}(x) = \mathbb{E}\left[K\left(\boldsymbol{X}_T\right)\right], \\
& \text{avec} \quad \boldsymbol{X}_{t_0} = x, \quad \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \phi_t(\boldsymbol{X}_t), \boldsymbol{W}_{t+1}\right)
\end{aligned}
$$

- Problème $(\mathcal{P}'_{t_0})$ démarrant en μ à t_0 :

$$
\begin{aligned}
\mathcal{V}_{t_0}(\mu) = \min_{\mu,\phi(\cdot)} \quad & \mathcal{V}^{\phi}_{t_0}(\mu) \\
& \mathcal{V}^{\phi}_{t_0}(\mu) = \sum_x \mu_T(x) K(x) \\
& \text{avec} \quad \mu_{t_0} = \mu, \quad \mu_{t+1} = \mu_t P^{\phi}_t
\end{aligned}
$$

- Dynamique $\mu_{t+1}(y) = \sum_x \mu_t(x) P^{\phi(x)}_{x,y}$

Liens entre les fonctions valeurs $\mathcal{V}^{\phi}_{t_0}(\cdot)$ et $V^{\phi}_{t_0}(\cdot)$

On a :

$$\mathcal{V}^{\phi}_{t_0}(\mu) = \left\langle \mu, V^{\phi}_{t_0} \right\rangle := \sum_x \mu(x) V^{\phi}_{t_0}(x)\,, \text{et} \quad V^{\phi}_{t_0}(x) = \mathcal{V}^{\phi}_{t_0}(\delta_x(\cdot))$$

En effet :

- Problème $(\mathcal{P}_{t_0})$ démarrant en x à t_0 :

  $V^{\phi}_{t}(x) = (P^{\phi}_{t} \cdots P^{\phi}_{T-1} K)(x)$

  (Par exemple pour $t = T-1 \quad V^{\phi}_{T-1}(x) = \sum_{y \in \mathbb{X}} P^{\phi}_{T-1}(x,y) K(y))$

- Problème $(\mathcal{P}'_{t_0})$ démarrant en μ à t :

  $\mathcal{V}^{\phi}_{t}(\mu) = \mu P^{\phi}_{t} \cdots P^{\phi}_{T-1} K$

  (Par exemple pour $t = T-1 \quad \mathcal{V}^{\phi}_{T-1}(\mu) = \sum_{x,y \in \mathbb{X}} \mu(x) P^{\phi}_{T-1}(x,y) K(y))$

LE CAS $T = 1$

$X_0 = x$

$$V(x) = \mathbb{E}\left[K(\boldsymbol{X}_1)\right] = \sum_y P_{x,y} K(y)$$

la loi de $\boldsymbol{X}_0$ est μ

$$\mathcal{V}(\mu) = \mathbb{E}\left[K(\boldsymbol{X}_1)\right] = \sum_x \mu(x) \sum_y P_{x,y} K(y)$$

On obtient

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

et

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

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 :

$$\mathcal{V}_t(\mu) = \min_{\phi(\cdot)} \quad \mu P_t^{\phi} \cdots P_{T-1}^{\phi} K$$

- À l'instant t, P_t^{ϕ} ne dépend que de la fonction ϕ_t.
- Le vecteur ligne $\mu P_{t_0}^{\phi}$ est à éléments positifs (c'est une loi de probabilité)

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

$$\mathcal{V}_t(\mu) = \min_{\phi_t} \left\langle \mu P_t^{\phi}, V_{t+1}(\cdot) \right\rangle = \min_{\phi_t} \mathcal{V}_{t+1}(\mu P_t^{\phi})$$

Formule récursive pour V_t avec $t \in \{0, \cdots, T\}$

On a la relation suivante (Équation de Bellman) :

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

$$u^\sharp(x) \in \operatorname*{Argmin}_{u \in \mathbb{U}} \mathbb{E}\Big(V_{t+1}(f_t(x, u, \mathbf{W}_{t+1}))\Big)$$

$$V_T(x) = K(x)$$

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

$$\mathcal{V}_t(\mu) = \min_{\phi_t} \mathcal{V}_{t_0+1}(\mu P_t^\phi)$$

- $V_t(x) = \mathcal{V}_t(\delta_x(\cdot))$
- $V_{t_0+1}(\delta_x(\cdot) P_t^\phi) = \mathbb{E}\Big(V_{t+1}(f_1(x, \phi_t, \mathbf{W}_{t+1}))\Big)$

