Markov Chain Models in Economics, Management and Finance

Intensive Lecture Course

in High Economic School, Moscow Russia

Alexander S. Poznyak

Cinvestav, Mexico

April 2017

Contact Data

Alexander S. Poznyak
CINVESTAV-IPN, Mexico
http://www.ctrl.cinvestav.mx/coordinacion/apoznyak/apoznyak.html
apoznyak@ctrl.cinvestav.mx

Resume

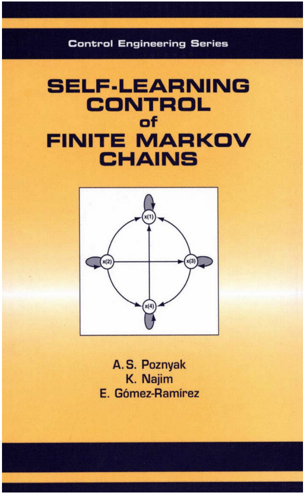
- This course introduces a newly developed optimization technique for a wide class of discrete and continuous-time finite Markov chains models.
- Along with a coherent introduction to the Markov models description (controllable Markov chains classifications, ergodicity property, rate of convergence to a stationary distribution) some optimization methods (such as Lagrange multipliers, Penalty functions and Extra-proximal scheme) are discussed.
- Based on the these models and numerical methods Marketing, Portfolio Optimization, Transfer Pricing as well as Stackleberg-Nash Games, Bargaining and Other Conflict Situations are profoundly considered.
- While all required statements are proved systematically, the emphasis is on understanding and applying the considered theory to real-world situations.

Structure of the course

- **1-st Lecture Day:**
  Basic Notions on Controllable Markov Chains Models, Decision Making and Production Optimization Problem.
- **2-nd Lecture Day:**
  The Mean-Variance Customer Portfolio Problem: Bank Credit Policy Optimization.
- **3-rd Lecture Day:**
  Conflict Situation Resolution: Multi-Participants Problems, Pareto and Nash Concepts, Stackelberg equilibrium.
- **4-th Lecture Day:**
  Bargaining (Negotiation).
- **5-th Lecture Day:**
  Partially Observable Markov Chain Models and Traffic Optimization Problem.

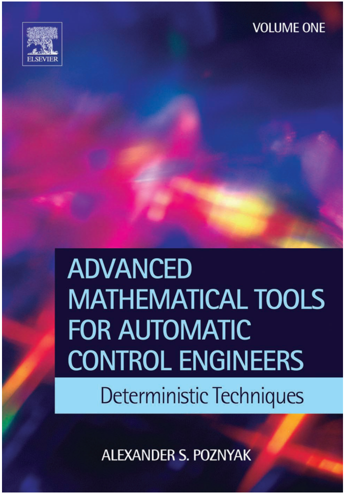
Recommended Bibliography

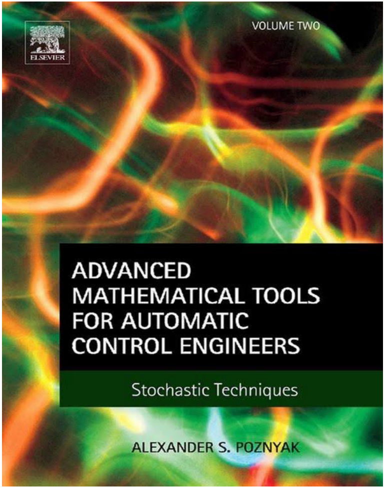
Books:1



Recommended Bibliography

Books:2





Papers recently published (1)

- Julio B. Clempner and **Alexander S. Poznyak**, Simple Computing of the Customer Lifetime Value: a Fixed Local-OoptimalL Policy Approach. J. Syst. Sci. Syst. Eng. December 2014, v.23, issue 4, pp. 439-459.
- Kristal K.Trejo, Julio B..Clempner, **Alexander S. Poznyak**. A Stackelberg security game with random strategies based on the extraproximal theoretic approach. Engineering Applications of Artificial Intelligence, 37 (2015), 145–153.
- Kristal K.Trejo, Julio B..Clempner, **Alexander S. Poznyak**. A Computing the Stckleberge/Nash Equilibria Using the Extraproximal Method. Int. J. Appl. Math. Comput. Sci., 2015, Vol. 25, No. 2, 337–351.
- Julio B. Clempner, **Alexander S. Poznyak**. Modeling the multi-traffic signal-control synchronization: A Markov chains game theory approach. Engineering Applications of Artificial Intelligence, 43 (2015) 147–156.
- Julio B. Clempner, **Alexander S. Poznyak**. Stackelberg security games: Computing the shortest-path equilibrium. Expert Systems with Applications, 42 (2015), 3967–3979.

Recommended Bibliography

Papers recently published (2)

- Emma M. Sanchez, Julio B. Clempner and **Alexander S. Poznyak**. Solving The Mean-Variance Customer Portfolio In Markov Chains Using Iterated Quadratic/Lagrange Programming: A Credit-Card Customer Limits Approach. Expert Systems with Applications. 42 (2015) pp. 5315–5327.
- Emma M. Sanchez, Julio B. Clempner and **Alexander S. Poznyak**. A priori-knowledge/actor-critic reinforcement learning architecture for computing the mean–variance customer portfolio: The case of bank marketing campaigns. Engineering Applications of Artificial Intelligence. Volume 46, Part A, November 2015, Pages 82–92.
- Julio B. Clempner, **Alexander S. Poznyak**. Computing the strong Nash equilibrium for Markov chains games. Applied Mathematics and Computation, Volume 265, 15 August 2015, Pages 911–927.
- Julio B. Clempner, **Alexander S. Poznyak**. Convergence analysis for pure stationary strategies in repeated potential games: Nash, Lyapunov and correlated equilibria. Expert Systems with Applications. Volume 46, 15 March 2016, Pages 474–484.

