

经济最优化的随机模型^{*}(I)

陈 木 法

(北京师范大学, 北京, 100875)

摘 要

本文首先提供华罗庚“计划经济大范围最优化的数学理论”的基本定理的另一证法。然后说明, 对于这一理论, 随机因素是不可轻视的。本文中, 我们建议一种相当一般的概率模型和两种具体模型: 正态模型与 2/3 模型。

§1. 引 言

假设社会生产的各种重要产品编号为 $1, 2, \dots, m$. 以 $x = (x^{(1)}, \dots, x^{(m)}) \in \mathbb{R}^m$ 表相应的产量所构成的向量(单位固定), 称为产综。假定根据目前的生产能力, 已知每生产一个单位的 i 类产品, 需要消耗 $a_{ij}^{(0)}$ 个单位的 j 类产品。矩阵 $A^{(0)} = (a_{ij}^{(0)})$ 表示目前的生产效率或生产水平。称之为消耗系数方阵或结构方阵。从初始产综 x_0 出发, 一年后的产综 x_1 应满足

$$x_0 = x_1 A^{(0)}.$$

若不计消费(带消费情形见下节), 即把第 $n-1$ 年的产综全部投入再生产那么到了第 n 年, 就有

$$x_{n-1} = x_n A^{(n-1)}.$$

作为基本模型, 假定 $A^{(n)}$ 与 n 无关。即关于时间齐次。并记作 A 。此时, 有 $x_n = x_0 A^{-n}$, $n \geq 1$ 。这样, 由初始产综 x_0 和结构方阵 A 刻画出了第 n 年的产综。问题是: 对于何种 x_0 与 A , x_n 的增长速度最快? 而对于其它的 x_0 与 A , 经济如何发展?

作为产综, 当然应要求 $x_n > 0$ (即各分量均为正)。但因 A^{-n} 未必非负, 因而有可能出现 x_n 的某个分量 $x_n^{(j)} \leq 0$ 。

定义 1 称 $T = \min \{n \geq 1: \text{存在 } j \text{ 使 } x_n^{(j)} \leq 0\}$ 为崩溃时, 约定 $\min \phi = +\infty$ 。

定义 2 称 m 阶非负方阵 $A = (a_{ij})$ 不可约, 如果对于每一对 i 和 $j \neq i$, 存在有限多个 $i_1, \dots, i_k$ 使

$$a_{ii_1} a_{i_1 i_2} \cdots a_{i_{k-1} i_k} a_{i_k j} > 0.$$

熟知, 任一不可约非负方阵 A 总有唯一正的左(右)特征向量(不计常数因子), 它们对应于 A 的最大特征根(总是正的)。下面是华罗庚的主要结果

霍英东基金和国家自然科学基金的资助,

* 本文 1989 年 9 月 4 日收到, 1991 年 8 月 8 日收到修改稿

基本定理 设 $A=(a_{ij})$ 为不可约非负方阵, 它的左正特征向量为 u , 对应于最大特征根 $p=\rho(A)$. 假定 A 可逆. 那么 x_n 的最好的增长速度是 $1/p$. 此时 $x_n=\rho^{-n}x_0$, $x_0=u$. 反之, 若 $x_0\neq u$ (忽略正常数因子) 且 A^{-1} 不是非负的, 则 $T<\infty$.

例 1^[1]. 取
$$A=\frac{1}{100}\begin{pmatrix}25 & 14\\40 & 12\end{pmatrix}.$$

则
$$u=\left(\frac{5}{7}(\sqrt{2409}+13), 20\right)\doteq(44.34397483, 20).$$

下表给出取 x_0 为 u 的不同精确值时, 崩溃时 T 的变化情况:

x_0	T
$(44, 20)$	3
$(44.344, 20)$	8
$(44.34397483, 20)$	13

由此看出经济系统的敏感性.

