第七讲　动态最优简介与数学基础

樊潇彦

复旦大学经济学院

本讲主要内容

1. 动态最优问题简介

1.1 什么是动态最优问题？

1.2 问题分类、示例和求解方法

2. 数学基础（一）：差分（微分）方程

2.1 为什么需要差分（微分）方程？

2.2 方程组的解与平稳性条件

3. 数学基础（二）：随机过程

3.1 随机过程基本概念

3.2 平稳与非平稳时间序列

3.3 马尔科夫过程

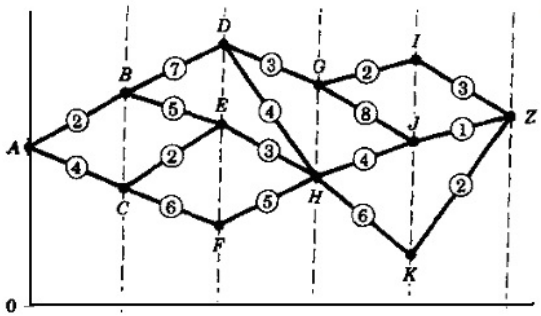
什么是动态最优问题？

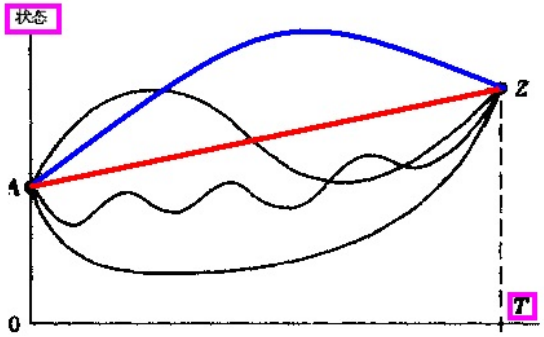
- 静态最优：给定目标函数和约束条件，选择当期最优的决策规则，或最优解 x^*。

$$\underset{x\in X}{Max}\ U(x,z;\theta)$$
$$s.t.\ \ G(x,z;\theta)=0$$

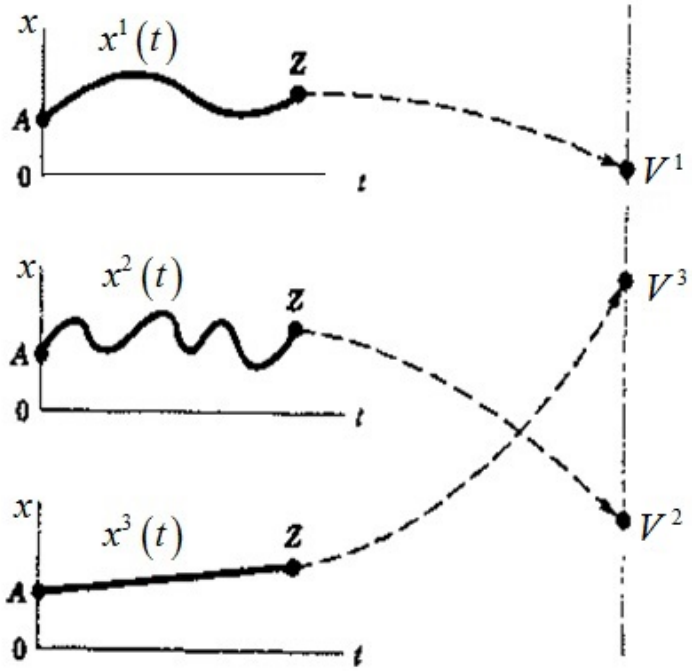
- 动态最优：给定目标函数和约束条件，包括在某些时刻必须要达到的状态，选择一条最优时间路径 $x^*(t)$ 。

$$\underset{x(t)\in X_T}{Max}\ \ U(x(t),z(t);\theta)$$
$$s.t.\ \ G(x(t),z(t);\theta)=0$$