Recommended Bibliography

Papers recently published (3)

- Julio B. Clempner, **Alexander S. Poznyak**. Solving the Pareto front for multiobjective Markov chains using the minimum Euclidean distance gradient-based optimization method. Mathematics and Computers in Simulation. Volume 119, January 2016, Pages 142–160.
- Julio B. Clempner and **Alexander S. Poznyak**. Constructing the Pareto front for multi-objective Markov chains handling a strong Pareto policy approach. Comp. Appl. Math. DOI 10.1007/s40314-016-0360-6.
- Julio B. Clempner, **Alexander S. Poznyak**. Multiobjective Markov chains optimization problem with strong Pareto frontier: Principles of decision making. Expert Systems With Applications 68 (2017) 123–135.
- J. Clempner and **A. Poznyak**. Analyzing An Optimistic Attitude For The Leader Firm In Duopoly Models: A Strong Stackelberg Equilibrium Based On A Lyapunov Game Theory Approach. Economic Computation And Economic Cybernetics Studies And Research. 4 (2017), 50, 41-60.

1-st Lecture Day

Basic Notions on Controllable Markov Chains Models, Decision Making and Production Optimization Problem

PART 1: MARKOV CHAINS AND DECISION MAKING

Definition

A stochastic dynamic system satisfies **the Markov property,** as it is accepted to say (this definition was introduced by A.A.Markov in 1906), "*if the probable (future) state of the system at any time* $t > s$ *is independent of the (past) behavior of the system at times* $t < s$*, given the present state at time* s".

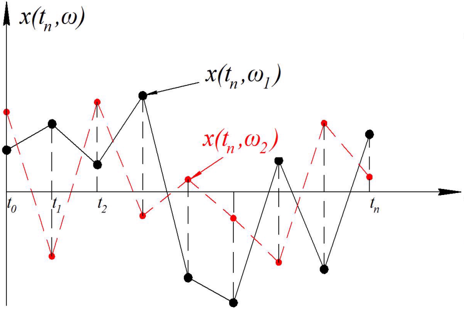
This property can be nicely illustrated by considering *a classical movement of a particle which trajectory after time* s *depends only on its coordinates (position) and velocity at time* s, so that its behavior before time s has no absolutely any affect to its dynamic after time s.

Definition

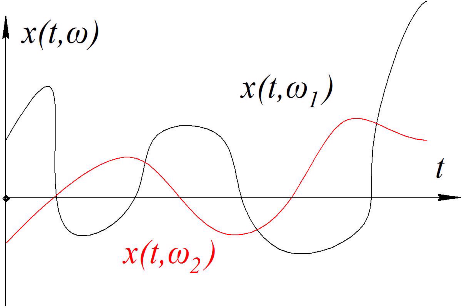
$x(t,\omega) \in \mathbb{R}^n$ is said to be a stochastic process defined on the probability space $(\Omega, \mathcal{F}, \mathsf{P})$ with state space $\mathbb{R}^n$ and the index time-set $J := [t_0, T] \subseteq [0, \infty)$. Here

- $\omega \in \Omega$ is an individual trajectory of the process, Ω a set of elementary events;
- $\mathcal{F}$ is the collection (σ - algebra) of all possible events arising from Ω;
- P is a probabilistic measure (probability) defined for any event $A \in \mathcal{F}$.

The time set J may be
- *discrete*, i.e., $J = [t_0, t_1, ..., t_n, ...)$ - then we talk on **a discrete-time stochastic process** $x(t_n, \omega)$;
- *continuous*, i.e., $J = [t_0, T)$ - then we talk on **a continuos-time stochastic process** $x(t, \omega)$



Discrete time process.



Continuos time process.

Definition

$\{x(t,\omega)\}_{t\in J}$ is called a **Markov process (MP)**, if the following **Markov property** holds: for any $t_0 \leq \tau \leq t \leq T$ an all $A \in \mathcal{B}^n$

$$\boxed{\mathsf{P}\left\{x(t,\omega) \in A \mid \mathcal{F}_{[t_0,\tau]}\right\} \overset{a.s.}{=} \mathsf{P}\{x(t,\omega) \in A \mid x(\tau,\omega)\}}$$

Let the *phase space* of a Markov process $\{x(t,\omega)\}_{t\in\mathcal{T}}$ be *discrete*, that is,

$$x(t,\omega) \in \mathcal{X} := \{(1,2,...,N) \text{ or } \mathbb{N} \cup \{0\}\}$$
$$\mathbb{N} = 1,2,\ldots \text{ is a countable set, or finite}$$

Definition

A Markov process $\{x(t,\omega)\}_{t\in\mathcal{T}}$ with a discrete phase space X is said to be a **Markov chain (or Finite Markov Chain** if $\mathbb{N}$ is finite)

a) in continuous time if

$$\mathcal{T} := [t_0, T), \quad T \text{ is admitted to be } \infty$$

b) in discrete time if

$$\mathcal{T} := \{t_0, t_1, ..., t_T\}, \quad T \text{ is admitted to be } \infty$$

Corollary

The main Markov property for this particular case looks as follows:

- *in continuose time: for any* $i, j \in \mathcal{X}$ *and any* $s_1 < \cdots < s_m < s \leq t \in \mathcal{T}$

$$P\{x(t,\omega) = j \mid x(s_1,\omega) = i_i, ..., x(s_m,\omega) = i_m, x(s,\omega) = i\} \overset{a.s.}{=} P\{x(t,\omega) = j \mid x(s,\omega) = i\}$$

- *in discrete time: for any* $i, j \in \mathcal{X}$ *and any* $n = 0, 1, 2, ...$

