ECN 6338 Cours 12

La Programmation Dynamique

William McCausland

2022-03-31

Un problème d'optimisation dynamique déterministe à horizon fini

- Objectif :

$$\sum_{t=1}^{T} \pi(x_t, u_t, t) + W(x_{T+1}),$$

  où
  - t est l'index du temps,
  - x_t est l'état en période t, $t = 1, \ldots, T+1$,
  - u_t est la commande (une variable de choix, control en anglais),
  - $\pi(\cdot)$ est le flux de valeur (souvent profit où utilité),
  - $W(\cdot)$ est la valeur terminale.
- Contraintes :
  - x_1 donné,
  - $x_{t+1} = F(x_t, u_t, t)$, $t = 1, \ldots, T$,
  - $u_t \in D(x_t, t)$, $t = 1, \ldots, T$.

La version à horizon infini

- L'objectif devient
$$\sum_{t=0}^{\infty} \beta^t \pi(x_t, u_t).$$
  - le temps commence à zéro,
  - il n'y a pas de valeur terminale,
  - cas très spécial de $\pi(x_t, u_t, t)$: $\beta^t \pi(x_t, u_t)$.
- Quant aux contraintes :
  - $F(x_t, u_t)$ au lieu de $F(x_t, u_t, t)$ (pas vraiment une restriction)
  - $u_t \in D(x_t)$ au lieu de $D(x_t, t)$
  - x_0 donnée au lieu de x_1.

Un exemple : accumulation de la richesse

- Le modèle :

$$\max_{c_t} \sum_{t=0}^{\infty} \beta^t u(c_t) \quad \text{t.q.} \quad k_{t+1} = F(k_t) - c_t,\ k_0 \text{ donné},$$

- Notes :
  - L'état est le stock du capital : $x_t = k_t$
  - La commande est la consommation : $u_t = c_t$
  - F est une fonction de production d'un bien homogène (pas de distinction entre le capital et le seul bien de consommation).
  - On peut inclure l'amortissement du capital dans F : $F(k) = f(k) - \delta k$, où $f(k)$ est la fonction de production grosse.
  - Une interprétation du problème : celui du planificateur central dans un modèle simple de croissance néoclassique : il y a un équilibre concurrentiel qui donne la même résultat.

Autres exemples

- Problème d'un monopôle :
  - la commande est la quantité de travail l et le dividende c
  - l'état est la quantité de capital k dans la firme
  - l'état prochain est le profit (comme fonction de l et k) moins le dividende
  - le flux de valeur est le dividende versé dans une période
- Problème d'inventaire agricole (Judd, 428)
  - la commande est le choix d'intrants de production agricole
  - l'état est l'inventaire agricole
  - l'état prochain est la production moins la consommation
  - le flux de valeur est l'utilité de la consommation d'une période
- Problème de gestion d'un forêt
  - la commande est la quantité récoltée de chaque type (age) d'arbre,
  - l'état est la quantité de chaque type d'arbre,
  - l'état prochain est déterminé par la croissance des arbres et les quantités récoltées,
  - le flux de valeur est le profit apporté par la récolte d'une période

La fonction de valeur

La fonction de valeur est comme une fonction d'utilité indirecte :

- Soit $u(x, y)$ la fonction d'utilité pour deux biens, en quantités x et y.
- La contrainte budgétaire est $p_x x + p_y y = m$.
- La fonction d'utilité indirecte est

$$v(p_x, p_y, m) = \max_{x,y} u(x, y) \quad \text{t.q.} \quad p_x x + p_y y = m.$$

La fonction de valeur pour le problème à horizon fini est

$$V(x, t) = \sup_{u_t,\ldots,u_T} \sum_{s=t}^{T} \pi(x_s, u_s, s) + W(x_{T+1}),$$

sous les contraintes

- $x_t = x$,
- $x_{s+1} = F(x_s, u_s, s)$, $s = t, \ldots, T$,
- $u_s \in D(x_s, s)$, $s = t, \ldots, T$.

Le principe d'optimalilté de Bellman

Une commande optimale $u_t, \ldots, u_T$ a la propriété que quelque soit l'état initial x_t et la décision initiale u_t, les décisions restantes $(u_{t+1}, \ldots, u_T)$ doivent être une commande optimale pour l'état $x_{t+1} = F(x_t, u_t, t)$ résultant de la décision u_t.