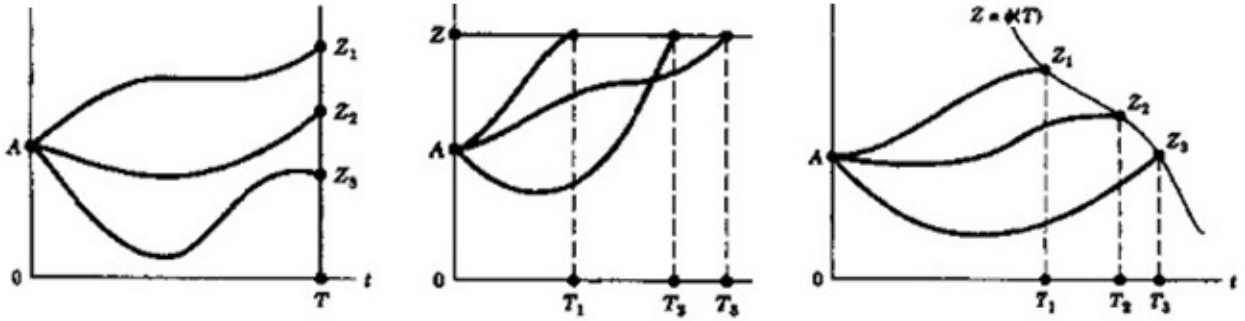
重要概念



- $x_0 = A$, $x_T = Z$ 称为边界条件或横截条件（transversality condition）$h(x, T) = 0$;
- $x(t)$ 为状态变量（state variable），假定状态空间（state space） X_T 非空，即存在允许路径;
- 总效用、总利润等实值函数 $V : X_T \to \mathbb{R}$ 也被称为价值函数（value function）;
- 约束条件中包括描述状态转移的控制方程，或控制变量（control variable）$u(t) = g[x(t), x(t'), z(t); \theta]$, $t' > t$

问题分类

- 离散时间 vs. 连续时间；
- 有限期界 vs. 无限期界；
- 确定性问题 vs. 随机性问题；
- 固定终结点 vs. 可变终结点：
  “初始时间—初始状态”和“终结时间—终结状态”记为有序对 $(0, A)$ 和 (T, Z)，若两者均明确给定，即为固定终结点问题；只给定 $(0, A)$，(T, Z) 不定为可变终结点问题，此时我们需要一个横截条件 $h(x, T) = 0$ 来代替（意为最优路径如何“穿过”终结线）。可变终结点又可分为三种情况：



垂直终结线　　水平终结线　　终结曲线

问题示例

1. 生命周期储蓄

$$
\begin{aligned}
&Max \int_0^T e^{-\rho t} \ln(c)\, dt \\
&s.t.\ \tfrac{dk}{dt} \equiv \dot{k} = w + rk - c \\
&\qquad k(0) = k(T) = 0
\end{aligned}
$$

2. 确定性和不确定性情况下的最优增长

$$
\begin{aligned}
&Max\ \ U_0 = \sum_{t=0}^{\infty} \beta^t \ln c_t \\
&s.t.\ c_t + k_{t+1} = zf(k_t) + (1-\delta) k_t \\
&\qquad k_0 = 1,\ \ \lim_{t\to\infty} \beta^t u'(c_t) k_{t+1} = 0
\end{aligned}
$$

$$
\begin{aligned}
&Max\ \ E(U_0) = E\left[\sum_{t=0}^{\infty} \beta^t u(c_t)\right] \\
&s.t.\ c_t + k_{t+1} = z_t f(k_t) + (1-\delta) k_t \\
&\qquad k_0 = 1,\ \ \lim_{t\to\infty} E_0\left[\beta^t u'(c_t) k_{t+1}\right] = 0 \\
&\qquad \ln z_t = \rho \ln z_{t-1} + \varepsilon_t, z_0 = 1, \rho \in (0,1)
\end{aligned}
$$

