第十讲　动态最优问题求解：（III）　动态规划

樊潇彦

复旦大学经济学院

本讲主要内容

1. 动态规划基本概念

1.1 基本思想：递归求解

1.2 问题示例

1.3 重要概念

2. 确定性动态规划：重要定理与求解方法

2.1 最优化原理

2.2 压缩映射定理

2.3 价值函数迭代

3. 随机动态规划：重要定理与求解方法

3.1 最优化原理和价值函数迭代

3.2 价值函数和策略函数的性质定理

4. 离散选择动态规划（Discrete Choice DP）

动态规划的基本思想

- 我们用拉格朗日方法和最优控制的最大值原理（maximum principle）来分析动态最优问题时，核心工具是关于内生变量的一阶条件，但上述方法很难拓展到内生变量离散、含有外生随机变量，或存在策略性行为等复杂的动态最优问题，因此我们需要更一般化的求解方法。
- Rust(2008): Dynamic programming has enabled economists to formulate and solve a huge variety of problems involving sequential decision making under uncertainty, and as a result it is now widely regarded as the single most important tool in economics.
- Bellman(1957)提出了动态规划的最优化原理（the principle of optimality），核心思想是递归算法（recursive algorithm）：An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute and optimal policy with regard to the state resulting from the first decision.

例1：最短路径的逆向求解（BACKWARD INDUCTION）

从 A 到 E 每一步的路程如右图所示，第 t 步第i 个节点到达 E 的最短路径长度分别为：

$$V_4(E) = 0$$

$$V_3(D_i) = \min\left[l(D_i, E) + V_4(E)\right]$$

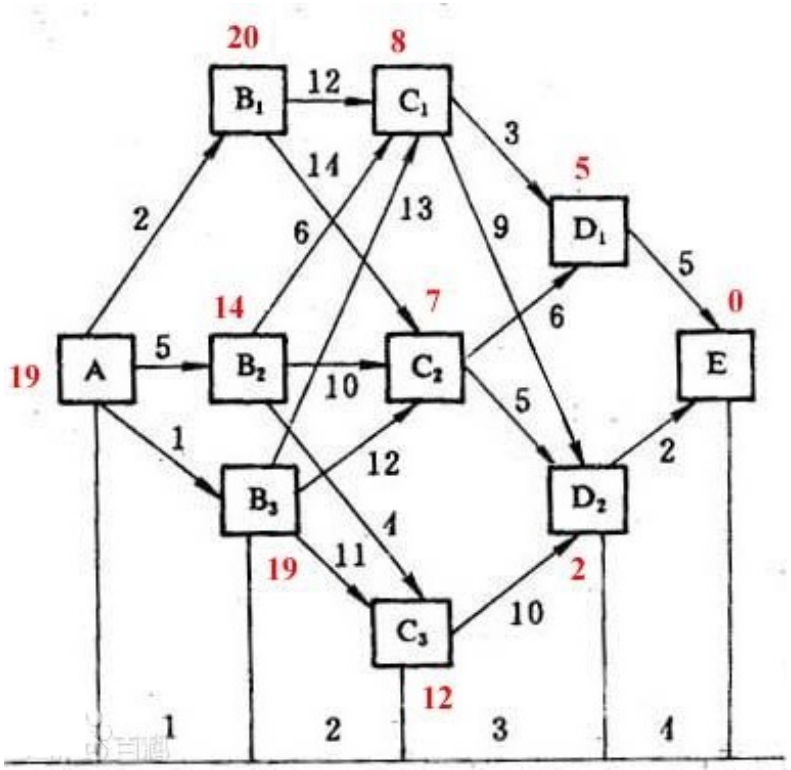
$$V_2(C_i) = \min\left[l(C_i, D_j) + V_3(D_j)\right]$$

$$V_1(B_i) = \min\left[l(B_i, C_j) + V_2(C_j)\right]$$

$$V_0(A) = \min\left[l(A, B_j) + V_1(B_j)\right] = 19$$

上述过程可以表示为逆向求解的递归公式：

$$V_t(i) = \min[l(i, j) + V_{t+1}(j)]$$



图片来源：百度百科“动态规划”词条。

例2：两期生命周期储蓄问题的逆向求解

$$\max \ln c_1 + \beta \ln c_2$$
$$s.t. k_1 = w + Rk_0 - c_1$$
$$k_2 = w + Rk_1 - c_2 = 0$$

