

6. Programmation dynamique

Programmation dynamique

- Supposons un problème décomposé en **étapes**
- Chaque étape possède un certain nombre d'**états** correspondant aux conditions possibles au début d'une étape
- A chaque étape, on doit prendre une **décision** qui associe à l'état courant l'état au début de la prochaine étape
- Dans le cas d'un problème d'optimisation, la programmation dynamique identifie une **politique optimale** : une décision optimale à chaque étape pour chaque état possible

Principe d'optimalité de Bellman

- Étant donné un état courant, la politique optimale pour les étapes à venir est indépendante des décisions prises aux étapes précédentes
- Lorsque ce principe est satisfait, on peut recourir à l'approche suivante pour résoudre un problème ayant N étapes :
 - Identifier une décision optimale pour la dernière étape (N)
 - Pour chaque étape n, de N-1 à 1 : Étant donné une politique optimale pour l'étape n+1, identifier une politique optimale pour l'étape n au moyen d'une **relation de récurrence**
- Lorsque n = 1, on obtient une politique optimale pour le problème

Programmation dynamique : notation

- N = nombre d'étapes
- n = indice de l'étape courante
- s_n = état au début de l'étape n
- x_n = variable de décision associée à l'étape n
- $f_n(s_n, x_n)$ = contribution des étapes n, n+1, ..., N à la valeur de l'objectif lorsqu'on se trouve à l'état s_n au début de l'étape n, que la décision est x_n , et que des décisions optimales sont prises par la suite
- x_n^* = solution optimale à l'étape n
- $f_n^*(s_n) = f_n(s_n, x_n^*) = \min\{f_n(s_n, x_n)\}$ ou $\max\{f_n(s_n, x_n)\}$

Programmation dynamique : tableau

- A chaque itération de la méthode, on construira un tableau représentant :
 - La liste des états possibles à l'étape n
 - Les valeurs possibles de la variable x_n
 - La contribution à la valeur de l'objectif, $f_n(s_n, x_n)$
 - La valeur optimale $f_n^*(s_n)$ et la solution optimale x_n^*

s_n	$x_n=a$	$x_n=b$	$x_n=c$	$f_n^*(s_n)$	x_n^*
1	$f_n(1,a)$	$f_n(1,b)$	$f_n(1,c)$	$f_n^*(1)$	
2	$f_n(2,a)$	$f_n(2,b)$	$f_n(2,c)$	$f_n^*(2)$	

6. Programmation dynamique

5

Exemple 1 : production-inventaire

- Une compagnie doit fournir à son meilleur client trois unités du produit P durant chacune des trois prochaines semaines
- Voici les coûts de production :

Semaine	Production max, temps régulier	Production max, temps supp.	Coût unitaire, temps régulier
1	2	2	300 \$
2	3	2	500 \$
3	1	2	400 \$

6. Programmation dynamique

6

Exemple 1 (suite)

- Le coût pour chaque unité produite en temps supplémentaire est 100 \$ de plus que le coût par unité produite en temps régulier
- Le coût unitaire d'inventaire est de 50 \$ par semaine
- Au début de la première semaine, il y a 2 unités du produit dans l'inventaire
- La compagnie ne veut plus rien avoir dans son inventaire au bout des trois semaines
- Combien doit-on produire d'unités à chaque semaine afin de rencontrer la demande du client, tout en minimisant les coûts?

6. Programmation dynamique

7

Exemple 1 (suite)

- Étape = semaine
- $N = 3$
- $n = 1, 2, 3$
- s_n = état au début de la semaine n = unités de produit dans l'inventaire
- x_n = nombre d'unités produites lors de la semaine n
- $s_{n+1} = s_n + x_n - 3$ (puisque l'on doit livrer 3 unités au client à chaque semaine)
- $s_1 = 2$ (puisque il y a 2 unités au début)

6. Programmation dynamique

8

Exemple 1 (suite)

- c_n = coût unitaire de production, semaine n
- r_n = production max en temps régulier, semaine n
- m_n = production max, semaine n
- $p_n(s_n, x_n) = c_n x_n + 100 \max(0, x_n - r_n) + 50 \max(0, s_n + x_n - 3)$
- $f_n(s_n, x_n) = p_n(s_n, x_n) + f_{n+1}^*(s_{n+1})$
- $f_n^*(s_n) = \min\{p_n(s_n, x_n) + f_{n+1}^*(s_{n+1}) \mid 3 - s_n \leq x_n \leq m_n\}$
- Si $n = 3$, on doit poser $f_4^*(s_4) = 0$
- De plus, $s_4 = 0 = s_3 + x_3 - 3 \rightarrow x_3 = 3 - s_3$
- Calculons d'abord les valeurs $f_3^*(s_3)$ et x_3^*