作为本节的结尾, 我们给出基本定理的一种证法. 它的好处是比原先的证明更容易看出这个定理的渊源. 先看 A 为转移概率矩阵 P 的特殊情形. 如果 P 还是非周期的, 则由马氏链的遍历性定理知

$$P^n\rightarrow\mathbb{1}\mu, \quad n\rightarrow\infty.$$

此处 $\mu=(\mu_1, \cdots, \mu_m)$ 为 P 的平稳分布而 $\mathbb{1}$ 是由 m 个 1 构成的列向量. 今设 $x_0>0$, $x_0\mathbb{1}=1$, 则由 $x_nP^n=x_0$ 知

$$1=x_0\mathbb{1}=x_nP^n\mathbb{1}=x_n\mathbb{1}, \quad n\geqslant 1.$$

这样, 若 $x>0$ 对一切 n 成立, 则存在子列 x_{n_k} 使得 $x_{n_k}\rightarrow x^*$, $k\rightarrow\infty$ 而且 $x^*\geqslant 0$, $x^*\mathbb{1}=1$. 进而
$$x_0=(x_0P^{-n_k})P^{n_k}\rightarrow x^*\mathbb{1}=\mu=\mu, \quad k\rightarrow\infty.$$

故必定有 $x_0=\mu$.

易证 P 的非周期性导致 $P^{-1}\not\geqslant 0$. 若 P 有周期 $r>1$, 则以 P^r 代替 P 即可. 只需留意 P^r 可拆成 r 个不可约、非周期子方阵. 此时, 条件 $P^{-1}\not\geqslant 0$ 导出每一子方阵的阶 >1 .

最后, 对任给的不可约非负方阵 A , 总存在对角方阵 $\Lambda=\text{diag}(\lambda_i>0; i=1, 2, \cdots, m)$ 和转移概率矩阵 P , 使得

$$A/\rho(A)=\Lambda P^{-1}\Lambda^{-1}$$

(详见[2; 第 160 页, 基本定理]). 因而一般情形可化成上述特殊情形.

§ 2. 一般模型

本节讨论三个问题. 一是说明研究随机模型的必要性; 二是说明一般的经济模型可归结为马尔可夫过程; 三是建议使用失调时以代替崩溃时.

由于随机因素的干扰, 结构方阵 A 是随机的. 假定 $\{A_n\}_{n=1}^\infty$ 为独立同分布方阵列, 那么上节所讨论的模型成为 $x_{n+1}=x_nA_n$, $n\geqslant 0$. 或者 $x_n=x_0A_1^{-1}\cdots A_n^{-1}$, $n\geqslant 1$. 假如小的随机干扰对于经济模型的影响不大, 那么作为一种近似, 只研究决定性模型也就可以了. 然而事实远非

如此. 请看

例2 设 $A = (a_{ij})$ 如上例. 定义 $A_n = (a_{ij}^{(n)})$ 为

$$a_{ij}^{(n)} = \begin{cases} a_{ij}, & \text{概率 } 2/3 \\ (1 \pm 0.01)a_{ij}, & \text{概率分别为 } 1/6. \end{cases}$$

取初始产综为 $x_0 = (44.344, 20)$. 经济在第 n 步失去平衡的概率分别为

n	1	2	3
概 率	0	0.09	0.65

这样, 在前三年内就失去平衡的概率达 0.74. 与决定性情形相比, 相差甚远. 这个例子正是我们工作的出发点.

如同文[1]中所指出, 经济模型不同于通常的马尔可夫链. 对于后者, $x_n = x_0 P^n$, $n \geq 1$, 此处 P 为转移概率矩阵. 因此, 只要 $x_0 > 0$, 则恒有 $x_n > 0$, 但我们已知经济模型 $x_n = x_0 A^{-n}$, $n \geq 1$ 无此性质. 然而, 我们将证明: 经济模型仍然是一个马尔可夫过程.

让我们先把决定性情形的理想模型加以推广, 使之包括消费. 即允许一部分生产资料转化为消费资料而不用于再生产. 假如每年都拿出增产的第 i 种产品的 $\theta_i \in [0, 1]$ ($1 \leq i \leq m$) 倍用于消费, 那么第 1 年可用于再生产的产综是

$$y_1 = x_0 + (x_1 - x_0)(I - \Theta).$$

此处 $\Theta = \text{diag}(\theta_i)$. 若记 $y_0 = x_0$, 则

$$y_1 = x_1(I - \Theta) + x_0\Theta = y_0(A^{-1}(I - \Theta) + \Theta).$$

递推下去, 得出第 n 年可用于再生产的产综为

$$y_n = y_0(A^{-1}(I - \Theta) + \Theta)^n, \quad n \geq 1.$$

自然会想到应取 y_0 为矩阵 $B = A^{-1}(I - \Theta) + \Theta$ 的左正特征向量. 依据是[3]中的基本定理及下述

引理 设 $0 \leq \theta_i < 1$, $1 \leq i \leq m$. 假定 A 为 m 阶可逆非负原方阵(即对某 N 有 $N^n > 0$)且 $\rho = \rho(A) \leq 1$. 则方阵

$$B = A^{-1}(I - \Theta) + \Theta$$

有唯一正特征根 $\lambda = \lambda(B)$, 它对应于唯一左(右)正特征向量. 此外, 当 $\rho = 1$ 时 $\lambda = 1$; 当 $\rho < 1$ 时 $\lambda > 1$.