- T=2：$c_2 = w + Rk_1$，$V_2(k_1) = \max \ln(w + Rk_1)$
- T=1：

$$\begin{aligned} V_1(k_0) &= \max[\ln c_1 + \beta V_2(k_1)] \\ &= \max[\ln c_1 + \beta \ln(w + R(w + Rk_0 - c_1))] \end{aligned}$$

由一阶条件 $\frac{\partial V_1(k_0)}{\partial c_1} = 0$ 有 $c_1 = \phi_1 w + \phi_2 k_0$（其中：$\phi_1 = \frac{1+R}{(1+\beta)R}$，$\phi_2 = \frac{R}{1+\beta}$），代入可得：

$$V_1(k_0) = (1+\beta)\ln(\phi_1 w + \phi_2 k_0) + \beta \ln(\beta R)$$

例3：确定性最优增长问题

$$\max \sum_{t=0}^{\infty} \beta^t \ln c_t$$
$$s.t.\ k_{t+1} = zk_t^{\alpha} - c_t$$
$$k_0 = 1,\ \lim_{t\to\infty} \beta^t u'(c_t) k_t = 0$$

把目标函数记为期初资本的函数 $V(k_0) = \ln c_0 + \beta \ln c_1 + ... = \ln c_0 + \beta V(k_1)$，以此类推，从任意 $t = 0, 1, 2...$ 开始有：

$$V(k_t) = \max\left[\ln c_t + \beta V(k_{t+1})\right]$$

猜测 $V(k_t) = \theta_0 + \theta_1 \ln k_t$，则：

$$V(k_{t+1}) = \theta_0 + \theta_1 \ln k_{t+1} = \theta_0 + \theta_1 \ln(zk_t^{\alpha} - c_t)$$

由一阶条件 $\frac{\partial V(k_t)}{\partial c_t} = 0$ 有 $c_t = \frac{zk_t^{\alpha}}{1+\beta\theta_1}$。代入价值函数，用待定系数法可得：

$$V(k_t) = \theta_0 + \theta_1 \ln k_t,\ c_t = (1-\alpha\beta)zk_t^{\alpha},\ k_{t+1} = \alpha\beta zk_t^{\alpha}$$

其中：$\theta_0 = \frac{1}{1-\beta}[\frac{\ln z + \alpha\beta \ln(\alpha\beta)}{1-\alpha\beta} + \ln(1-\alpha\beta)]$、$\theta_1 = \frac{\alpha}{1-\alpha\beta}$。

例4：随机性最优增长问题

$$
\begin{aligned}
&\max E_0[\sum_{t=0}^{\infty} \beta^t \ln c_t] \\
&s.t.\ k_{t+1} = z_t k_t^{\alpha} - c_t \\
&\qquad k_0 = 1,\ \ \lim_{t\to\infty} \beta^t u'(c_t) k_t = 0 \\
&\qquad \ln z_t = \rho \ln z_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma^2)
\end{aligned}
$$

当 $t = 0, 1, 2...$ 时，类似地定义：

$$V(k_t, z_t) = \max\{\ln c_t + \beta E_t[V(k_{t+1}, z_{t+1})]\}$$

猜测 $V(k_t, z_t) = \theta_0 + \theta_1 \ln k_t + \theta_2 \ln z_t$，同样利用一阶条件和待定系数法可得：

$$V(k_t, z_t) = \theta_0 + \theta_1 \ln k_t + \theta_2 \ln z_t$$

$$c_t = (1 - \alpha\beta) z_t k_t^{\alpha},\ k_{t+1} = \alpha\beta z_t k_t^{\alpha}$$

其中：$\theta_0 = \frac{1}{1-\beta}[\frac{\alpha\beta \ln(\alpha\beta)}{1-\alpha\beta} + \ln(1-\alpha\beta)]$、$\theta_1 = \frac{\alpha}{1-\alpha\beta}$、$\theta_2 = \frac{1}{(1-\rho\beta)(1-\alpha\beta)}$