Notes :

- Un relation entre fonctions de valeur de différentes périodes :

$$V(x,t) = \sup_{u \in D(x,t)} \pi(x,u,t) + V(F(x,u,t), t+1).$$

- Si $u \in D(x,t)$ atteint le sup, la fonction de politique est

$$U(x,t) = \arg \max_{u \in D(x,t)} \pi(x,u,t) + V(F(x,u,t), t+1).$$

- $U(x,t)$, $V(x,t)$ sans mémoire, fonctions seulement de x et t.

Trouver la fonction de valeur $V(x, 1)$

- Une condition terminal : $V(x, T+1) = W(x)$, une fonction connue.
- Les autres $V(\cdot, t)$ par raisonnement rétrograde : application de l'opérateur définie par

$$V(x, t) = \sup_{u \in D(x,t)} \pi(x, u, t) + V(F(x, u, t), t+1).$$

- Attention : l'opérateur prend une fonction de x et donne une fonction de x.

Le problème déterministe à horizon infini

- La fonction de valeur est maintenant définie de façon récursive :

$$V(x) = \sup_{u \in D(x)} \pi(x, u) + \beta V(F(x, u)),$$

  ou $V = TV$; V est une fonction de $x \in X$, T est un opérateur.
- La solution résout une équation *fonctionnelle*.
- Le théorème des applications contractantes (contraction mapping theorem) : si $0 < \beta < 1$, $\pi(x, u)$ est borné et X est compacte,
  - T est monotone ($y_1 \geq y_2 \Rightarrow Ty_1 \geq Ty_2$)
  - T est une contraction de module β ($\|Ty_1 - Ty_2\| \leq \beta \|y_1 - y_2\|$)
  - $V = TV$ a une solution (T a un point fixe) unique.
- La fonction de politique est maintenant

$$U(x) \in \arg\max_{u \in D(x)} \pi(x, u) + \beta V(F(x, u)).$$

Comment appliquer l'operateur en pratique

- Prenons l'exemple de l'accumulation de richesse.
- La fonction de valeur est un point fixe de l'opérateur T, où

$$(TV)(k) \equiv \max_{0 \leq c \leq F(k)} u(c) + \beta V(F(k) - c).$$

- Comment trouver une approximation de $V(k)$?
- Une méthode : permettre seulement les valeurs de k sur une grille $K = \{k^m, \ldots, k^M\}$.
- En pratique il faut prendre $k^+ \equiv F(k) - c$, le capital à la prochaine période, comme la variable de commande.
- L'équation Bellman avec le changement de variables :

$$V(k) = \max_{k^+ \geq 0} u(F(k) - k^+) + \beta V(k^+).$$

- L'équation Bellman avec $V \colon K \to K$ au lieu de $V \colon \mathbb{R}_+ \to \mathbb{R}_+$:

$$V(k) = \max_{k^+ \in K} u(F(k) - k^+) + \beta V(k^+).$$

- Elle est un système d'équations non-linéaire; la solution est un vecteur.

Exemple

Prenons :

- $\beta = 0.96$
- $u(c) = c^{\gamma+1}/(\gamma+1)$, $\gamma = -2$.
- $F(k) = k + f(k)$, où $f(k) = \frac{1-\beta}{\alpha\beta} k^{\alpha}$, $\alpha = 0.25$
- $K = \{0.8, 0.8+\kappa, 0.8+2\kappa, \ldots, 1.2\}$.

Notes :

- k^{ss} est un état stationnaire si $k^{ss} + F(k^{ss}) - C(k^{ss}) = k^{ss}$.
- β dans la fonction de production est étrange, le coefficient $(1-\beta)/(\alpha\beta)$ fait en sorte que $k^{ss} = 1$.

Choisir une fonction de valeur initiale

- La commande $C(k) = f(k)$ est faisable, mais pas optimal si $k \neq 1$.
- Autour de $k = k^{ss}$, $C(k)$ devrait être près de la commande optimale.
- La commande tient constant l'état k.
- La fonction de valeur $V^c(k)$ associée à la commande $C(k)$ vérifie

$$V^c(k) = u(f(k)) + \beta V^c(k), \text{ ou } V^c(k) = u(f(k))/(1-\beta).$$

- Notez que $V^c(k) \leq V(k)$, où $V(k)$ est la fonction de valeur de la politique optimale.
- Considérez aussi $V_z(k) = 0$ comme fonction de valeur initiale.