证 设 $\lambda u = uB = u(A^{-1}(I - \Theta) + \Theta)$.

则 $u(\lambda I - \Theta)(I - \Theta)^{-1}A = u$. 改记

$$A = \text{diag}\left(\frac{\lambda - \theta_i}{1 - \theta_i}\right) = (\lambda I - \Theta)(I - \Theta)^{-1},$$

则有 $u = uAA$. 这样, 问题归结为求 $\lambda > 0$ 和 $u > 0$ 使上式成立. 显然, 对每 $\lambda \geq 1$, AA 依然是非负原方阵, 记它的最大特征根为 $\rho(\lambda)$. 那么 $\rho(1) = \rho(A) = \rho \leq 1$. 其次, 由[4]引理 2.3 和定理 2.2 知, 当 $\lambda \uparrow \infty$ 时, $\rho(\lambda)$ 严格上升趋于 ∞ . 再则, $\rho(\lambda)$ 关于 λ 连续. 因此, 存在唯一的 $\lambda_0 \geq 1$ 使 $\rho(\lambda_0) = 1$. 改记相应于 λ_0 的 A 为 A_0 . 再记对应于 $\rho(\lambda_0)$ 的 $A_0 A$ 的唯一左正特征向量为 u_0 . 那么

$$u_0 = \rho(\lambda_0)u_0 = u_0 A_0 A.$$

另一方面, 由

$$\begin{aligned} \det(I - (\lambda_0 I - \Theta)(I - \Theta)^{-1}A) &= 0 \\ \Leftrightarrow \det((I - \Theta)(\lambda_0 I - \Theta)^{-1} - A) &= 0 \\ \Leftrightarrow \det(I - (I - \Theta)^{-1}A(\lambda_0 I - \Theta)) &= 0 \end{aligned}$$

及 $(I - \Theta)^{-1}A(\lambda_0 I - \Theta)$ 为非负原方阵知

$$\rho((I - \Theta)^{-1}A(\lambda_0 I - \Theta)) = 1$$

而且存在唯一右正特征向量 v^* 使

$$v^* = (I - \Theta)^{-1}A(\lambda_0 I - \Theta)v^*.$$

等价地,

$$\lambda_0 v^* = Bv^*.$$

现在考虑一般的随机模型. 设 A 为定义在 $(\Omega, \mathcal{F}, \mathbb{P})$ 上的 m 阶随机方阵, 满足

$$\mathbb{P}[\det A = 0] = 0. \quad (1)$$

对每个 $n \geq 2$, 设 A_n 为 $A_1 = A$ 的独立复制品. 再令

$$B_n = A_n^{-1}(I - \Theta) + \Theta.$$

那么, 诸 B_n 独立同分布. 熟知, 对于任给的与诸 B_n 独立的随机变量 x_0 , 若令 $x_n = x_{n-1}B_n$, $n \geq 1$, 则

命题 1 $\{x_n\}_{n \geq 0}$ 是一个时齐马尔可夫过程. 记其一步转移概率为 $P(z, \Gamma)$, $z \in \mathbb{R}^m$, $\Gamma \in \mathcal{B}(\mathbb{R}^m)$, 再记

$$H = \{(z^{(1)}, \dots, z^{(m)}) \in \mathbb{R}^m; z^{(j)} > 0, j = 1, \dots, m\}.$$

那么, 对于每个 $x_0 \in H$ (非随机), 崩溃时

$$T(\omega) = \min \{n \geq 1; x_n(\omega) \notin H\}, \quad \omega \in \Omega$$

的分布由下式决定:

$$\mathbb{P}[T > n] = \int_H P(x_0, dz_1) \int_H P(z_1, dz_2) \cdots \int_H P(z_{n-2}, dz_{n-1}) P(z_{n-1}, H), \quad n \geq 1. \quad (2)$$

在无消费的情形下, 我们有形式上较为简单的表达式.