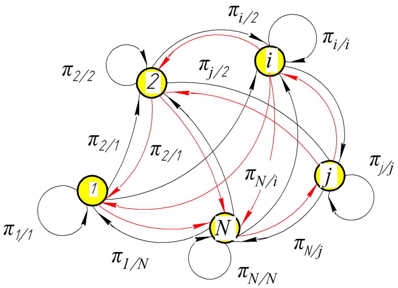
$$P\{x(t_{n+1},\omega) = j \mid x(t_0,\omega) = i_0, ..., x(t_m,\omega) = i_m, x(t_n,\omega) = i\} \overset{a.s.}{=} P\{x(t_{n+1},\omega) = j \mid x(t_n,\omega) = i\} := \pi_{j|i}(n)$$

Homogeneous

Definition

A Markov Chain is said to be Homogeneous **(Stationary)** if the transition probabilities are constant, that is,

$$\pi_{j|i}(n) = \pi_{j|i} = \text{const for all } n = 0, 1, 2, \ldots$$



- Transition matrix Π:

$$\Pi = \begin{bmatrix} \pi_{1|1} & \pi_{2|1} & \cdots & \pi_{N|1} \\ \pi_{1|2} & \pi_{2|2} & \cdots & \pi_{N|2} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{1|N} & \pi_{2|N} & \cdots & \pi_{N|N} \end{bmatrix} = \left[\pi_{j|i}\right]_{i,j=1,\ldots,N}$$

- Stochastic property

$$\sum_{j=1}^{N} \pi_{j|i} = 1 \text{ for all } i = 1, \ldots, N$$

By the Bayes formula

$$\mathrm{P}\{A\} = \sum_i \mathrm{P}\{A \mid B_i\}\,\mathrm{P}\{B_i\}$$

it follows

$$P\{x(t_{n+1},\omega) = j\} = \sum_{i=1}^{N} \underbrace{P\{x(t_{n+1},\omega) = j \mid x(t_n,\omega) = i\}}_{\pi_{j|i}} P\{x(t_n,\omega) = i\}$$

Defining $p_i(n) := P\{x(t_n,\omega) = i\}$, we can write the *Dynamic Model of Finite Markov Chain* as

$$\boxed{p_j(n+1) = \sum_{i=1}^{N} \pi_{j|i}\, p_i(n)}$$

In the vector format, Dynamic Model of Finite Markov Chain can be represented as follows

$$\boxed{p(n+1) = \Pi^{\intercal} p(n), \; p(n) := (p_1(n), ..., p_N(n))^{\intercal}}$$

Iteration back implies

$$\boxed{p(n+1) = \Pi^{\intercal} p(n) = \left(\Pi^{\intercal}\right)^{n+1} p(0)}$$

Definition

A Markov Chain is called **ergodic** if all its states are returnable.

The result below shows that homogeneos ergodic Markov chains posses some additional property:
after a long time such chains "forget" the initial states from which they have started.

Theorem (the ergodic theorem)

Let for some state $j_0 \in X$ *of a homogeneous Markov chain and some* $n_0 > 0$, $\delta \in (0, 1)$ *for all* $i \in (1, ..., N)$

$$\boxed{\left(\Pi^{n_0}\right)_{j_0|i} \geq \delta > 0}$$

i.e., after n_0*-times multiplications* Π *by itself at least one column of the matrix* Π^{n_0} *has all nonzero elements. Then for any initial state distribution* $P\left\{x\left(t_0, \omega\right) = i\right\}$ *and for any* $i, j \in (1, ..., N)$ *there exists the limit*

$$\boxed{p_j^* := \lim_{n\to\infty} \left(\Pi^n\right)_{j|i} > 0}$$

such that for any $t \geq 0$ *this limit is reachable with an exponential rate, namely,*

$$\left|\left(\Pi^n\right)_{j|i} - p_j^*\right| \leq (1-\delta)^{[t_n/n_0]} = e^{-\alpha[t_n/n_0]}, \alpha := \left|\ln(1-\delta)\right|$$

Show that the Finite Markov Chain with the transition matrix

$$\Pi := \begin{bmatrix} 0 & 0.3 & 0 & 0.7 \\ 1 & 0 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

is *ergodic*. Indeed, after 2 steps $(n_0 = 2)$

$$\Pi^2 = \begin{bmatrix} 0.3 & 0.7 & 0 & 0 \\ 0 & 0.3 & 0 & 0.7 \\ 0.09 & 0.03 & 0.81 & 0.07 \\ 1.0 & 0 & 0 & 0 \end{bmatrix}, \Pi^3 = \begin{bmatrix} 0.7 & \mathbf{0.09} & 0 & 0.21 \\ 0.3 & \mathbf{0.7} & 0 & 0 \\ 0.111 & \mathbf{0.097} & 0.729 & 0.063 \\ 0 & \mathbf{0.3} & 0 & 0.7 \end{bmatrix}$$

$\Pi^3 = \Pi^{1+n_0}$ contains the column $j = 2$ with strictly positive elements.

Finite Markov Chains

Ergodicity coefficient

Corollary

If for a homogeneous finite Markov chain with transition matrix Π *the* **ergodicity coefficient** $k_{erg}(n_0)$ *is strictly positive,that is,*

$$k_{erg}(n_0) := 1 - \frac{1}{2} \max_{i,j=1,\dots,N} \sum_{m=1}^{N} \left| (\Pi^{n_0})_{m|i} - (\Pi^{n_0})_{m|j} \right| > 0$$

then this chain is ergodic.