6. Programmation dynamique

9

Exemple 1 (suite)

s_3	$f_3(3 - s_3, s_3)$	$f_3^*(s_3)$	x_3^*
0	1400	1400	3
1	900	900	2
2	400	400	1
≥ 3	0	0	0

6. Programmation dynamique

10

Exemple 1 (suite)

- Voyons maintenant comment on peut calculer les valeurs $f_2^*(s_2)$ et x_2^* , lorsque $s_2 = 0$
- D'abord, il est clair qu'on doit avoir $x_2 \geq 3$ (car on doit livrer au moins 3 unités du produit)
- $f_2(0, 3) = p_2(0, 3) + f_3^*(0) = 1500 + 1400 = 2900$
- $f_2(0, 4) = p_2(0, 4) + f_3^*(1) = 2150 + 900 = 3050$
- $f_2(0, 5) = p_2(0, 5) + f_3^*(2) = 2800 + 400 = 3200$
- $f_2^*(0) = \min\{f_2(0, 3), f_2(0, 4), f_2(0, 5)\} = f_2(0, 3)$
- D'où $x_2^* = 3$
- De la même façon, on calcule pour $s_2 = 1, 2, 3$

6. Programmation dynamique

11

Exemple 1 (suite)

s_2	$x_2=0$	$x_2=1$	$x_2=2$	$x_2=3$	$x_2=4$	$x_2=5$	$f_2^*(s_2)$	x_2^*
0	-	-	-	2900	3050	3200	2900	3
1	-	-	2400	2450	2600	2850	2400	2
2	-	1900	1950	2000	2250	2900	1900	1
≥ 3	1400	1450	1500	1650	2300	2950	1400	0

6. Programmation dynamique

12

Exemple 1 (suite)

- Pour la première étape ($n = 1$), on a $s_1 = 2$ (il y a 2 unités au départ dans l'inventaire)
- On doit donc avoir $x_1 \geq 1$
- $f_1(2,1) = p_1(2,1) + f_2^*(0) = 300 + 2900 = 3200$
- $f_1(2,2) = p_1(2,2) + f_2^*(1) = 650 + 2400 = 3050$
- $f_1(2,3) = p_1(2,3) + f_2^*(2) = 1100 + 1900 = 3000$
- $f_1(2,4) = p_1(2,4) + f_2^*(3) = 1550 + 1400 = 2950$
- $f_1^*(2) = f_1(2,4)$ et $x_1^* = 4$

Exemple 1 (suite)

- Sous forme de tableau, cela donne :

s_1	$x_1=1$	$x_1=2$	$x_1=3$	$x_1=4$	$f_1^*(s_1)$	x_1^*
2	3200	3050	3000	2950	2950	4

- La politique optimale est donc : $x_1^* = 4$ (d'où $s_2 = 3$), $x_2^* = 0$ (d'où $s_3 = 0$), $x_3^* = 3$
- Le coût total est : $f_1^*(2) = 2950$ \$

Exemple 2 : affectation de ressources

- Trois équipes de recherche travaillent sur un même problème avec chacune une approche différente
- La probabilité que l'équipe E_i échoue est : $P(E1) = 0,4$; $P(E2) = 0,6$; $P(E3) = 0,8$
- Ainsi, la probabilité que les trois équipes échouent simultanément est : $(0,4)(0,6)(0,8) = 0,192$
- On veut ajouter deux autres scientifiques afin de minimiser la probabilité d'échec
- A quelles équipes faut-il affecter ces deux scientifiques afin de minimiser la probabilité d'échec?

Exemple 2 : données

- Voici le tableau représentant les probabilités d'échec pour chaque équipe en fonction du nombre de scientifiques ajoutés

Scientifiques ajoutés	E1	E2	E3
0	0,4	0,6	0,8
1	0,2	0,4	0,5
2	0,15	0,2	0,3

Exemple 2 : modèle

- Étape n = équipe E_n , $n = 1, 2, 3$ ($N = 3$)
- État s_n = nombre de scientifiques non encore affectés
- Décision x_n = nombre de scientifiques affectés à l'équipe E_n
- $p_n(x_n)$ = probabilité d'échec de l'équipe E_n si on lui affecte x_n scientifiques
- $f_n(s_n, x_n) = p_n(x_n) \cdot f_{n+1}^*(s_n - x_n)$ ($f_4^*(0) = 1$)
- $f_n^*(s_n) = \min\{f_n(s_n, x_n) \mid x_n = 0, 1, \dots, s_n\}$
- Calculons d'abord $f_3^*(s_3)$ et x_3^*
- On doit avoir $s_3 = x_3$ (on affecte tous les scientifiques non encore affectés)

