Complex Probability and Markov Stochastic Process

Bijan Bidabad[1], Behrouz Bidabad

Abstract:
    This note discusses the existence of "complex probability" in the real world sensible problems. By defining a measure more general than conventional definition of probability, the transition probability matrix of discrete Markov chain is broken to the periods shorter than a complete step of transition. In this regard, the complex probability is implied.

1- Introduction

   Somtimes analytic numbers coincide with the mathematical modeling of real world and make the real analysis of problems complex. All the measures in our everyday problems belong to R, and mostly to R^+. Probability of occurrence of an event always belongs to the range [0,1]. In this paper, it is discussed that to solve a special class of Markov chain which should have solution in real world, we are confronted with "analytic probabilities"!. Though the name probability applies to the values between zero and one, we define a special analogue measure of probability as complex probability where the conventional probability is a subclass of this newly defined measure.

   Now define the well-known discrete time Markov chain $\{Y_n\}$ a Markov stochastic process whose state space is $s = \{1,2,...,N\}$ for which $T = \{0,1,2,...\}$. Refer to the value of Y_n as the outcome of the nth trial. We say Y_n being in state i if $Y_n = i$. The probability of Y_{n+1} being in state j, given that Y_n is in state i (called a one–step transition probability) is denoted by $P_{ij}^{n,n+1}$, i.e.,

$$P_{ij}^{n,n+1} = P_{ij} = \Pr\{Y_{n+1} = j \,|\, Y_n = i\} \tag{1}$$

Therefore, the Markov or transition probability matrix of the process is defined by

$$\mathbf{P} = \left\| P_{ij} \right\| \tag{2}$$

$$P_{ij} \geq 0 \quad \forall\, i,j \in s$$

$$\sum_{j=1}^{N} P_{ij} = 1 \quad \forall\, i \in s$$

---

[1]No.18, 12th St., Mahestan Ave., Shahrak Gharb, Tehran 14658-33659, Iran bijan_bidabad@msn.com

The n-step transition probability matrix $\mathbf{P}^{(n)} = \left\| P_{ij}^{n} \right\|$, which P_{ij}^{n} denotes the probability that the process goes from state i to state j in n transitions. Formally,

$$P_{ij}^{n} = \Pr\left\{ Y_{n+m} = j \mid Y_{m} = i \right\} \qquad \forall i, j \in S$$
(3)

According to Chapman – Kolmogorov relation for discrete Markov matrices (Karlin and Taylor (1975)), it can be proved that

$$\mathbf{P}^{(n)} = \mathbf{P}^{n} \qquad\qquad n \in N \ (Natural\ numbers) \qquad\qquad (4)$$

$\mathbf{P}^{n}$ that is $\mathbf{P}$ to the power n is a Markov matrix if $\mathbf{P}$ is Markov.

Now, suppose that we intend to derive the t-step transition probability matrix $\mathbf{P}^{(t)}$ where t≥0 from the above (3) and (4) definition of n-step transition probability matrix $\mathbf{P}$. That is, to find the transition probability matrix for incomplete steps. On the other hand, we are interested to find the transition matrix $\mathbf{P}^{(t)}$ when t is between two sequential integers. This case is not just a tatonnement example. To clarify the application of this phenomenon, consider the following example.

Example 1. Usually in population census of societies with N distinct regions, migration information is collected in an NxN migration matrix for a period of ten years. Denote this matrix by $\mathbf{M}$. Any element of $\mathbf{M}$, m_{ij} is the population who left region i and went to region j through the last ten years. By deviding each m_{ij} to sum of the ith row of $\mathbf{M}$, a value of P_{ij} is computed as an estimate of probability of transition from i^{th} to j^{th} regions. Thus, the stochastic matrix $\mathbf{P}$ gives the probabilities of going from region i to region j in ten years (which is one–step transition probability matrix). The question is: how we can compute the transition probability matrix for one year or one-tenth step and so on.

If we knew the generic function of probabilities in very small period of time we would be able to solve problems similar to example 1. But the generic function (5) is not obtainable. If it were, we would apply the continous time Markov procedure using the generic NxN matrix $\mathbf{A}$ as

$$\mathbf{A} = \lim_{h \to o^{+}} \frac{\mathbf{P}(h) - \mathbf{I}}{h} \qquad\qquad (5)$$

Where $\mathbf{P}(h)$ denotes transition probability matrix at time h. Then the transition probability matrix at any time $t \geq 0$ might be computed as follows. (Karlin and Taylor (1975)).

$$\mathbf{P}(t) = e^{\mathbf{A}t} \qquad\qquad (6)$$

Therefore a special procedure should be adopted to find the transition probability matrix $\mathbf{P}^{(t)}$ at any time t from discrete Markov chain information. As it will be show later the adopted procedure coincides with transition probability matrix with complex elements.

2- Breaking the time in discrete Markov chain

Consider again matrix **P** defined in (2). Also, assume **P** is of full rank.