Th following simple estimate holds

$$k_{erg}(n_0) \geq \min_{i=1,\dots,N} \max_{j=1,\dots,N} (\Pi^{n_0})_{j|i} := k_{erg}^{-}(n_0)$$

Corollary

Si, if $k_{erg}^{-}(n_0) > 0$, *then the chain is ergodic.*

Corollary

For any $j \in (1, 2, ..., N)$ of an ergodic homogeneous finite Markov chain the components p_j^* of the **stationary distribution,** *satisfy the folowing* **ergodicity relations**

$$\left.\begin{array}{c} p_j^* = \sum\limits_{i \in \mathcal{X}} \pi_{j|i}\, p_i^* \\ \sum\limits_{i \in \mathcal{X}} p_i^* = 1, p_i^* > 0 \ \ (i = 1, 2, ..., N) \end{array}\right\}$$

or equivalently, in the vector format

$$p^* = \Pi^{\intercal} p^*, \qquad p^* := \left(p_1^*, ..., p_N^*\right), \quad \Pi := \left\| \pi_{j|i} \right\|_{i,j=1,...,N}$$

that is, the positive vector p^* is the eigenvector of the matrix $\Pi^{\intercal}(t)$ corresponding to its eigenvalue equal to 1.

Let $\Pi_k(n) := \left\|\pi_{j|i,k}(n)\right\|_{i,j=1,...,N}$ be the transition matrix with the elements

$$\pi_{j|i,k}(n) := P\left\{x(t_{n+1},\omega) = j \mid x(t_n,\omega) = i,\ a(t_n,\omega) = k\right\}, k = 1, ..., K$$

where the variable $a(t_n,\omega)$ is associated with a control action (decision making) from the given set of possible controls $(1, ..., K)$. Each control action $a(t_n,\omega) = k$ may be selected (realized) in state $x(t_n,\omega) = i$ with the probability

$$d_{ki}(n) := P\left\{a(t_n,\omega) = k \mid x(t_n,\omega) = i\right\}$$

fulfilling the *stochastic constraints*

$$d_{ki}(n) \geq 0, \sum_{k=1}^{N} d_{ki}(n) = 1 \text{ for all } i = 1, ..., N$$

Controllable Markov Chains

What are control strategies (decision making) for Finite Markov Chain processes?

Definitions

A sequence $\{d(0), d(1), ...\}$ of a stochastic matrices

$$d(n) := \|d_{ki}(n)\|_{i=1,...,N;k=1,...,K}$$

with the elements satisfying the stochastic constrains is called **a control strategy** or **decision making process.**

If $d(n) = d$ is a constant stochastic matrix such startegy is named **stationary** one.

Definition

If each row of the matrix d contains one element equal to 1 and others equal to zero, i.e., $d_{ki} = d_{k_0 i}\delta_{k,k_0}$ where δ_{k,k_0} is the Kronecker symbol $\delta_{k,k_0} := \begin{cases} 1 & \text{if} \quad k = k_0 \\ 0 & \text{if} \quad k \neq k_0 \end{cases}$, then the strategy is referred to as **pure**, if at least in one row this is not true, then strategy is called **mixed**.

Example

$$d = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ - a pure strategy} \; ; \; d = \begin{bmatrix} 0 & 0.2 & 0 & 0.8 \\ 1 & 0 & 0 & 0 \\ 0.3 & 0 & 0.7 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ - a mixed strategy}$$

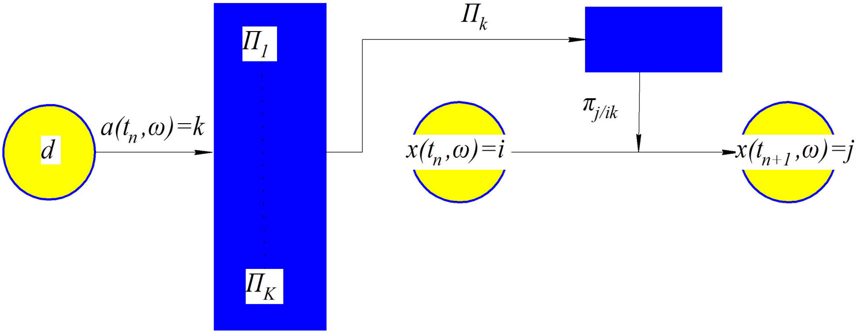


Figure: Structure of a controllable Markov Chain.

Again by the Bayes formula $\mathrm{P}\{A\} = \sum_i \mathrm{P}\{A \mid B_i\} \mathrm{P}\{B_i\}$ we have

$$\pi_{j|i}(n) := P\{x(t_{n+1},\omega) = j \mid x(t_n,\omega) = i\} =$$
$$\sum_{k=1}^{N} \underbrace{P\{x(t_{n+1}) = j \mid x(t_n) = i,\ a(t_n) = k\}}_{\pi_{j|i,k}(n)} \underbrace{P\{a(t_n) = k \mid x(t_n) = i\}}_{d_{k|i}(n)}$$

so,

$$\pi_{j|i}(n) = \sum_{k=1}^{N} \pi_{j|i,k}(n)\, d_{k|i}(n)$$

For homogenous Finite Markov models and stationary under stationary strategies $d_{k|i}(n) = d_{k|i}$ one has

$$\boxed{\pi_{j|i}(d) = \sum_{k=1}^{N} \pi_{j|i,k} d_{k|i}}$$

For stationary startegy $d = \|d_{ki}\|_{i=1,\ldots,N;k=1,\ldots,K}$ we have

$$p_j(n+1) := P\{x(t_n,\omega) = i\} = \sum_{i=1}^{N} \pi_{j|i}(d)\, p_i(n) = \sum_{i=1}^{N} \left(\sum_{k=1}^{N} \pi_{j|i,k} d_{k|i} \right) p_i(n)$$

which represents the *Dynamic Model of Controllable Finite Markov Chain* under a stationary starategy d. If for each d the chain is ergodic, then $p_j(n) \underset{n\to\infty}{\longrightarrow} p_j$ satisfying

$$p_j = \sum_{i=1}^{N} \sum_{k=1}^{N} \pi_{j|i,k} d_{k|i} p_i$$

or

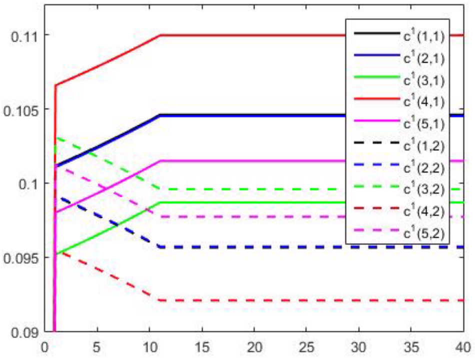


Figure: Convergence to stationary distribution.