重要概念

- 状态变量（state variable）$s_t = (x_t, z_t)$，如内生的资本存量、外生的技术冲击；
- 控制变量（control variable）u_t，如连续的消费决策、离散的是否参加工作的选择等；
- （即期）回报函数（instantaneous payoff function）$F(s_t, u_t)$，如每一期的效用函数、利润函数等；
- 目标函数（objective function）或价值函数（value function）$V(s_t)$，如总效用、总利润等；
- （最优）策略函数（policy function）$u_t^* = \text{argmax} V(s_t)$，如最优消费决策、最优投资策略等。

重要概念

- 内生变量序列形成的历史向量（history）：

$$H_{t-1} = (s_1, u_1, s_2, u_2, ...s_{t-1}, u_{t-1})$$

一般假定时间序列具有马尔科夫性质，历史向量中将只有上一期的变量 (s_{t-1}, u_{t-1}) 会影响当期，从而使问题的分析大大简化。

- 非空的状态变量集合 $s_t \in \mathbb{S}(H_{t-1})$，构成条件包括：
  - 内生状态变量的转移方程（transfer function）$x_t = G(s_{t-1}, u_{t-1})$，如预算约束方程
  - 外生状态变量的条件概率（conditional probability），如马尔科夫链的转移概率矩阵（transition matrix）$P(z_t|z_{t-1})$
- 非空的决策变量约束集（constraint set）$u_t \in \mathbb{U}(s_t)$，如消费的区间 $c_t \in (0, z_t k_t^\alpha)$。$\mathbb{U}$ 是从状态变量到控制变量的实值映射或对应集合（a set-valued mapping or a correspondence）。

最优化原理（Principle of Optimality）

给定初始状态 x_0，动态最优化问题可以表述为序列决策问题（sequential problem, SP）：

$$V^*(x_0) = \sup_{\{u_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t F(x_t, u_t)$$
$$s.t.\ x_{t+1} \in \mathbb{X}(x_t, u_t),\ u_t \in \mathbb{U}(x_t)$$

也可以写为贝尔曼方程（Bellman equation）所描述的方程决策问题（functional problem, FP）（约束条件不变，$t = 0, 1, \ldots$）：

$$V(x_t) = \sup_{u_t} \{F(x_t, u_t) + \beta V(x_{t+1})\}$$

假定 $V^*(s_0)$ 存在且有限，则：

- 贝尔曼方程所定义的价值函数（value function）$V(x_t)$ 存在；
- 存在满足贝尔曼方程的最优策略，即存在策略函数（policy function）：$u_t^* = \text{argmax} V(x_t)$；
- $V^*(x_0) = V(x_0)$，SP与FP等价，点点最优保证序列最优。

动态规划方法的等价性：以最优增长问题为例

最优增长问题的贝尔曼方程为：

$$V(k_t) = \max_{c_t} [\ln c_t + \beta V(k_{t+1})]$$

根据预算约束条件 $k_{t+1} = zk_t^{\alpha} - c_t$，上式等价于：

$$V(k_t) = \max_{k_{t+1}} [\ln(zk_t^{\alpha} - k_{t+1}) + \beta V(k_{t+1})]$$

当价值函数可导时有：

- 一阶条件（FOC）：

$$\frac{\partial V(k_t)}{\partial k_{t+1}} = \frac{-1}{c_t} + \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}} = 0 \Leftrightarrow \frac{1}{c_t} = \beta \frac{\partial V(k_{t+1})}{\partial k_{t+1}}$$