Exemple 2 : résolution

s_3	$f_3(s_3, s_3)$	$f_3^*(s_3)$	x_3^*
0	0,8	0,8	0
1	0,5	0,5	1
2	0,3	0,3	2

Exemple 2 : résolution (suite)

- Voyons maintenant comment on peut calculer les valeurs $f_2^*(s_2)$ et x_2^* , lorsque $s_2 = 1$
- Il est clair qu'on doit avoir $x_2 \leq 1$
- $f_2(1, 0) = p_2(1, 0) \cdot f_3^*(1) = 0,6 \cdot 0,5 = 0,3$
- $f_2(1, 1) = p_2(1, 1) \cdot f_3^*(0) = 0,4 \cdot 0,8 = 0,32$
- $f_2^*(1) = \min\{f_2(1, 0), f_2(1, 1)\} = f_2(1, 0)$
- D'où $x_2^* = 0$
- De la même façon, on calcule pour $s_2 = 0$ et 2

Exemple 2 : résolution (suite)

s_2	$x_2=0$	$x_2=1$	$x_2=2$	$f_2^*(s_2)$	x_2^*
0	0,48	-	-	0,48	0
1	0,3	0,32	-	0,3	0
2	0,18	0,2	0,16	0,16	2

Exemple 2 : résolution (suite)

- Pour la dernière étape ($n = 1$), on a $s_1 = 2$ (il y a 2 scientifiques à affecter)
- $f_1(2,0) = p_1(2,0).f_2^*(2) = 0,4.0,16 = 0,064$
- $f_1(2,1) = p_1(2,1).f_2^*(1) = 0,2.0,3 = 0,06$
- $f_1(2,2) = p_1(2,2).f_2^*(0) = 0,15.0,48 = 0,072$
- $f_1^*(2) = f_1(2,1)$ et $x_1^* = 1$

Exemple 2 : résolution (suite et fin)

s_1	$x_1=0$	$x_1=1$	$x_1=2$	$f_1^*(s_1)$	x_1^*
0	0,064	0,06	0,072	0,06	1

- La politique optimale est donc :
 $x_1^* = 1$ (d'où $s_2 = 1$), $x_2^* = 0$ (d'où $s_3 = 1$), $x_3^* = 1$
- La probabilité d'échec est : $f_1^*(1) = 0,06$

Affectation de ressources

- Programmation dynamique : souvent utilisée pour résoudre des problèmes d'affectation de ressources
- C'était le cas dans notre dernier exemple, où il y avait *une* ressource à affecter (les scientifiques)
- Que faire lorsqu'il y a plus d'une ressource?
- Dans ce cas, chaque état est un vecteur dont chaque coordonnée i représente la quantité de la ressource i non encore affectée
- Voyons un exemple bien connu d'affectation de ressources multiples : le problème *Wyndor Glass*

Exemple 3 : problème Wyndor Glass

- Deux types de produits (produit 1, produit 2)
- Trois usines (usine 1, usine 2, usine 3)
- Ressource i : Capacité de production pour l'usine i (h/par semaine)
- Profit par lot (20 unités) de chaque produit
- Chaque lot du produit 1 (2) est le résultat combiné de la production aux usines 1 et 3 (2 et 3)
- Déterminer le taux de production pour chaque produit (nombre de lots/semaine) de façon à maximiser le profit total

Exemple 3 : données

- Voici le tableau représentant les données du problème

	Produit 1 (tps de production, h/lot)	Produit 2 (tps de production, h/lot)	Capacité de production (h)
Usine 1	1	0	4
Usine 2	0	2	12
Usine 3	3	2	18
Profit(\$)/lot	3000	5000	

Exemple 3 : modèle

- Étape $n =$ produit n , $n = 1, 2$ ($N = 2$)
- État $s_n = (R_1, R_2, R_3)_{n,r}$ où R_i est la quantité de la ressource i (temps de production à l'usine i) non encore affectée
- Décision $x_n =$ nombre de lots du produit n
- $f_n(s_n, x_n) = C_n x_n + f_{n+1}^*(s_{n+1})$ ($f_3^*(s_3) = 0$)
- $f_2^*(s_2) = \max\{5x_2 \mid 2x_2 \leq 12, 2x_2 \leq 18 - 3x_1, x_2 \geq 0\}$
- $f_1^*(s_1) = \max\{3x_1 + f_2^*(s_2) \mid x_1 \leq 4, 3x_1 \leq 18, x_1 \geq 0\}$
- $s_1 = (4, 12, 18)$, $s_2 = (4 - x_1, 12, 18 - 3x_1)$