Controllable Markov Chains

Dynamics of state probabilities: the vector form

In the vector form the *Dynamic Model* of Controllable Finite Markov Chain (or Decision Making process) under a stationary strategy d looks as

$$p = \Pi^{\intercal}(d)\, p$$

$$\Pi(d) = \left\| \sum_{k=1}^{N} \pi_{j|i,k} d_{ki} \right\|_{i=1,\dots,N; j=1,\dots,N}$$

Fact

So, the final distribution p depends also on the strategy d, that is, $p = p(d)$, so that

$$p(d) = \Pi^{\intercal}(d)\, p(d)$$

PART 2: Simplest Production Optimization Problem

Suppose that some company obtains for the transition

$$x(t_n,\omega) = i,\ \ a(t_n,\omega) = k \ \ \rightarrow x(t_{n+1},\omega) = j$$

from state i to the state j, applying the control k, the following *income*

$$\boxed{W_{j|i,k},\, i,j = 1,...,n,\, k = 1,...,K}$$

Then *the average income* of this company in stationary state is

$$\boxed{J(d) := \sum_{i=1}^{N}\sum_{j=1}^{N} W_{j|i,k}\left(\sum_{k=1}^{N}\pi_{j|i,k}d_{k|i}\right)p_i = \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{N} W_{j|i,k}\pi_{j|i,k}d_{k|i}p_i}$$

where the components p_i satisfies the ergodicity condition

$$p_j(d) = \sum_{i=1}^{N}\sum_{k=1}^{N}\pi_{j|i,k}d_{k|i}p_i(d)$$

The rigorous mathematical problem formulation is as follows:

Problem

$$\left.\begin{array}{c} J(d) = \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k} d_{k|i} p_i(d) \to \max_{d \in \mathcal{D}_{adm}} \\ \textit{under the constrains} \\ \mathcal{D}_{adm} := \left\{ d_{k|i} : p_j(d) = \sum_{i=1}^{N} \sum_{k=1}^{N} \pi_{j|i,k} d_{k|i} p_i(d),\ j = 1, \ldots N \right. \\ \left. d_{k|i} \geq 0,\ \sum_{k=1}^{N} d_{k|i} = 1,\ i = 1, \ldots N \right\} \end{array}\right\}$$

Definition

The matrix d^{br} is called the **best-reply strategy** if

$$d^{br}_{\beta|\alpha} = \begin{cases} 1 & \text{if } \sum_{j=1}^{N} W_{j|\alpha,\beta}\pi_{j|\alpha,\beta} \geq \sum_{j=1}^{N} W_{j|i,k}\pi_{j|i,k} \\ 0 & \text{if } \quad \text{not} \end{cases}$$

Indeed, the upper bound for $J(d)$ can be estimated as

$$J(d) = \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{j=1}^{N} W_{j|i,k}\pi_{j|i,k} d_{k|i} p_i(d) \leq \sum_{i=1}^{N} \max_k \left(\sum_{j=1}^{N} W_{j|i,k}\pi_{j|i,k} \right) p_i(d)$$

which is reachable for $d^{br}_{k|i} = d^{br}_{k|i}$. It is **optimal** if and only if

$$\max_k \left(\sum_{j=1}^{N} W_{j|i,k}\pi_{j|i,k} \right) = \max_k \left(\sum_{j=1}^{N} W_{j|s,k}\pi_{j|s,k} \right) \; \forall i, s \qquad (1)$$

Example (State and action spaces interpretation)

Let

- the state $x(t_n, \omega) = i$ be associated with a number of working unites (staff places);
- the action $a(t_n, \omega) = k$ is related with the financial schedule (possible wage increase,decreasing or no changes): $k = (-1, 0, 1)$;
- the incomes for these actions may be calculated as

$$\boxed{W_{j|i,k} = \left[v_0 - (v + \Delta v k) - v_1\right](j - i)}$$

where v_0 the price of the product, produced by a single working unit with the salary v, its Δv adjustment and the production costs v_1 supporting this process.

Example (State and action spaces interpretation (continuation-1))

For example, for $N = 3$, $i = (10, 20, 30)$ and $v_0 = 400,000.00$, $v_1 = 20,000.00$, $v = 80,000.00$, $\Delta v = 5,000.00$ we have

$$\left\| W_{j|i,k=-1} \right\| = \begin{bmatrix} 0 & 3050,000.00 & 6100,000.00 \\ -3050,000.00 & 0 & 3050,000.00 \\ -6100,000.00 & -3050,000.00 & 0 \end{bmatrix}$$

$$\left\| W_{j|i,k=0} \right\| = \begin{bmatrix} 0 & 3000,000.00 & 6000,000.00 \\ -3000,000.00 & 0 & 3000,000.00 \\ -6000,000.00 & -3000,000.00 & 0 \end{bmatrix}$$

$$\left\| W_{j|i,k=1} \right\| = \begin{bmatrix} 0 & 2950,000.00 & 5900,000.00 \\ -2950,000.00 & 0 & 2950,000.00 \\ -5900,000.00 & -2950,000.00 & 0 \end{bmatrix}$$

Example (State and action spaces interpretation (continuation-2))