3. 工作搜寻

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

求解方法

以确定性固定终结点问题为例 （$x(0)=A, x(T)=Z$）:

求解方法	问题描述	数学工具
拉格朗日 古典变分	$Max\ \ V=\sum_{t=0}^{T} F(x_t, x_{t+1})$ $Max\, V=\int_{t=0}^{T} F(t, x, x')\, dt$	欧拉方程
最优控制	$Max\, V=\int_{t=0}^{T} F(t, x, u)\, dt$ $s.t.\ x'=g(t, x, u)$	汉密尔顿方程
动态规划	$V(x)=\underset{x'\in g(t,x,u)}{Max}\ F(x)+\beta V(x')$	贝尔曼方程

其中：状态变量 x 、控制变量 u、价值函数 V、策略函数（或转移方程）g，以及当期回报函数 F。

为什么需要差分（微分）方程？

以确定性的最优增长问题为例：

$$\begin{aligned} \underset{\{k_{t+1}\}_{t=0}^{\infty}}{Max} \quad & U_0 = \sum_{t=0}^{\infty} \beta^t \ln c_t \\ s.t.\ c_t + k_{t+1} &= zf(k_t) + (1-\delta) k_t \\ k_0 = 1, \quad & \lim_{t\to\infty} \beta^t u'(c_t) k_{t+1} = 0 \end{aligned}$$

等价于：

$$\underset{\{k_{t+1}\}_{t=0}^{\infty}}{Max} \quad U_0 = \sum_{t=0}^{\infty} \beta^t \ln \left[zf(k_t) + (1-\delta) k_t - k_{t+1} \right]$$

由一阶条件 $\frac{dU_0}{dk_{t+1}} = 0$ 可得欧拉方程（Euler equation）：

$$\frac{c_{t+1}}{c_t} = \beta \left[zf'(k_{t+1}) + 1 - \delta \right]$$

为什么需要差分（微分）方程？

欧拉方程和预算约束方程联立得到的二元一阶差分方程即为解条件：

$$\begin{cases} \frac{c_{t+1}}{c_t} = \beta\left[zf'\left(k_{t+1}\right) + 1 - \delta\right] \\ c_t + k_{t+1} = f\left(k_t\right) + \left(1-\delta\right)k_t \end{cases}$$

我们关心的是：

- 给定一个人的初始资本存量 k_0，他应该在每一期消费多少、储蓄多少，以实现一生的效用最大化，即最优路径（optimal path）$\{c_t, k_{t+1}\}_{t=0}^{\infty}$ 是什么？
- 经过长期的动态演化能否达到稳态（steady states）$\lim\limits_{t\to\infty} k_t = k^*, \lim\limits_{t\to\infty} c_t = c^*$？
- 从当前状态出发，经济向稳态过渡将遵循怎样的转移动态（transitional dynamics）过程？

综上，在离散时间动态最优问题中，解条件是差分方程组，对于连续时间问题，解条件就是微分方程组，因此差分（微分）方程是我们分析动态最优问题的基本工具。

一元一阶差分方程的解与平稳性条件[1]

y_t 的动态过程可以表述为一元一阶差分方程（其中 $\varepsilon_t \sim N\left(0, \sigma^2\right)$，$y_0 = \mu + \varepsilon_0$）：