Code initial : préférences et technologies

```r
# Valeurs des paramètres
gamma <- -2.0   # Paramètre de préférence
beta <- 0.96    # Paramètre d'impatience
alpha <- 0.25   # Paramètre de production

# Fonction d'utilité
u <- function(c) {
  ifelse(c>0, c^(gamma + 1)/(gamma + 1), -Inf)
}

# Fonction de production
f <- function(k) {
  ((1-beta)/(alpha*beta)) * k^alpha
}
```

Code initial : grille et précomputation

```r
# Grille de capital
kappa <- 0.001
k <- seq(0.5, 1.5, by=kappa)
N <- length(k)

# Matrice de valeurs de u(k + f(k) - k+)
u_k <- function(k, kplus) {
  u(k + f(k) - kplus)
}
u_tab <- outer(k, k, u_k)

# Fonction de valeur initial, épargne zéro
V_c <- u(f(k))/(1-beta)

# Fonction de valeur initial, zéro
V_z <- rep(0, N)
```

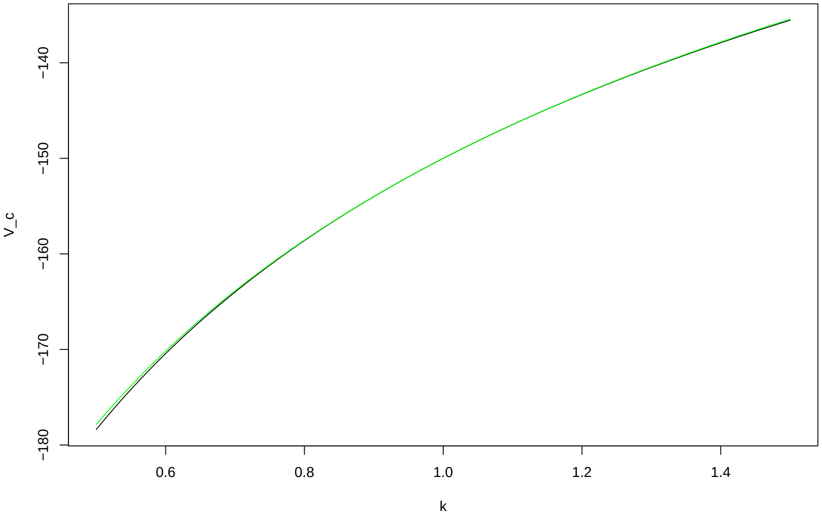
Code, itération de la fonction de valeur

```
# Iteration de la fonction de valeur
VU_suiv <- function(V) {
  V_suiv <- rep(0, N);
  U_suiv <- rep(0, N); U_suiv_i <- vector('integer', N)
  for (i in 1:N) {
    # k[i_plus] est la commande optimale à l'état k[i]
    i_plus = which.max(u_tab[i,] + beta*V)
    U_suiv[i] = k[i_plus]; U_suiv_i = i_plus
    # V_suiv[i] est la fonction V suivante à l'état k[i]
    V_suiv[i] = u_tab[i, i_plus] + beta*V[i_plus]
  }
  list(V=V_suiv, U=U_suiv, U_i=U_suiv_i)
}

# Une itération de T à partir de V_c
VUs_c <- VU_suiv(V_c)
```

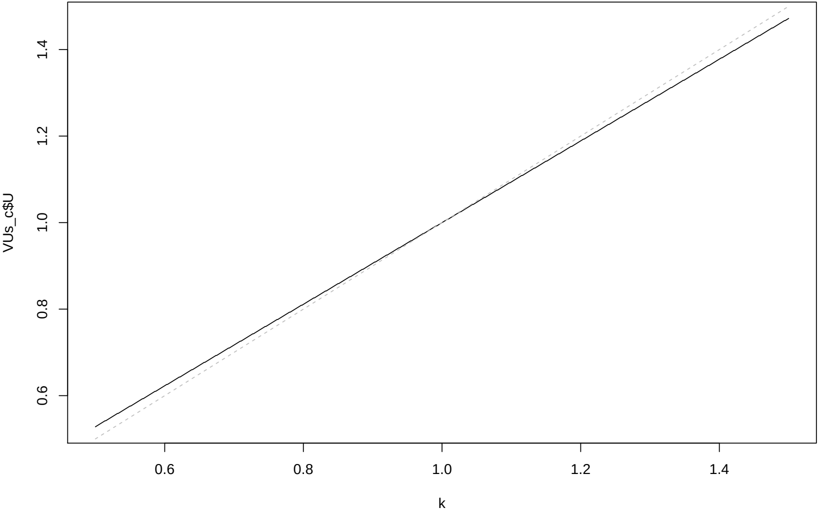
La fonction de valeur $TV_c(k)$ (une itération)

```
plot(k, V_c, type='l')
lines(k, VUs_c$V, col='green')
```