Let the transition matrices $\left\|\pi_{j|i,k}\right\|$ be as follows:

$$\underbrace{\begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}}_{k=-1}, \underbrace{\begin{bmatrix} 0 & 0.1 & 0.9 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}}_{k=0}, \underbrace{\begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0 & 0.25 & 0.75 \\ 1 & 0 & 0 \end{bmatrix}}_{k=1}$$

Example (State and action spaces interpretation (continuation-3))

Then the matrix $\left\| \sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k} \right\|$, participating in the average income, is

$$\left\| \sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k} \right\| = \left\| \begin{array}{ccc} 2\,745\,000 & 5\,700\,000 & 2\,360\,000 \\ 1\,525\,000 & 3\,000\,000 & 2\,212\,500 \\ -1\,525\,000 & 0 & -5\,900\,000 \end{array} \right\|$$

and the best reply strategy is (it is non-optimal)

$$d_{br} = \left\| \begin{array}{ccc} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{array} \right\|$$

So, no changes are recommended since $k^* = 0$ for all states i.

Optimization Problem with Additional Constrains

Problem formulation with additional constrains

Problem

$$
\left.
\begin{array}{c}
J(d) = \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k} d_{k|i} p_i(d) \to \max_{d \in \mathcal{D}_{adm}} \\
\textit{under the constrains} \\
\mathcal{D}_{adm} := \left\{ d_{k|i} \colon p_j(d) = \sum_{i=1}^{N} \sum_{k=1}^{N} \pi_{j|i,k} d_{k|i} p_i(d),\ j = 1, \ldots N \right. \\
d_{k|i} \geq 0,\ \sum_{k=1}^{N} d_{k|i} = 1,\ i = 1, \ldots N \\
\left. \sum_{i=1}^{N} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} A_{j|i,k}^{(l)} \pi_{j|i,k} \right) d_{k|i} p_i(d) \leq b_l,\ l = 1, \ldots, L \right\}
\end{array}
\right\}
$$

The additional constrains may be interpreted as some *financial limitations*.

Fact

- ***The best reply strategy** in general is **non optimal**: it may not satisfy condition (1) and the additional constrans.*
- *The functional*

$$J(d) = \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k} d_{k|i} p_i(d)$$

as well as the constrains

$$p_j(d) = \sum_{i=1}^{N} \sum_{k=1}^{N} \pi_{j|i,k} d_{k|i} p_i(d), \sum_{i=1}^{N} \sum_{k=1}^{K} \sum_{j=1}^{N} A_{j|i,k}^{(l)} \pi_{j|i,k} d_{k|i} p_i(d) \leq b_l$$

*are extremely **nonlinear functions** of d.*

The question is: what can we do in this complex situation?

Optimization Problem with Additional Constrains

What can we do in this complex situation? Answer: c-variables!

Definition

Define new variables

$$\boxed{c_{ik} := d_{k|i} p_i (d)}$$

Then the Production Optimization Problem can be express as **a Linear Programming Problem** solved by standard Matlab Toolbox:

$$J(d) = \sum_{i=1}^{N} \sum_{k=1}^{K} \underbrace{\sum_{j=1}^{N} W_{j|i,k} \pi_{j|i,k}}_{W_{ik}^{\pi}} \underbrace{d_{k|i} p_i (d)}_{c_{ik}} = \sum_{i=1}^{N} \sum_{k=1}^{K} W_{ik}^{\pi} c_{ik} := J(c) \to \min_{c \in \mathcal{C}_{adm}}$$

$$\mathcal{C}_{adm} := \left\{ c_{ik} : \sum_{k=1}^{N} c_{jk} = \sum_{i=1}^{N} \sum_{k=1}^{K} \pi_{j|i,k} c_{ik},\ j = \overline{1, N},\ c_{ik} \geq 0,\ \sum_{i=1}^{N} \sum_{k=1}^{K} c_{ik} = 1, \right.$$
$$\left. \sum_{i=1}^{N} \sum_{k=1}^{K} \left(\sum_{j=1}^{N} A_{j|i,k}^{(l)} \pi_{j|i,k} \right) c_{ik} \leq b_l,\ l = \overline{1, L} \right\}$$

Corollary

For **c-variables** *defined as* $c_{ik} := d_{k|i} p_i(d)$ *and found as the solution* c^* *of the LPP above*

1) *we can recuperate the state distribution* $p_i(d^*)$ *as*

$$p_i(d^*) = \sum_{k=1}^{K} c^*_{ik} > 0 \text{ (by the ergodicity property)}$$

2) *and the optimal control strategy (or decision making)* $d^*_{k|i}$ *can be recuperated as*

$$d^*_{k|i} = \frac{c^*_{ik}}{p^*_i(d)} = \frac{c^*_{ik}}{\sum_{k=1}^{K} c^*_{ik}}$$

Soltion of LPP

Numerical solution of LPP with the data of the previous example

$$c_{ik}^{*} = \begin{bmatrix} 0.0001 & 0.0001 & 0.0001 \\ 0.0001 & 0.0001 & 0.3996 \\ 0.5996 & 0.0001 & 0.0002 \end{bmatrix}$$

$$p_i^{*}(d) = (0.0003, 0.3998, 0.5999)$$

$$d_{k|i}^{*} = \begin{bmatrix} 0.3333 & 0.3333 & 0.3333 \\ 0.0002 & 0.0003 & 0.9995 \\ 0.9995 & 0.0002 & 0.0003 \end{bmatrix}$$