Écriture avec la matrice de transition de la chaîne de Markov

Équation de Bellman

$$V_t(x) = \min_{u \in \mathbb{U}} \sum_y P^u_{x,y} V_{t+1}(y)$$
$$V_T(x) = K(x)$$
$$u^\sharp(x) \in \operatorname*{Argmin}_{u \in \mathbb{U}} \sum_y P^u_{x,y} V_{t+1}(y)$$

Preuve :

$$\sum_y P^u_{x,y} V_{t+1}(y) = \sum_y \mathbb{P}\big(f(x,u,\mathbf{W}_{t+1}) = y\big) V_{t+1}(y)$$
$$= \mathbb{E}\Big(V_{t+1}\big(f_t(x,u,\mathbf{W}_{t+1})\big)\Big)$$

Horizon fini avec coût instantané

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

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

Équation de programmation dynamique

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

PREUVE

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

$$
\begin{aligned}
\widetilde{V}_0(z, x) = \min_{\boldsymbol{X}, \boldsymbol{U} \in \mathcal{U}} \quad & \mathbb{E}\left[Z_T + K\left(\boldsymbol{X}_T\right)\right], \\
\text{s.c.} \quad & \boldsymbol{X}_0 = x, \boldsymbol{Z}_0 = z, \text{ fixées}, \\
& \boldsymbol{X}_{t+1} = f_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \\
& \boldsymbol{Z}_{t+1} = \boldsymbol{Z}_t + L_t\left(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\right), \qquad \forall t = 0, \ldots, T-1,
\end{aligned}
$$

Le couple $(\boldsymbol{Z}, \boldsymbol{X})$ est markovien pour des feedbacks : $\phi_t(z, x)$.

Équation de Bellman

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

Preuve

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

$$\widetilde{V}_t(z,x) = \min_{u \in \mathbb{U}} \mathbb{E}\Big(z + L_t(x,u,\boldsymbol{W}_{t+1}) + V_{t+1}\big(f_t(x,u,\boldsymbol{W}_{t+1})\big)\Big)$$

$$\widetilde{V}_t(z,x) = z + \underbrace{\min_{u \in \mathbb{U}} \mathbb{E}\Big(L_t(x,u,\boldsymbol{W}_{t+1}) + V_{t+1}\big(f_t(x,u,\boldsymbol{W}_{t+1})\big)\Big)}_{V_t(x)}$$

- 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 $\widetilde{V}_t(0,x) = V_t(x)$, 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.

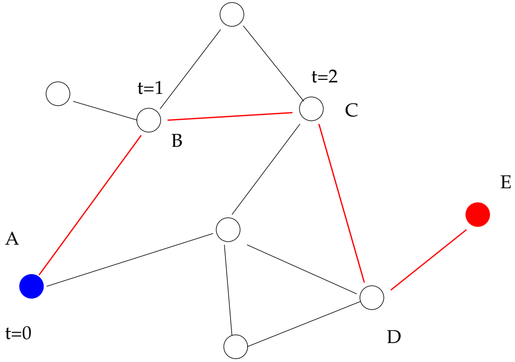

Figure – Plus court chemin

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

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

- Si pas de cycles de poids négatif alors $d^{(n-1)}$ ou n est le nombre de noeuds 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 $(d^{(m)}, m = 0, \dots, n-1)$.
- 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 :

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

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

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

  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

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

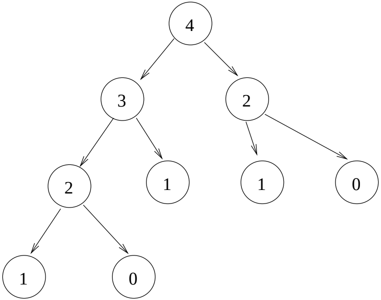

FIGURE – Fibonacci

Récursion

- Complexité exponentielle si l'on y prends 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 »