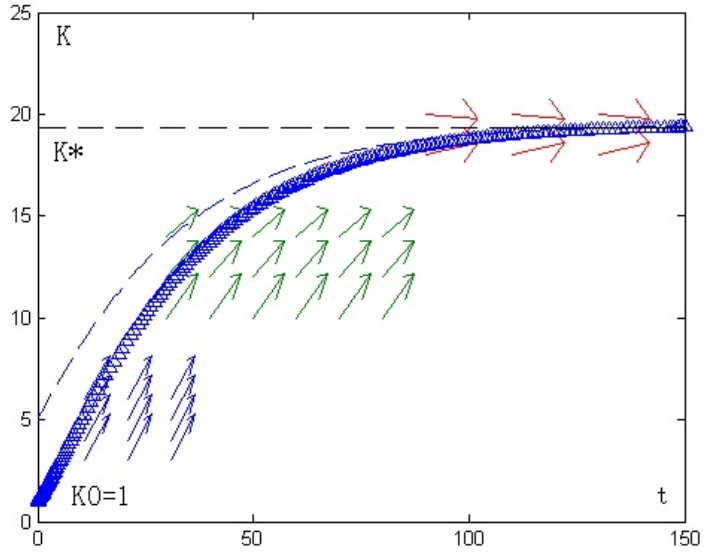
$$y_t = \mu + \phi y_{t-1} + \varepsilon_t = \mu \sum_{s=0}^{t} \phi^s + \sum_{s=0}^{t} \phi^s \varepsilon_{t-s}$$

由于 $\frac{\partial y_t}{\partial \varepsilon_0} = \phi^t$，当时间 $t \to \infty$ 时，y_t 的平稳性意味着当期的随机冲击影响有限，这等价于 $|\phi| < 1$。否则 $\phi > 1$ 时 y_t 会单调发散，$\phi < -1$ 时会震荡发散。

例如在 Solow-Swan 模型中，一个国家的资本积累可以表述为非线性一元一阶差分方程：

$$K_{t+1} = (1-\delta) K_t + sY_t = (1-\delta) K_t + sAK_t^{\alpha}$$

给定参数 $\delta = 0.05$，$s = 0.4$，$A = 1$，$\alpha = 0.3$，右图展示了 $K_0 = 1$ 和 $K_0 = 5$ 两种情况下资本 K_t 的动态演进过程。



[1] 在此重点讨论线性差分方程，关于微分方程的求解可参见龚六堂、苗建军（2014）。

多元一阶差分方程组的解与平稳性条件

一元高阶差分方程 $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} ... + \phi_p y_{t-p} + \varepsilon_t$ 可以用降阶增元的方法改写为多元一阶差分方程组 $Y_t = \tilde{\mu} + \Phi Y_{t-1} + \tilde{\varepsilon}_t$：

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

上述方程组也可以写为 $(I - \Phi L) Y_t = \tilde{\mu} + \tilde{\varepsilon}_t$，根据汉密尔顿(1999)，当 Φ 的所有特征根的模都小于1（$|\lambda_i| < 1$，$(i = 1, 2...p)$）时，方程组的解为：

$$Y_t = (I - \Phi L)^{-1} (\tilde{\mu} + \tilde{\varepsilon}_t)$$

其中 $(I - \Phi L)^{-1} = I + \Phi L + (\Phi L)^2 + ... = \sum_{j=0}^{+\infty} (\Phi L)^j$。

以 $AR(2)$ 为例的稳定性条件

$AR(2)$ 过程 $y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$ 可写为二元一阶方程组：

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} \mu \\ 0 \end{bmatrix} + \begin{bmatrix} \phi_1 & \phi_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix}$$

求解 Φ 的特征方程得到特征根：

$$|\lambda I - \Phi| = \begin{vmatrix} \lambda - \phi_1 & -\phi_2 \\ -1 & \lambda \end{vmatrix}$$

$$= \lambda^2 - \phi_1 \lambda - \phi_2 = 0$$

$$\Rightarrow \lambda_{1,2} = \frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{2}$$