推论 1 如 $\Theta = 0$ 且条件 (1) 和条件

$$\mathbb{P}[A \geq 0] = 1 \quad (3)$$

满足, 则 $\mathbb{P}[T > n] = P^n(x_0, H)$, $n \geq 1$, $x_0 \in H$. 此处,

$$P^0(x, \Gamma) = I_\Gamma(x), \quad P^n(x, \Gamma) = \int P^{n-1}(x, dz) P(z, \Gamma), \quad x \in \mathbb{R}^m, \Gamma \in \mathcal{B}(\mathbb{R}^m), \quad n \geq 1.$$

这是因为, 若 $x_n \geq 0$, $\mathbb{P}$ -a. s., 则由 $x_{n-1} = x_n A_n$ 知 $x_{n-1} > 0$, $\mathbb{P}$ -a. s.

我们假定 $\{A_n\}$ (或 $\{B_n\}$) 独立同分布是基于经济计划常是短期的 (例如五年计划), 因而可以假定技术水平的提高不显著而不至于造成太大的失真. 如果要考虑较长期的计划, 则需研究非时齐的马尔可夫过程.

再则, 仅仅从避免经济崩溃的角度来看方案的可行性是不够的. 即使在溃崩时的前一年停下来, 进行大调整, 可能也已为时太晚. 因此, 我们建议考虑另一种停止时间.

定义 3 设 $\mathbb{E}B_1$ 有左正特征向量 u 而 D 为 $\mathbb{R}^m$ 中包含 u 的一个给定的区域, 称为控制域. 称

$$S = \min \{n \geq 1; x_n \notin D\}$$

为失调时.

对于例 1, 如取

$$D = \{(z^{(1)}, z^{(2)}) \in \mathbb{R}^2; 0 < z^{(2)} < z^{(1)} < 3.5z^{(2)}\},$$

则有

x_0	S
(44, 20)	3
(44.344, 20)	8
(44.34397483, 20)	12

与原来的结果相比, 只是最后一行的 13 改为 12, 变化不大. 但对于随机情形的例 2, 当 $x_0 = (44.344, 20)$ 时, 例 2 中所给出的概率正是 $S = 1, 2, 3$ 时的概率.

现在, 将我们的做法小结如下: 由 B_1 决定 $\mathbb{E}B_1$ 及其左正特征向量 u (严格地讲, 需假定这种向量的存在性), 取 $x_0 = u$ 并求出

$$n(\alpha) = \max \{n \geq 1: \mathbb{P}(S \geq n) \geq \alpha\} - 1.$$

然后在第 $n(\alpha)$ 年停止计划进行调整. 这里 $\alpha \in \left(\frac{1}{2}, 1\right]$ 是预先确定的可靠性标准. 如 $\alpha = 0.9$, 那么在第 1 年至第 $n(\alpha)$ 年间, 经济失调的可能性是 $1 - \alpha = 0.1$. 称 $1 - \alpha$ 为信度.

§ 3. 正态模型与 2/3 模型

从理论上讲, 计算 T 或 S 的分布, 关键在于算出 $P(z, I)$. 今设 λ 为 $\mathbb{R}^m$ 上的测度 (如 Lebesgue 测度, 计数测度等). 使用积分变换定理可得

命题 2 假定关于 λ , A 有概率密度 $p(Z)$ (Z 为 m 阶实方阵) 且条件 (1) 满足. 命

$$Z(z, I) = \{Z: \det Z \neq 0 \text{ 且 } z(Z^{-1}(I - \Theta) + \Theta) \in I\}, z \in \mathbb{R}^m, I \in \mathcal{B}(\mathbb{R}^m).$$

则
$$P(z, I) = \int_{Z(z, I)} p(Z) \lambda(dZ).$$

现在, 转入更为具体的正态模型. 一种自然的假设是: $a_{ij} = \bar{a}_{ij} + \varepsilon_{ij}$, 其中 ε_{ij} 是均值为零、方差为 σ_{ij}^2 的正态分布而且相互独立, $\bar{a}_{ij} = \mathbb{E}a_{ij}$. 那么, $A = (a_{ij})$ 有密度函数

$$p(Z) = c(m, \sigma) \exp \left[-\frac{1}{2} \sum_{i,j=1}^m (z_{ij} - \bar{a}_{ij})^2 / \sigma_{ij}^2 \right],$$

其中
$$c(m, \sigma) = (2\pi)^{-m^2/2} \left(\prod_{i,j=1}^m \sigma_{ij} \right)^{-1}.$$