- 包络条件（envelope condition）：

$$\frac{\partial V(k_t)}{\partial k_t} = \frac{z\alpha k_t^{\alpha-1}}{c_t}$$

联立即可得到欧拉方程：

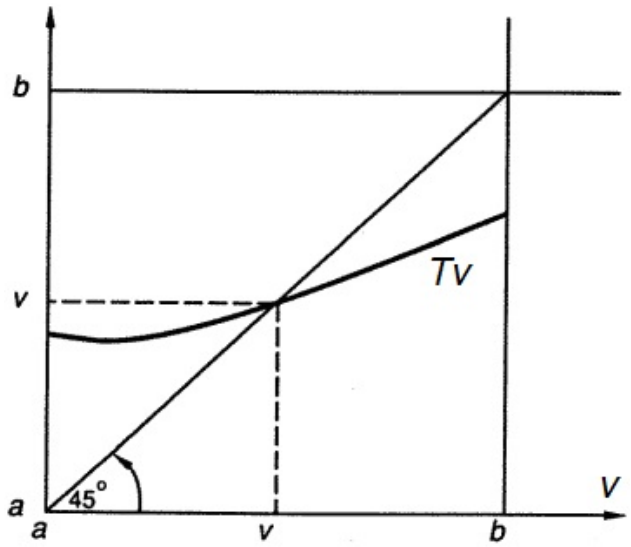
$$c_{t+1} = \beta z\alpha k_{t+1}^{\alpha-1} c_t$$

压缩映射定理（Theorem of Contraction Mapping）

如果 (M,d) 是一个完备度量空间[1]，$T: M \to M$ 是一个模为 β 的压缩映射[2]，则：

- T 在 M 中有唯一的不动点 $V = T(V)$；
- 对任意 $V^{(0)} \in M$ 和 $n = 0, 1...$：

$$d(T^n(V^{(0)}), V) \leq \beta^n d(V^{(0)}, V)$$



图片来源：Stocky and Lucas(1999)。

[1] 度量空间(M,d) 是一个定义了度量（距离函数）d 的集合 M，对 $\forall x, y, z \in M$ 有：(1) $d(x,y) \geq 0$，当且仅当 $x = y$ 时为零；(2) $d(x,y) = d(y,x)$；(3) $d(x,z) \leq d(x,y) + d(y,z)$。如果 S 中每个柯西序列都收敛到 M 中的一个元素 $\lim_{n\to\infty} x_n = x \in M$，则称$(M,d)$ 为完备度量空间（complete metric space）。对于我们关心的问题，在所有有界连续函数 $V: \mathbb{S} \to R$ 构成的集合 $\mathbb{V}$ 上定义上确界范数 $\|V\| = \sup_{s\in\mathbb{S}} |V(s)|$，则 $(\mathbb{V}, \|\cdot\|)$ 是一个完备度量空间。

[2] 给定度量空间 (M,d) 和 $T: M \to M$，若存在 $\beta \in (0,1)$ 对 $\forall x, y \in M$ 有 $d(Tx, Ty) \leq \beta d(x,y)$，则称 T 是一个模为 β 的压缩映射。贝尔曼方程的右边就定义了一个模为 β 的压缩映射。

价值函数迭代（VALUE FUNCTION ITERATION，VFI）

- 算法：
  1. $n=0$，在状态变量的格点 $\{x_i\}_{i=1}^{NX}$ 上对价值函数做初始猜测 $V^{(n)}(x_i)$；
  2. 在整个定义域 $x \in [x_{min}, x_{max}]$ 上拟合出连续的价值函数 $\hat{V}^{(n)}(x)$；
  3. 在格点 $\{x_i\}_{i=1}^{NX}$ 上求解：
  $$V^{(n+1)}(x_{i,t}) = \max_{u_t} \left[F(x_{i,t}, u_t) + \beta \hat{V}^{(n)}(x_{t+1}) \right]$$
  4. 判断是否有 $\left| V^{(n+1)}(x_i) - V^{(n)}(x_i) \right| < \varepsilon$。是：$V = V^{(n+1)}$；否：$n = n+1$，重复步骤2～4。
- 评价：
  - 优点是根据压缩映射定理，只要求 $V(x_t)$ 有界连续，而且从任意初始猜测开始，可以得到全局最优解；
  - 缺点除了收敛速度慢（以 β 为模的压缩映射）以外，计算量将随状态变量个度的增加而迅速上升。比如有K个状态变量、每个状态变量取N个格点，就要在 $NX = N^K$ 个格点上搜索 V，因此工作量将随K的增加呈几何级数增长，出现维数灾难（curse of dimensionality）。