Continuous-time controllable Markov chains

Main definition

Definition

A controllable continuous-time Markov chain is a 4-tuple $CTMC = (S, A, \mathbb{K}, Q)$ where S is the finite state space $\left\{s_{(1)}, ..., s_{(N)}\right\}$, A is the set of actions: for each $s \in S$, $A(s) \subset A$ is the non-empty set of admissible actions at state $s \in S$, $\mathbb{K} = \left\{(s, a) \middle| s \in S, a \in A(s)\right\}$ is the class of admissible state-action pairs and $Q =$ is the **transition rates** $\left[q_{(j|i,k)}\right]$ with the elements defined as

$$q_{(j|i,k)} = \begin{cases} q_{(i|i,k)} = -\sum\limits_{i \neq j}^{N} q_{(j|i,k)} \leq 0 & \text{if} \quad i = j \\ \geq 0 & \text{if} \quad i \neq j \end{cases}$$

Fact

For each fixed k *the matrix of the transition rates is assumed to be* **conservative**, *i.e., from the definition above it follows that* $\sum_{j=1}^{N} q_{(j|i,k)} = 0$ *and* **stable**, *which means that* $q_{(i)} := \max_{a_{(k)} \in A\left(s_{(i)}\right)} q_{(j|i,k)} < \infty \quad \forall i.$

Example

$$q_{(j|i,k=1)} = \begin{bmatrix} -0.5366 & 0.0888 & 0.0611 & 0.1893 & 0.1409 \\ 0.0416 & -0.5689 & 0.0588 & 0.1331 & 0.0942 \\ 0.2358 & 0.1929 & -0.3784 & 0.1878 & 0.2084 \\ 0.0942 & 0.1929 & 0.1244 & -0.5963 & 0.0570 \\ 0.1649 & 0.0942 & 0.1342 & 0.0861 & -0.5005 \end{bmatrix}$$

Definition

Let $X_s := \{i \in \mathcal{X} : \mathrm{P}\{x(s,\omega) = i\} \neq 0,\ s \in \mathcal{T}\}$. For $s \leq t$ $(s, t \in \mathcal{T})$ and $i \in X_s$, $k \in A, j \in X$ define the conditional probabilities

$$\boxed{\pi_{j|i,k}(s,t) := \mathrm{P}\{x(t,\omega) = j \mid x(s,\omega) = i,\ a(s) = k\}}$$

which we will call the transition probabilities of a given Markov chain defining the conditional probability for a process $\{x(t,\omega)\}_{t\in\mathcal{T}}$ to be in the state j at time t under the condition that it was in the state i at time $s < t$ and in the same tame the decision $a(s) = k$ was done. The transition probabilities to be in the state j at time t under the condition that it was in the state i at time $s < t$

$$\boxed{\pi_{ij}(s,t \mid d) := \sum_{k=1}^{K} \pi_{j|i,k}(s,t) \underbrace{\mathrm{P}\{a(s) = k \mid x(s,\omega) = i\}}_{d_{k|i}} = \sum_{k=1}^{K} \pi_{j|i,k}(s,t)\, d_{k|i}}$$

The function $\pi_{ij}(s, t \mid d)$ for any $i \in X_s$, $j \in X$ and any $s \leq t$ $(s, t \in \mathcal{T})$ should satisfy the following *four conditions*:

1) $\pi_{ij}(s, t \mid d)$ is a conditional probability, and hence, is nonnegative, that is, $\pi_{i,j}(s, t) \geq 0$.
2) starting from any state $i \in X_s$ the Markov chain will obligatory occur in some state $j \in X_t$, i.e., $\sum_{j \in \mathcal{X}_t} \pi_{ij}(s, t \mid d) = 1$.
3) if no transitions, the chain remains to in its starting state with probability one, that is, $\pi_{ij}(s, s \mid d) = \delta_{ij}$ for any $i, j \in X_s$, $j \in X$ and any $s \in T$;
4) the chain can occur in the state $j \in X_t$ passing through any intermediate state $k \in X_u$ $(s \leq u \leq t)$, i.e.,

$$\boxed{\pi_{ij}(s, t \mid d) = \sum_{k \in \mathcal{X}_u} \pi_{ik}(s, u \mid d)\, \pi_{kj}(u, t \mid d)}$$

This relation is known as the Markov (or Chapman-Kolmogorov) equation.

Corollary

Since for **homogeneous Markov** *chains the transition probabilities* $\pi_{i,j}(s,t)$ *depend only on the difference* $(t-s)$, *below we will use the notation*

$$\boxed{\pi_{ij}(s-t \mid d) := \pi_{ij}(s,t \mid d)} \tag{2}$$

In this case the Markov equation becomes

$$\boxed{\pi_{i,j}(h_1+h_2 \mid d) = \sum_{k\in\mathcal{X}} \pi_{i,k}(h_1 \mid d)\,\pi_{k,j}(h_2 \mid d)} \tag{3}$$

valid for any $h_1, h_2 \geq 0$.

- Consider now the time τ (after the time s) just before changing the current state i, i.e., $\tau > s$.
- By the *homogeneity property* it follows that distribution function of the time τ_1 (after the time $s_1 := s + u$, $x(s + u, \omega) = i$) is the same as for the τ (after the time s, $x(s, \omega) = i$) that leads to the following identity

$$\mathsf{P}\{\tau > v \mid x(s, \omega) = i\} = \mathsf{P}\{\tau_1 > v \mid x(s_1, \omega) = i\}$$

$$\mathsf{P}\{\tau > v + u \mid x(s + u, \omega) = i\} =$$

$$\mathsf{P}\{\tau > u + v \mid x(s, \omega) = i, \tau > u \geq s\}$$

since the event $\{x(s, \omega) = i, \tau > u\}$ includes as a subset the event $\{x(s + u, \omega) = i\}$.