y_t 的平稳性条件 $|\lambda_{1,2}| < 1$ 需要同时满足：

$-1 \le \phi_2 \le 1$、$\phi_2 < 1 - \phi_1$、$\phi_2 < 1 + \phi_1$，

即(ϕ_1, ϕ_2)落在右图所示的红色三角型区域内。

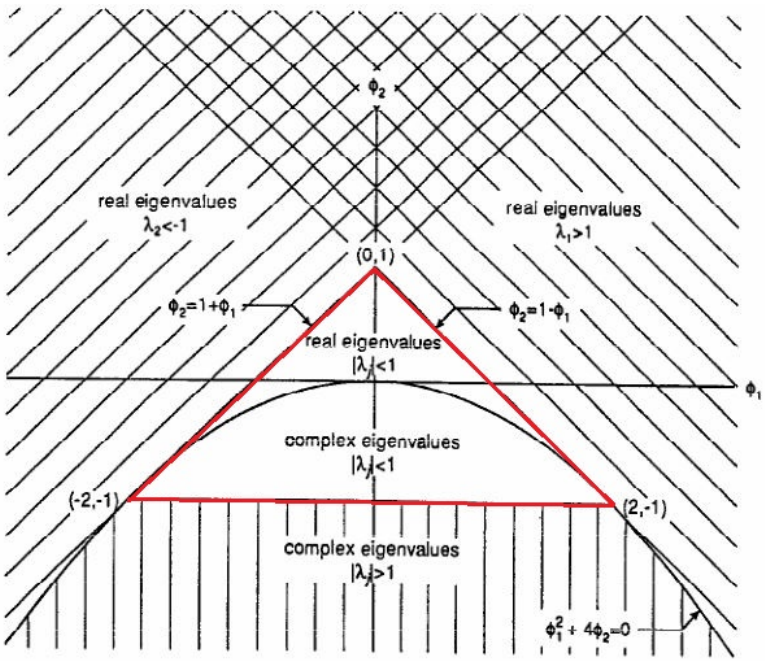


FIGURE 1.5 Summary of dynamics for a second-order difference equation.

随机过程的定义与数字特征

- 随机过程（stochastic process）是随机变量按时间编排的集合。设对每个参数 $t \in T$，$Y(t,\omega)$ 都是一个随机变量，则称下面的随机变量族为随机过程：

$$Y_T = \{Y(t,\omega) \,|\, Y : T \times \Omega \to \mathbb{R}\}$$

- 数字特征：

数字特征	定义	样本值
均值 μ_t	$E(Y_t)$	$\bar{Y} = \frac{1}{T}\sum_{t=1}^{T} Y_t$
方差 $Var(Y_t)$	$E\left[(Y_t-\mu_t)^2\right]$	$s_Y^2 = \frac{1}{T-1}\sum_{t=1}^{T}(Y_t-\bar{Y})^2$
协方差 $\gamma_{t,s}$	$E\left[(Y_t-\mu_t)(Y_s-\mu_s)\right]$	$\frac{1}{T-l-1}\sum_{t=1}^{T-l}(Y_t-\bar{Y})(Y_s-\bar{Y}),\ s=t+l$
自相关系数（ACF） $\rho_{t,s}$	$\frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}}\sqrt{\gamma_{s,s}}}$	$\frac{\sum_{t=1}^{T-l}(Y_t-\bar{Y})(Y_s-\bar{Y})}{\sqrt{\sum_{t=1}^{T-l}(Y_t-\bar{Y})^2}\sqrt{\sum_{s=l+1}^{T}(Y_s-\bar{Y})^2}},\ s=t+l$

随机过程的分类：时间参数与状态空间

1. 时间参数 t 是否连续：序列（series）vs. 过程（process）
   - $t \in \{...-2,-1,0,1,2...\}$ 为可列的离散变量时，称 $\{Y_t\}$ 为时间序列；
   - $t \in (-\infty,\infty)$ 为不可列的连续变量时，称 $\{Y(t)\}$ 为随机过程。
2. 状态空间 $Y_t \in S$ 是否连续
   - 离散状态空间：$S = \{S_1, S_2, ...\}$，如马尔科夫链。
   - 连续状态空间：$S = \mathbb{R}$，如维纳过程和马尔科夫过程等。