问题描述

给定初始状态 $s_0 = (x_0, z_0)$ 和具有马尔科夫性质 $P(z_t|H_{t-1}) = P(z_t|z_{t-1})$ 的随机变量 z_t，随机动态最优问题描述为：

$$V^*(s_0) = \sup_{\{u_t\}_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta^t F(s_t, u_t) \right]$$
$$s.t.\ x_{t+1} = G(s_t, u_t),\ u_t \in \mathbb{U}(s_t)$$

将转移方程 $x_{t+1} = G(s_t, u_t)$ 和可行集约束 $u_t \in \mathbb{U}(s_t)$ 合并为 $x_{t+1} \in \mathbb{G}(s_t)$。随机动态规划问题描述为：

$$V(s_t) = \sup_{u_t} \{F(s_t, u_t) + \beta E[V(s_{t+1})|s_t]\}$$
$$s.t.\ x_{t+1} \in \mathbb{G}(s_t),\quad (t = 0, 1, 2...)$$

最优化原理

对于一阶马尔科夫链 $z_t \in \mathbb{Z} = \{Z_1, Z_2 ... Z_N\}$，假定：

1. $V^*(s_0)$ 存在且有限；
2. 回报函数 $F(s_t, u_t)$ 连续；
3. 内生状态变量集合 $\mathbb{X}$ 为 $\mathbb{R}^K$ 的紧子集，且实值映射（或对应集合）$\mathbb{G}: \mathbb{S} \to \mathbb{X}$ 非空、紧值且连续（non-empty compact-valued and continuous）

对于一般的马尔科夫随机过程 $z_t \in \mathbb{Z} \subseteq \mathbb{R}^L$，增加假定：

4. z_t 具有费勒性质（Feller property），即对任意有界连续函数 $V(\cdot, z_t)$，$E[V(\cdot, z_{t+1})|z_t]$ 也是 z_t 的有界连续函数

根据 Acemoglu(2009)，有：

- $V(s_t)$ 存在、唯一，且有界连续；
- $u_t^* = \text{argmax} V(s_t)$ 存在；
- $V^*(s_0) = V(s_0)$。

价值函数迭代

1. 将随机过程 z_t 离散化为 NZ 个格点，计算转移矩阵：

$$P_{NZ\times NZ} = [p_{ij}],\ p_{ij} = \Pr\left(z_{t+1} = Z_j | z_t = Z_i\right)$$

2. $n = 0$，在 $NS = NX \times NZ$ 个状态变量的格点 $s_i = (x_i, z_i)$ 上猜测价值函数的值 $V^{(n)}(s_i)$；
3. 给定 z_i，在 $\mathbb{X}$ 上拟合出连续的价值函数 $\hat{V}^{(n)}(x, z_i)$，相应有：

$$E[\hat{V}^{(n)}\left(x_{t+1}, z_{j,t+1}\right)|z_{it}] = \hat{V}^{(n)} \times P^T$$

4. 在格点 s_i 上求解：

$$V^{(n+1)}\left(s_{it}\right) = \max_{u_t}\left\{F\left(s_{it}, u_t\right) + \beta E[\hat{V}^{(n)}\left(x_{t+1}, z_{j,t+1}\right)|z_{it}]\right\}$$

5. 判断是否有 $\left|V^{(n+1)}(s_i) - V^{(n)}(s_i)\right| < \varepsilon$。是：$V = V^{(n+1)}$；否：$n = n + 1$，重复步骤3～5。

价值函数和策略函数的性质定理

- 在假定1～4的基础上，增加假定5：$F(s,u)$ 是 u 的严格凹函数，对应 $\mathbb{G}(x,z)$ 是 x 的凸集（convex in x），则价值函数 $V(s)$ 为严格凹函数，策略函数 u^* 唯一且对 x 连续。
- 在假定1～4的基础上，增加假定6：给定 z,u，$F(x,z,u)$ 是 x 的严格增函数，对应 $\mathbb{G}(x,z)$ 对 x 单调（即对于 $x \leq x'$ 有 $\mathbb{G}(x,z) \subseteq \mathbb{G}(x-,z)$），则 $V(x,z)$ 是 x 的严格增函数。
- 在假定1～5的基础上，增加假定7：给定 z，$F(x,z,u)$ 在 $\mathbb{X} \times \mathbb{U}$ 的内部对 x 和 u 连续可微，则对于 $x \in \mathrm{Int}\mathbb{X}$，$u^* \in \mathrm{Int}\mathbb{U}(s)$，有：

$$V_x(x,z) = F_x(x,z,u^*(x,z))$$

离散选择动态规划：以工作搜索问题为例

对于工作搜索问题：

$$V(x) = Max\left\{\frac{x}{1-\beta},\ \beta\int_0^{\infty} V(x')dF\left(x'\right)\right\}$$