Exemple 3 : résolution

- On cherche d'abord $f_2^*(s_2)$ et x_2^*
- $f_2^*(s_2) = \max\{5x_2 \mid 2x_2 \leq 12, 2x_2 \leq 18 - 3x_1, x_2 \geq 0\}$
- $2x_2 \leq 12$ et $2x_2 \leq 18 - 3x_1 \rightarrow x_2 \leq \min\{6, (18 - 3x_1)/2\}$
- $f_2^*(s_2) = \max\{5x_2 \mid 0 \leq x_2 \leq \min\{6, (18 - 3x_1)/2\}\}$
- $x_2^* = \min\{6, (18 - 3x_1)/2\}$
- $f_2^*(s_2) = 5 \min\{6, (18 - 3x_1)/2\} \mid 3x_1 \leq 18\}$
- Calculons maintenant $f_1^*(s_1)$ et x_1^*

Exemple 3 : résolution (suite)

- $f_1^*(s_1) = \max\{3x_1 + f_2^*(s_2) \mid x_1 \leq 4, 3x_1 \leq 18, x_1 \geq 0\}$
- $x_1 \leq 4 \rightarrow 3x_1 \leq 18$
- $f_1^*(s_1) = \max\{3x_1 + 5 \min\{6, (18 - 3x_1)/2\} \mid 0 \leq x_1 \leq 4\}$
- $\min\{6, (18 - 3x_1)/2\} = 6$, si $0 \leq x_1 \leq 2$
 $= (18 - 3x_1)/2$, si $2 \leq x_1 \leq 4$
- $f_1^*(s_1) = \max\{f_1(s_1, x_1) \mid 0 \leq x_1 \leq 4\}$
- $f_1(s_1, x_1) = 3x_1 + 30$, si $0 \leq x_1 \leq 2$
 $= 45 - (9/2)x_1$, si $2 \leq x_1 \leq 4$
- Le maximum est atteint en $x_1^* = 2$ et $f_1^*(s_1) = 36$

Exemple 3 : résolution (suite et fin)

- $x_2^* = \min\{6, (18-3x_1)/2\}$ et $x_1^* = 2 \rightarrow x_2^* = 6$
- La politique optimale est donc :
 $x_1^* = 2, x_2^* = 6$ de valeur $f_1^*(s_1) = 36$
- En utilisant la même approche, on peut résoudre par la programmation dynamique des modèles où :
 - L'objectif ou les contraintes sont non linéaires
 - Certaines variables sont entières

Cas probabiliste

- Dans la programmation dynamique probabiliste, la décision optimale à l'étape n dépend de la *loi de probabilité* de l'état à l'étape $n+1$
- Ainsi, on passera de l'étape n à l'étape $n+1$, se retrouvant ainsi dans l'état i , avec une probabilité p_i
- Étant donné S états à l'étape $n+1$, on aura $p_i \geq 0$ et $\sum_{i=1}^S p_i = 1$
- La relation de récurrence dépend de cette loi de probabilité et de l'objectif à optimiser

Exemple 4 : jeu de hasard

- Le jeu consiste à miser un nombre quelconque de jetons
- Si on gagne, on gagne le nombre de jetons misés
- Si on perd, on perd le nombre de jetons misés
- Un statisticien croit pouvoir gagner chaque jeu avec une probabilité égale à $2/3$
- Ses collègues parient avec lui qu'en misant au départ 3 jetons, il aura moins de 5 jetons après 3 parties
- Combien de jetons miser à chacune des 3 parties?

Exemple 4 : modèle

- Étape n = partie $n, n = 1, 2, 3 (N = 3)$
- État s_n = nombre de jetons au début de la partie n
- Décision x_n = nombre de jetons à parier à la partie n
- On veut maximiser la probabilité d'avoir au moins 5 jetons après 3 parties
- $f_n(s_n, x_n)$: probabilité de terminer avec au moins 5 jetons, étant donné qu'on est à l'état s_n à l'étape n , qu'on mise x_n jetons et qu'on effectue des décisions optimales aux étapes $n+1, \dots, N$
- $f_n^*(s_n) = \max\{f_n(s_n, x_n) \mid x_n = 0, 1, \dots, s_n\}$

Exemple 4 : modèle (suite)