随机过程的分类：概率特征

3. 联合分布是否变化：平稳（stationary）vs. 非平稳（non-stationary）
   - $\forall t_1, t_2 ... t_N \in T$ 和 $s > 0$，如果 $Y_{t_1+s}, Y_{t_2+s} ... Y_{t_N+s}$ 和 $Y_{t_1}, Y_{t_2} ... Y_{t_N}$ 有相同的联合分布，则 Y_T 严平稳。
   - $\forall t \in T$，均值 $E(X) = \mu$ 和协方差 $Cov(Y_t, Y_{t+j}) = \gamma_j$ 不随时间 t 而变化，则 Y_T 弱平稳（或宽平稳、协方差平稳）；
   - 不满足上述两种定义的随机过程，是非平稳的。
4. 条件概率是否独立。例如马尔科夫链的定义为：

$$P(Y_{t+1} = s_{t+1} | Y_1 = s_1, Y_2 = s_2, \ldots, Y_t = s_t) = P(Y_{t+1} = s_{t+1} | Y_t = s_t)$$

维纳过程和马尔科夫过程也具有该性质，但 $AR(2)$ 就不具备马氏性质。

平稳时间序列分类

- $AR(p)$：p 阶自回归（autoregressive）

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

- $MA(q)$：q 阶移动平均（moving average）

$$Y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \ldots - \theta_q \varepsilon_{t-q}, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

- $ARMA(p,q)$：(p,q) 阶自回归移动平均（autoregressive moving average）

$$\begin{aligned} Y_t = \phi_0 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} \\ + \varepsilon_t - \theta_1 \varepsilon_{t-1} \ldots - \theta_q \varepsilon_{t-q} \end{aligned}, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

非平稳时间序列分类

- 不带漂移的随机游走（random walk）或单位根（unit-root）过程：

$$Y_t = Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

- 带漂移的随机游走（random walk with drift）：

$$Y_t = \delta + Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

- 带趋势的随机游走（random walk with trend）

$$Y_t = \beta t + Y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N\left(0, \sigma^2\right)$$

平稳与非平稳序列的关系

- 转换：积分与差分
  - 平稳序列经过 d 阶积分后可以变为非平稳序列，例如白噪声过程的1阶积分为 $Y_t = \sum_{s=0}^{t} \varepsilon_0$，其方差 $Var(Y_t) = t\sigma^2$ 将随着 t 的增加而增加，不存在稳态分布。由平稳序列经过 d 阶积分得到的非平稳序列称为 d 阶单积（integrated of order d）序列，记为 $Y_t \sim I(d)$。
  - 反之，非平稳序列 Y_t 经过 d 阶差分后可变为平稳序列，如单位根过程经一阶差分后变为白噪声过程：$\Delta Y_t = Y_t - Y_{t-1} = \varepsilon_t$。一般地，如果 Y_t 经过 d 阶差分后变为 $ARMA(p,q)$ 过程，则称为 $ARIMA(p,d,q)$ 过程。
- 回归：协整与谬误回归

  以线性回归模型 $y_t = X_t\beta + \varepsilon_t$ 为例，假定变量 y_t 和 X_t 是非平稳的。
  - 如果 y_t 可以表示为 X_t 的线性组合，则上述回归得到的残差项 ε_t 平稳，称 y_t 和 X_t 协整（cointegrated）。
  - 否则如果 ε_t 不平稳，大数定理和中心极限定理所要求的有限方差条件不满足，OLS 估计得到的结果不可信，称为谬误回归（spurious regression）。

基本概念

以一阶马尔科夫链 z_t 为例，记 $p_{ij}(t) = P(z_{t+1} = Z_j | z_t = Z_i)$ 为 t 时刻状态为 Z_i、$t+1$ 时刻状态为 Z_j 的概率。

