

# Décision dans l'incertain

## Le problème du vendeur de journaux en horizon fini

Jean-Philippe CHANCELIER

10 mars 2022





# LE PROBLÈME À HORIZON $T$

## Le problème sur un horizon fini

$$\min_{\mathbf{u} \in \mathcal{U}, \mathbf{X}} \mathbb{E}_{\mathbf{W}} \left[ \sum_{t=0}^{T-1} \alpha^t \tilde{\mathcal{J}}(\mathbf{u}_t, \mathbf{X}_{t+1}) \right]$$

sous contraintes

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

- ▶ La loi de la demande  $\mathbf{W}_t$ , est supposée connue (de moyenne finie) pour tout  $t$ .
- ▶ On sera amené à supposer que les demandes  $(\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_T)$  sont **indépendantes**.
- ▶  $\alpha \in (0, 1]$  est un taux d'actualisation (valoriser des gains futurs à l'instant présent).



# CALCUL DU COÛT

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

# LA CONTRAINTE DE NON ANTICIPATIVITÉ

- ▶ On suppose que le vendeur de journaux note chaque jour la valeur de la demande qui s'est réalisée et conserve cette valeur.
  - ▶ À l'instant  $t$ , il connaît  $(W_1, \dots, W_t)$  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  $(X_0, W_1, \dots, W_t)$ .
- ▶ Si les bruits sont indépendants et indépendants de  $X_0$ , on peut montrer que la stratégie optimale pour des fonctions de  $(X_0, W_1, \dots, W_t)$  ne dépend à l'instant  $t$  que du stock  $X_t$ . La forme de la fonction coût à son importance dans ce résultat.
- ▶ 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.

# PROBLÈME D'OPTIMISATION À L'INSTANT 0

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

$$V_0(x) = \min_{\mathbf{x}, \mathbf{u} \in \mathcal{U}} \mathbb{E} \left[ \sum_{s=0}^{T-1} L_s(\mathbf{X}_s, \mathbf{u}_s, \mathbf{w}_{s+1}) + \bar{K}(\mathbf{X}_T) \right],$$

s.c.  $\mathbf{X}_0 = x$  fixé,

$$\mathbf{X}_{s+1} = f_s(\mathbf{X}_s, \mathbf{u}_s, \mathbf{w}_{s+1}), \quad \forall s = 0, \dots, T-1,$$

avec

- ▶  $L_t(x, u, w) = \alpha^t c_t(x, u)$
- ▶  $\bar{K} = \alpha^T K$



# L'ÉQUATION DE BELLMAN

Équation de programmation dynamique

$$V_T = \bar{K}$$

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

- ▶ Il faut résoudre autant de problème d'optimisation qu'il y a de  $x$  possibles
- ▶ Pour chaque  $x$  possible, on trouve un  $u^\#$  solution
- ▶ On construit ainsi une stratégie

$$\lambda_t : \mathbb{X} \rightarrow \mathbb{U}$$

$$x \mapsto u^\#(x)$$

# ON COMMENCE PAR UN PROBLÈME AVEC JUSTE UN COÛT FINAL

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

$$V_0(x) = \min_{\mathbf{X}, \mathbf{U} \in \mathcal{U}} \mathbb{E} [K(\mathbf{X}_T)] ,$$

$$\text{s.c. } \mathbf{X}_0 = x \text{ fixé,}$$

$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) , \quad \forall t = 0, \dots, T-1,$$

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

# DYNAMIQUE MARKOVIENNE

Hypothèse Markovienne : bruits  $\mathbf{X}_0, \mathbf{W}_1, \dots, \mathbf{W}_T$  indépendants et les  $\mathbf{W}_t$  ont même loi (homogénéité).

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

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

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

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

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

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

## UNE FAMILLE DE PROBLÈMES

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

$$V_{t_0}(x) = \min_{\mathbf{X}_{t_0}, \dots, \mathbf{X}_T, \phi(\cdot)} V_{t_0}^\phi(x)$$

$$V_{t_0}^\phi(x) = \mathbb{E} [K(\mathbf{X}_T)],$$

$$\text{avec } \mathbf{X}_{t_0} = x, \quad \mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \phi_t(\mathbf{X}_t), \mathbf{W}_{t+1})$$

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

$$\mathcal{V}_{t_0}(\mu) = \min_{\mu_{t_0}, \dots, \mu_T, \phi(\cdot)} \mathcal{V}_{t_0}^\phi(\mu)$$

$$\mathcal{V}_{t_0}^\phi(\mu) = \sum_x \mu_T(x) K(x)$$

$$\text{avec } \mu_{t_0} = \mu, \quad \mu_{t+1} = \mu_t P_t^\phi$$

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

# LIENS ENTRE LES FONCTIONS VALEURS $\mathcal{V}_{t_0}^\phi(\cdot)$ ET $V_{t_0}^\phi(\cdot)$

On a :

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

En effet :

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

$$V_{t_0}^\phi(x) = (P_{t_0}^\phi \cdots P_{T-1}^\phi K)(x)$$

$$(\text{Par exemple, pour } t_0 = T - 1, \quad V_{T-1}^\phi(x) = \sum_{y \in \mathbb{X}} P_{T-1}^\phi(x, y) K(y))$$