Lemma

The **expectation time** τ *(of the homogenous Markov chain* $\{x(t,\omega)\}_{t\in\mathcal{T}}$ *with a discrete phase space* X*) to be in the current state* $x(s,\omega)=i$ *before its changing has the exponential distribution*

$$\boxed{\mathsf{P}\left\{\tau > v \mid x(s,\omega)=i\right\} = e^{-\lambda_i v}} \tag{4}$$

where λ_i *is a nonnegative constant which inverse value characterizes* **the average expectation time** *before the changing the state* $x(s,\omega)=i$*, namely,*

$$\boxed{\frac{1}{\lambda_i} = \mathsf{E}\left\{\tau \mid x(s,\omega)=i\right\},\ \lambda_i = \left|q_{(i|i,k)}\right| = \sum_{i\neq j}^{N} q_{(j|i,k)}} \tag{5}$$

The constant λ_i *is usually called the* **"exit density"**.

Proof.

Define the function $f_i(u)$ as $f_i(u) := P\{\tau > u \mid x(s,\omega) = i\}$. By the Bayes formula

$$\begin{gathered}f_i(u+v) := P\{\tau > u+v \mid x(s,\omega) = i\} = \\ P\{\tau > u+v \mid x(s,\omega) = i,\, \tau > u\} P\{\tau > u \mid x(s,\omega) = i\} \\ = P\{\tau > u+v \mid x(s,\omega) = i,\, \tau > u\} f_i(u)\end{gathered}$$

By the homogeneous property one has

$$\begin{gathered}f_i(u+v) := P\{\tau > u+v \mid x(s,\omega) = i\} = \\ P\{\tau > v \mid x(s,\omega) = i\} f_i(u) = f_i(v) f_i(u)\end{gathered}$$

which means that

$$\begin{gathered}\ln f_i(u+v) = \ln f_i(u) + \ln f_i(v) \\ f_i(\tau = 0) = P\{\tau > 0 \mid x(s,\omega) = i\} = 1\end{gathered}$$

Proof.

[Continuation of the proof] Differentiation the logarithmic identity by u gives $\frac{f_i'(u+v)}{f_i(u+v)} = \frac{f_i'(u)}{f_i(u)}$ which for $u = 0$ becomes

$$\frac{f_i'(v)}{f_i(v)} = \frac{f_i'(0)}{f_i(0)} = f_i'(0) := -\lambda_i \rightarrow f_i(v) = e^{-\lambda_i v}$$

To prove (5) it is sufficient to notice that

$$\mathsf{E}\{\tau \mid x(s,\omega) = i\} = \int_{t=0}^{\infty} t d[-f_i(t)] =$$

$$\left[-te^{-\lambda_i t}\right]_{t=0}^{\infty} - \int_{t=0}^{\infty} \left[-e^{-\lambda_i t}\right] dt = \int_{t=0}^{\infty} e^{-\lambda_i t} dt = \lambda_i^{-1}$$

Lemma is proven. □

For homogenous Markov chains $\pi_{ij}(s,t \mid d) = \pi_{ij}(t-s \mid d)$ and stationary strategies $\mathsf{P}\{a(s) = k \mid x(s,\omega) = i\} = d_{k|i}$ the Markov equation becomes (taking $s = 0$)

$$\frac{d}{dt}\pi_{ij}(t \mid d) = -\left(\sum_{i}^{N} q_{(j|i,k)}\right)\pi_{ij}(t \mid d) + q_{(j|i,k)}\pi_{il}(t \mid d)$$

can be written as the matrix differential equation as follows:

$$\Pi'(t \mid d) = \Pi(t \mid d)\,Q(d)\,; \qquad \Pi(0) = I_{N\times N}$$

$$\Pi(t \mid d) = \|\pi_{i,k}(t \mid d)\| \in \mathbb{R}^{N\times N},\ \ Q(d) = \left\|\sum_{k=1}^{K} q_{(j|i,k)} d_{k|i}\right\|$$

This system can be solved by

$$\Pi(t \mid d) = \Pi(0)e^{Q(d)t} = e^{Qt} := \sum_{t=0}^{\infty} \frac{t^n Q^n(d)}{n!} \tag{6}$$

At the stationary state, the probability transition matrix is

$$\boxed{\Pi\left(d\right) = \lim_{t\to\infty} \Pi(t \mid d)}$$

Definition

The vector $P \in R^N$ $\left(\sum_{i=1}^{N} P_i = 1\right)$ is called **the stationary distribution vector** if

$$\Pi^\top\left(d\right) P = P$$

Claim

This vector can be seen as the long run proportion of time that the process is in state $s_{(i)} \in S$.

Theorem (Xianping Guo, Onesimo Hernandez Lerma, 2009)

The the stationary distribution vector P satisfies the linear equation

$$\boxed{Q^{\top}(d)\,P = 0} \tag{7}$$

Fact

The Production Optimization Problem, described by a continuous-time controllable Markov chain in stationary states, is **the same Linear programming problem** *(LPP) as for a discrete-time model but* **with the additional linear constraint** *(7).*

Conclusion

Which topics we have discussed today?

1-st Lecture Day: *Basic Notions on Controllable Markov Chains Models, Decision Making and Production Optimization Problem.*

- Markov Processes: Classical Definition (Markov), Mathematical Definition (Kolmogorov), Markov property in a general format
- Finite Markov Chains: Main definition, Homogeneous (Stationary) Markov Chains, Transition matrix, Dynamic Model of Finite Markov Chains,
  Ergodic Markov Chains, Ergodic Theorem, Ergodicity coefficient.
- Controllable Markov Chains: Transition matrices for controllable Finite Markov Chain processes, Pure and mixed strategies.
- Simplest Production Optimization Problem: c-variables, Linear Programming Problem.
- Continuous-time controllable Markov chains: Distribution function of the time just before changing the current state, the transition rates,
  expectation time, Additional linear constraint and LPP problem.

Next Lecture Day: *The Mean-Variance Customer Portfolio Problem: Bank Credit Policy Optimization.*

Thank you for your attention! See you soon!