- Étant donné qu'on est à l'état s_n à l'étape n et qu'on mise x_n jetons, on peut :
 - Perdre et se retrouver à l'état $s_n - x_n$ avec une probabilité $1/3$
 - Gagner et se retrouver à l'état $s_n + x_n$ avec une probabilité $2/3$
- $f_n(s_n, x_n) = (1/3) \cdot f_{n+1}^*(s_n - x_n) + (2/3) \cdot f_{n+1}^*(s_n + x_n)$
- $f_4^*(s_4) = 0$, si $s_4 < 5$
 $= 1$, si $s_4 \geq 5$
- Calculons d'abord $f_3^*(s_3)$ et x_3^*

Exemple 4 : résolution (n = 3)

s_3	$x_3=0$	$x_3=1$	$x_3=2$	$x_3=3$	$x_3=4$	$f_3^*(s_3)$	x_3^*
0	0	-	-	-	-	0	0
1	0	0	-	-	-	0	≥ 0
2	0	0	0	-	-	0	≥ 0
3	0	0	2/3	2/3	-	2/3	≥ 2
4	0	2/3	2/3	2/3	2/3	2/3	≥ 1
≥ 5	1					1	0

Exemple 4 : résolution (n = 2)

- Voyons maintenant comment on peut calculer les valeurs $f_2^*(s_2)$ et x_2^* , lorsque $s_2 = 3$
- Il est clair qu'on doit avoir $x_2 \leq 3$
- $f_2(3,0) = 1/3 \cdot f_3^*(3) + 2/3 \cdot f_3^*(3) = 2/3$
- $f_2(3,1) = 1/3 \cdot f_3^*(2) + 2/3 \cdot f_3^*(4) = 4/9$
- $f_2(3,2) = 1/3 \cdot f_3^*(1) + 2/3 \cdot f_3^*(5) = 2/3$
- $f_2(3,3) = 1/3 \cdot f_3^*(0) + 2/3 \cdot f_3^*(6) = 2/3$
- $f_2^*(3) = \max\{f_2(3,0), f_2(3,1), f_2(3,2), f_2(3,3)\} = f_2(3,0)$
- D'où $x_2^* = 0, 2$ ou 3
- De la même façon, on calcule pour $s_2 = 0, 1, 2, 4, 5$

Exemple 4 : résolution (n = 2)

s_2	$x_2=0$	$x_2=1$	$x_2=2$	$x_2=3$	$x_2=4$	$f_2^*(s_2)$	x_2^*
0	0	-	-	-	-	0	0
1	0	0	-	-	-	0	≥ 0
2	0	4/9	4/9	-	-	4/9	1,2
3	2/3	4/9	2/3	2/3	-	2/3	0,2,3
4	2/3	8/9	2/3	2/3	2/3	8/9	1
≥ 5	1					1	0

Exemple 4 : résolution (n = 1)

- Pour la première étape (n = 1), on a $s_1 = 3$ (3 jetons au départ affecter)
- $f_1(3,0) = 1/3 \cdot f_2^*(3) + 2/3 \cdot f_2^*(3) = 2/3$
- $f_1(3,1) = 1/3 \cdot f_2^*(2) + 2/3 \cdot f_2^*(4) = 20/27$
- $f_1(3,2) = 1/3 \cdot f_2^*(1) + 2/3 \cdot f_2^*(5) = 2/3$
- $f_1(3,3) = 1/3 \cdot f_2^*(0) + 2/3 \cdot f_2^*(6) = 2/3$
- $f_1^*(3) = \max\{f_1(3,0), f_1(3,1), f_1(3,2), f_1(3,3)\} = f_1(3,1)$
- D'où $x_1^* = 1$

6. Programmation dynamique

37

Exemple 4 : résolution (suite et fin)

s_1	$x_1=0$	$x_1=1$	$x_1=2$	$x_1=3$	$f_1^*(s_2)$	x_1^*
3	2/3	20/27	2/3	2/3	20/27	1

- La politique optimale est donc :
 - $x_1^* = 1$
 - Si on gagne ($s_2 = 4$), $x_2^* = 1$
 - Si on gagne ($s_3 = 5$), $x_3^* = 0$
 - Si on perd ($s_3 = 3$), $x_3^* = 2$ ou 3
 - Si on perd ($s_2 = 2$), $x_2^* = 1$ ou 2
 - Si on gagne ($s_3 = 3$ ou 4), $x_3^* = 2$ ou 3 (si $x_2^* = 1$) ou $1 \leq x_3^* \leq 4$ (si $x_2^* = 2$)
 - Si on perd ($s_3 = 1$ ou 0), $x_3^* \geq 0$ (mais le pari est perdu!)

6. Programmation dynamique

38