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

$$\mathcal{V}_{t_0}^\phi(\mu) = \mu P_{t_0}^\phi \cdots P_{T-1}^\phi K$$

$$(\text{Par exemple, pour } t_0 = T - 1, \quad \mathcal{V}_{T-1}^\phi(\mu) = \sum_{x, y \in \mathbb{X}} \mu(x) P_{T-1}^\phi(x, y) K(y))$$

# LIEN ENTRE $V$ ET $\mathcal{V}$ SUR UN PB D'HORIZON $T = 1$

$V(x)$

$$V(x) = \mathbb{E}_{X_0=x} [K(\mathbf{X}_1)] = \sum_y P_{x,y} K(y)$$

$\mathcal{V}(\mu)$

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

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  $(\mathcal{P}'_t)$  démarrnant en  $\mu$  à  $t$  :

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

- ▶ À l'instant  $t$ ,  $P_t^\phi$  ne dépend que de la fonction  $\phi_t$ .
- ▶ 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 $t \in \{0, \dots, T\}$

On a la relation suivante (Équation de Bellman) :

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

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

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

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

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

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

# ÉCRITURE AVEC LA MATRICE DE TRANSITION DE LA CHAÎNE DE MARKOV

## Équation de Bellman

$$V_t(x) = \min_{u \in \mathcal{U}} \sum_y P_{x,y}^u V_{t+1}(y)$$

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

$$u^\#(x) \in \operatorname{Argmin}_{u \in \mathcal{U}} \sum_y P_{x,y}^u V_{t+1}(y)$$

Preuve :

$$\begin{aligned} \sum_y P_{x,y}^u V_{t+1}(y) &= \sum_y \mathbb{P}(f(x, u, \mathbf{W}_{t+1}) = y) V_{t+1}(y) \\ &= \mathbb{E}(V_{t+1}(f_t(x, u, \mathbf{W}_{t+1}))) \end{aligned}$$

## HORIZON FINI AVEC COÛT INSTANTANÉ

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

$$V_0(x) = \min_{\mathbf{X}, \mathbf{U} \in \mathcal{U}} \mathbb{E} \left[ \sum_{t=0}^{T-1} L_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}) + K(\mathbf{X}_T) \right],$$

s.c.  $\mathbf{X}_0 = x$  fixé,

$$\mathbf{X}_{t+1} = f_t(\mathbf{X}_t, \mathbf{U}_t, \mathbf{W}_{t+1}), \quad \forall t = 0, \dots, T-1$$

Équation de programmation dynamique

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

## PREUVE

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

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

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

### Équation de Bellman

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

## 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 :

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

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

- ▶ 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  $\tilde{V}_t(0, x) = V_t(x)$ , ce qui donne l'équation de Bellman pour le problème avec coût instantané ( $V$ ).

# PLUS COURT CHEMIN DANS UN GRAPHE

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

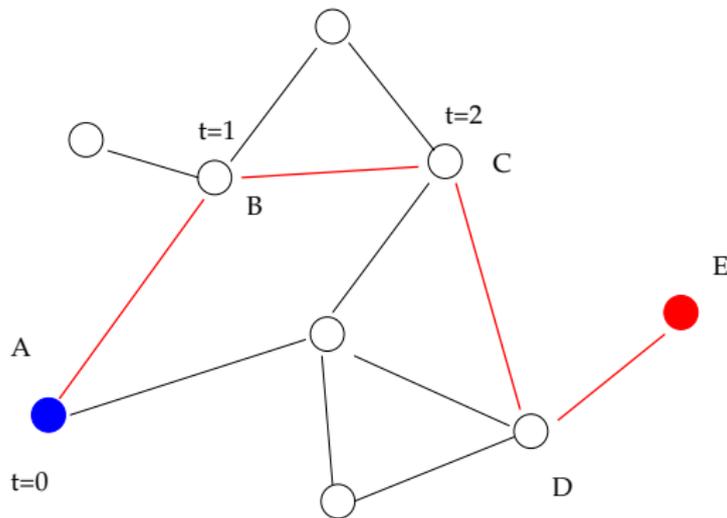


Figure – Plus court chemin

## 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 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 ( $d^{(m)}, m = 0, \dots, n - 1$ ).
- ▶ 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  $k^\#$  cela veut dire que le plus court chemin  $i \rightsquigarrow j = i \rightsquigarrow k^\# \rightarrow j$

- ▶ On regarde alors comment aller de  $i$  à  $k^\#$  en au plus  $n - 2$  arcs.
- ▶ Noter que  $d^{(n-1)} = (W)^{(n-1)}$  (dans l'algèbre  $(\max, +)$ ) : carré itérés.

## RÉCURSION

►  $Fib(n) = (n \leq 1)?1 : Fib(n - 1) + Fib(n - 2)$ .

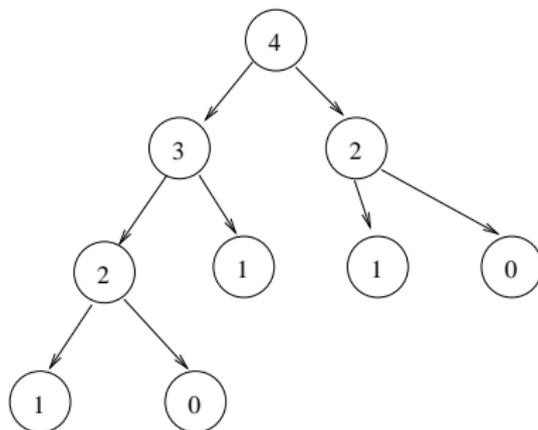


Figure – Fibonacci

# 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 ».