存在某个工资的分界点 $\tilde{x}$（cutting point）：

$$x < \tilde{x} \text{ 跳槽：} V\left(x\right) = \beta\int_0^{\infty} V\left(x'\right)dF\left(x'\right)$$

$$x = \tilde{x} \text{ 无差异：} \frac{\tilde{x}}{1-\beta} = \beta\int_0^{\infty} V\left(x'\right)dF\left(x'\right)$$

$$x > \tilde{x} \text{ 不跳槽：} V\left(x\right) = \frac{x}{1-\beta}$$

代入价值函数，经计算整理可得：

$$\int_0^{\infty} V\left(x\right)dF\left(x\right) = \int_0^{\tilde{x}} \frac{\tilde{x}}{1-\beta}dF\left(x\right) + \int_{\tilde{x}}^{\infty} \frac{x}{1-\beta}dF\left(x\right)$$

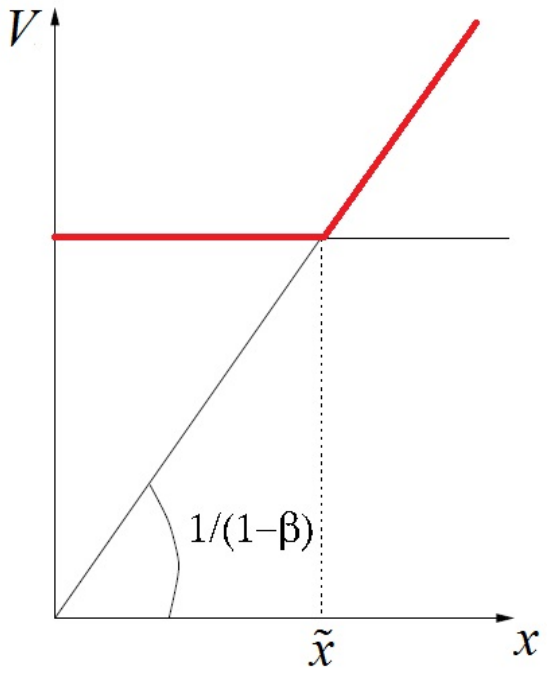
$$\frac{\tilde{x}}{\beta\left(1-\beta\right)} = \frac{\tilde{x}}{1-\beta} + \int_{\tilde{x}}^{\infty} \frac{x-\tilde{x}}{1-\beta}dF\left(x\right)$$

$$\tilde{x} = \frac{\beta}{1-\beta}\int_{\tilde{x}}^{\infty}\left(x-\tilde{x}\right)dF\left(x\right)$$

工作搜索问题

$$\tilde{x} = \frac{\beta}{1-\beta} \int_{\tilde{x}}^{\infty} (x - \tilde{x})\, dF(x)$$

- 从经济含义来看，上式左边是当前的工资收入，即跳槽的机会成本；右边则是跳槽的边际收益，两边相等时是否换工作是无差异的，因此 $\tilde{x}$ 也被称为保留工资。
- 给定参数 β 和工资的分布函数，由于右边是 $\tilde{x}$ 的减函数，因此可以解出唯一的 $\tilde{x}$。易证保留工资是 β 的增函数，但需要假定一些特殊的分布函数，才能得到 $\tilde{x}$ 的显示解。



参考文献

- Acemoglu, D. 2013: *Introduction to Modern Economic Growth*, Princeton University Press
- L.L. 扬奎斯特、T.J. 托马斯著，2013：《递归宏观经济理论（第二版）》，杨斌等译，中国人民大学出版社
- N.L. 斯托基、R.E. 卢卡斯、E.C.普雷斯克特，1999：《经济动态的递归方法》，王明舰译，王永宏校，秦宛顺主审，中国社会科学出版社