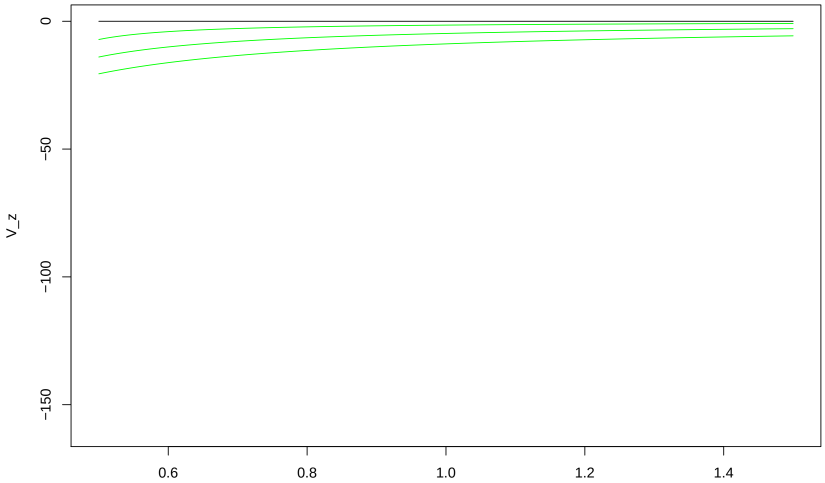
La fonction de politique après une itération de $V_c(k)$

```
plot(k, VUs_c$U, type='l')
lines(k, k, lty='dashed', col='grey')
```



Les fonctions de valeur $T^i V_z(k)$, $i = 1, 2, 3$

```
plot(k, V_z, type='l', ylim = c(-160, 0)); V = V_z
for (i in 1:3) {
  V <- VU_suiv(V)$V; lines(k, V, col='green')
}
```



Itération de la fonction de politique : approximation de V^U pour U donnée

- Pour une fonction de politique $k^+ = U(k)$, pas forcément optimale, on peut calculer, pour chaque $k_0 \in K$, la suite d'état
$$k_0, U(k_0), U^{(2)}(k_0), U^{(3)}(k_0), \ldots,$$
où $U^{(0)}(k_0) = k_0$, $U^{(i+1)}(k_0) = U(U^{(i)}(k_0))$.
- On peut approximer la fonction de valeur de cette politique, le vecteur $V^U(k_0)$, $k_0 \in K$ par $\hat{V}^U(k_0)$ défini comme
$$\sum_{t=0}^{T} \beta^t u(U^{(t)}(k_0) + f(U^{(t)}(k_0)) - U^{(t+1)}(k_0)) + \beta^{T+1} \tilde{V}(U^{(T+1)}(k_0)),$$
où $\tilde{V}$ est une approximation de la fonction de valeur optimale.
- Si on fait ce calcul directe pour chaque k_0 il peut y avoir des calculs redondants; pire, le calcul directe ne marche pas pour les problèmes analogues stochastiques.

La version rétrograde de l'approximation de V^U

Cette version rétrograde évite des calculs redondants et se généralise aux problèmes stochastiques, équation (12.4.2) de Judd :

- $W^0 = V$ (vecteur!)
- Pour $j = 0, \ldots, T$:
  - Pour $k \in K$,

$$W^{j+1}(k) = u(k + f(k) - U(k)) + \beta W^j(U(k)).$$

- $V^U \approx \hat{V}^U \equiv W^{k+1}$ (vecteur!)

Itération de la fonction de politique

Itérez les étapes suivantes jusqu'à ce que $\|V^{l+1} - V^{l}\| < \epsilon$:

1. Utilisez V^{l} pour calculer U^{l+1}, la fonction de politique sous-produite par l'itération de la fonction de valeur.
2. Utilisez U^{l+1} et V^{l} (pour la valeur résiduelle à $t = T + 1$) pour calculer V^{l+1} comme

$$V^{l+1} \equiv \hat{V}^{U^{l+1}}$$