推论 2 对于上述正态模型, 转移概率 $P(x, \cdot)$ 有分布密度

$$p(z) = c(m, \sigma) \int \cdots \int_{\substack{-\infty < z_{ij} < \infty \\ i=2, \dots, m \\ j=1, \dots, m}} \exp \left[-\frac{1}{2} J(x, Z, z, \Theta, \sigma) \right] |D(x, Z, z, \Theta)| \prod_{\substack{2 \leq i \leq m \\ 1 \leq j \leq m}} dz_{ij},$$

其中
$$D(x, Z, z, \Theta) = \frac{(\theta_1 - 1)^m}{\prod_{j=2}^m (1 - \theta_j) (z_{11} - \theta_1 x^{(1)})} \begin{vmatrix} x^{(1)} & z_{21} & \cdots & z_{m1} \\ x^{(2)} & z_{22} & \cdots & z_{m2} \\ \vdots & \vdots & & \vdots \\ x^{(m)} & z_{2m} & \cdots & z_{mm} \end{vmatrix},$$

而

$$J(z, Z, z, \Theta, \sigma)$$

$$= \sum_{\substack{2 \leq i \leq m \\ 1 \leq j \leq m}} (z_{ij} - \bar{a}_{ij})^2 / \sigma_{ij}^2 + \sum_{j=1}^m \left[\frac{1 - \theta_1}{z^{(1)} - \theta_1 x^{(1)}} \left(x^{(j)} - \sum_{k=2}^m \frac{z^{(k)} - \theta_k x^{(k)}}{1 - \theta_k} z_{kj} \right) - \bar{a}_{1j} \right]^2 / \sigma_{1j}^2.$$

若使用 $p(z)$ 来计算 $\mathbb{P}[S > n]$, 则需计算一个重数为 $n \cdot m^3$ 的重积分. 即便对于不太大的 n 和 m , 也需要求助于数值计算.

作为本文的结束, 我们建议一种有限步停止的马尔可夫链模型. 虽然从理论上讲, 它不如正态模型准确, 但可能较易于在计算机上实现. 为使模型尽可能地简单而又不太失去代表性, 我们建议只考虑每个 ε_{ij} 为如下离散型随机变量:

$$\begin{aligned} \varepsilon_{ij} &= 0, && \text{概率 } 2/3 \\ &= c_{ij} > 0, && \text{概率 } 1/6 \\ &= -d_{ij} < 0, && \text{概率 } 1/6 \end{aligned}$$

此处 c_{ij} 和 d_{ij} 为常数. 并假定 (ε_{ij}) 相互独立. 此时, 如只考虑前 4 步, 则样本轨道的总数是 $m^{2 \times 3 \times 4} = m^{24}$. 平凡的情况也有 $2^{24} = 16,777,216$ 条轨道. 当然难于一一计算. 这似乎会给人以错觉, 这种模型毫无实际意义. 然而, 正是样本轨道的总数非常之大, 使得每一轨道的贡献相对地很小. 因而成了统计方法的大有用武之地. 例如对于例 2, 笔者曾做 26 次试验, 得出 S 的统计分布如下

S	1	2	3	4	≥ 5
出现次数	0	5	17	3	1
百分比	0	.19	.65	.12	.04

与计算机算出的结果(见例 2)相比, 两者非常接近. 由此看出, 最迟在第 2 年(而不是决定性情形的第 7 年)就应停止计划进行调整.

致谢, 感谢审稿人对本文初稿的宝贵建议.

参 考 献 文

[1] 华罗庚, 计划经济大范围最优化的数学理论, (I) — (XI), 科学通报, 1984 年第 12 期 — 1985 年第 24 期.
[2] 华罗庚, 高等数学引论(余篇), 科学出版社, 1984.
[3] 华罗庚、华苏. 具有二正特征矢量的实方阵的研究, 数学通报, 1985 年第 8 期
[4] Varga, R. S., Matrix Iterative Analysis, Prentice-Hall, Inc. 1962.

STOCHASTIC MODEIS OF ECONOMICAL OPTIMIZATION (I)

CHEN MUFA

(Beijihg Normal University)

This paper first provides a different proof for Hua's fundamental theorem in the mathematical theory of global optimization on planned economy. Then, we show that for this theory, the randomness cannot be ignored. Finally, a quite general stochastic model and two concrete stochastic models (the normal distributed model and the 2/3 model) are proposed.