- 若 $p_{ij}(t) \equiv p_{ij}$ 与时间无关，称 z_t 为齐次马氏链，转移矩阵 $P = [p_{ij}]$，其中 $p_{ij} \geq 0$，且列加总 $\sum_{j=1}^{n} p_{ij} = 1$。
- 对于定义在 $\mathbb{Z}$ 上的概率分布 $\pi = (\pi_1, \pi_2 ... \pi_N)$，如果 $\pi = \pi P$，则 π 为 z_t 的平稳分布。显然，π 是 P 的等于 1 的特征值所对应的左特征向量。
- 从某一初始分布 $\pi^0 = \left(\pi_1^0, \pi_2^0 ... \pi_N^0\right)$ 出发，有 $\pi^{t+1} = \pi^t P = \pi^0 P^t$。如果 $\lim_{t \to \infty} \pi^t = \pi^*$ 存在，则称 π^* 为 z_t 的极限分布。如果 π^* 与 π^0 无关，则称该马氏链是遍历的（ergodic）。
- 如果 $\forall i, j,\ p_{ij} > 0$，则有 $\lim_{t \to \infty} P^t = P^* = \begin{bmatrix} \pi_1 & \pi_2 & \dots & \pi_N \\ \dots & \dots & \dots & \dots \\ \pi_1 & \pi_2 & \dots & \pi_N \end{bmatrix}$，$z_t$ 遍历平稳 $\pi = \pi^* = P_i^*$。

五种类型的极限分布[2]

1. 唯一平稳分布：

$$P = \begin{bmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{bmatrix}, \quad P^* = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}$$

2. 唯一平稳分布，但存在离开就不会返回的暂态（transient state）：

$$P = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}, \quad P^* = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}$$

3. 多个平稳分布，且存在暂态：

$$P = \begin{bmatrix} 0.4 & 0.3 & 0.3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad P^* = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

[2] 来自斯托基和卢卡斯（1999）《经济动态的递归方法》。

五种类型的极限分布（续）

4. 两个平稳分布：

$$P = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.7 & 0.3 \\ 0 & 0 & 0.3 & 0.7 \end{bmatrix}, \quad P^* = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

5. 周期：

$$P = \begin{bmatrix} 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{bmatrix}$$

$$P^{2t} \to \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \end{bmatrix}, \quad P^{2t+1} \to \begin{bmatrix} 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \end{bmatrix}$$

应用示例：稳态人口分布

David C. Lay著:《线性代数及其应用（原书第3版）》，刘深泉等译，机械工业出版社，2005

- 根据上图写出人口迁移的转移矩阵。
- 是否存在人口分布的稳态？
- 如果预计未来总人口为1000万，请问居住在城市和农村的各占多少？

应用示例：经济增长的平稳分布

以随机最优增长问题为例，内生变量的动态方程为：

$$\ln k_{t+1} = \ln(\alpha\beta) + \alpha \ln k_t + \ln z_t$$
$$\ln c_{t+1} = \ln(1-\alpha\beta) + \alpha \ln k_t + \ln z_t$$

当 $\varepsilon_t \sim N(0,\sigma^2)$ 时，$\ln z_t \sim N(0, \frac{\sigma^2}{1-\rho^2})$，因此有：

$$\lim_{t\to\infty} E(\ln k_t) = \ln k^* = \frac{\ln(\alpha\beta)}{1-\alpha}$$
$$\lim_{t\to\infty} E(\ln c_t) = \ln c^* = \ln(1-\alpha\beta) + \alpha \ln k^*$$
$$\lim_{t\to\infty} Var(\ln k_t) = \lim_{t\to\infty} Var(\ln c_t) = \frac{\sigma^2}{(1-\alpha^2)(1-\rho^2)}$$
$$\lim_{t\to\infty} Cov(\ln c_t, \ln k_t) = 1$$

经济达到稳态时，内生变量服从平稳分布 $\Phi(k,c)$，这是确定性系统中平稳点 (k^*, c^*) 的随机类似物。

参考文献

- 龚六堂、苗建军，2014:《动态经济学方法（第三版)》，北京大学出版社
- 汉密尔顿，J.D. 1999:《时间序列分析》，刘明志译、李学校、靳云汇主审，中国社会科学出版社
- 林元烈编著，2001:《应用随机过程》，清华大学出版社