Assumption 1: **P** is of full rank.
This assumption assures that all eigenvalues of **P** are nonzero and **P** is diagnalizable, Searle (1982), Dhrymes (1978). This assumption is not very restrictive, since; actually, most of Markov matrices have dominant diagonals. That is probability of transition from state i to itself is more than the sum of probabilities from state i to all other states. The matrices having dominant diagonals are non-singular, Takayama (1974). Threfore, **P** can be decomposed as follows (Searle (1982), Klein (1973)).

$$\mathbf{P} = \mathbf{X}\,\mathbf{\Lambda}\,\mathbf{X}^{-1} \tag{7}$$

Where **X** is an NxN matrix of eigenvectors $\mathbf{x}_i$, $i = 1,...,N$,

$$\mathbf{X} = [\mathbf{x}_1,....,\mathbf{x}_N] \tag{8}$$

and $\mathbf{\Lambda}$ the NxN diagonal matrix of corresponding eigenvalues,

$$\mathbf{\Lambda} = diag\,\{\lambda_1,...,\lambda_N\} \tag{9}$$

Using (7), (8) and (9) to break n-step transition probability matrix **P** to any smaller period of time $t \geq 0$, we do as follows. If $t_i \leq 0$ for all i$\in\{1,...,K\}$ are fractions of n–step period and $\sum_{i=1}^{k} t_i = n$ for any n belonging to natural numbers then,

$$\mathbf{P}^n = \prod_{j=1}^{k} \mathbf{P}^{t\,i} = \mathbf{P}^{\sum_{i=1}^{k} t_i} \tag{10}$$

On the other hand, transition probability matrix of n-step can be broken to fractions of n, if sum of them is equal to n. Therefore, any $t \geq 0$ fraction of one-step transition probability matrix can be written as,

$$\mathbf{P}^t = \mathbf{X}\mathbf{\Lambda}^t\,\mathbf{X}^{-1} \tag{11}$$

where,

$$\mathbf{\Lambda}^t = diag\,\{\lambda_1^t,....,\lambda_N^t\} \tag{12}$$

Before discussing on the nature of eigenvalues of **P** let us define the generalized Markov matrix.

Definition 1. Matrix **Q** is a generalized Markov matrix if the following conditions are fulfilled:

1) $q_{ij} \in C$ $\forall\, i\,,j \in S$

2) $\text{Re}\,(q_{ij}) \in [0,1]$ $\forall i\,,j \in S$

3) $\text{Im}(q_{ij}) \in [-1,1]$ $\forall i\,,j \in S$

4) $\sum\limits_{j=1}^{N} \text{Re}\,(q_{ij}) = 1$ $\forall i \in S$

5) $\sum\limits_{j=1}^{N} \text{Im}(q_{ij}) = 0$ $\forall i \in S$

<u>Remark 1</u>. According to definition 1, matrix **Q** can be written as:

$$\mathbf{Q} = \mathbf{U} + i\mathbf{V}$$
(13)

Where **U** and **V** are NxN matrices of real and imaginary parts of **Q** with $i = \sqrt{-1}$.

<u>Remark 2</u>. Matrix **U** has all Properties of **P** defined by (2), thus, $\mathbf{P} \subset \mathbf{Q}$.

<u>Treorem 1</u>. If **P** is a Markov matrix then $\mathbf{P}^t$ also satisfies Markovian properties.
Proof: According to Chapman–Kolmogorov relation for continous Markov chain (Karlin and Taylor (1975)), we have

$$\mathbf{P}(t+s) = \mathbf{P}(t)\,\mathbf{P}(s) \qquad t,s \geq 0 \tag{14}$$

That is, if **P**(t) and **P**(s), transition probability matrices at times t and s are Markovs, then the product of them **P**(t+s) is also Markov. Let t=1, then **P**(1) is a one-step transition probability matrix which is equivanlent to (2). Hence, our discrete Markov matrix **P** is equivalent to its contionous analogue **P**(1). So

$$\mathbf{P} = \mathbf{P}(1) \tag{15}$$

If we show that

$$\mathbf{P}^t = \mathbf{P}(t) \tag{16}$$

Then according to (14)

$$\mathbf{P}^{t+s} = \mathbf{P}^t\,\mathbf{P}^s \tag{17}$$

We can conclude that if **P** is Markov then $\mathbf{P}^t$, $\mathbf{P}^s$ and $\mathbf{P}^{t+s}$ are also Markovs for $t,s \geq 0$ and the theorem is proved.

Rewrite **P**(t) in (6) as (18).

$$\mathbf{P}(t) = \mathbf{X}\,\boldsymbol{\Lambda}(t)\,\mathbf{X}^{-1} \tag{18}$$

4

Where $\lambda_i, i \in S$ are the eigenvalues of **A** defined by (5), and

$$\mathbf{\Lambda}(t) = diag\left\{\exp(\lambda_1^t),...,\exp(\lambda_N^t)\right\} \tag{19}$$

And **X** is the corresponding eigenmatrix of **A**. Take the natural logarithm of (18),

$$ln\,\mathbf{P}(t) = \mathbf{X}\mathbf{\Phi}(t)\mathbf{X}^{-1} \tag{20}$$

Where,

$$\mathbf{\Phi}(t) = t\,diag\left\{\lambda_1,...,\lambda_N\right\} \tag{21}$$

So,

$$ln\,\mathbf{P}(t) = t\,\mathbf{X}\mathbf{\psi}\mathbf{X}^{-1} \tag{22}$$

where

$$\mathbf{\psi} = diag\left\{\lambda_1,...,\lambda_N\right\} \tag{23}$$

Write (22) for t=1 and multiply both side by t,

$$t\,ln\,\mathbf{P}(1) = t\,\mathbf{X}\mathbf{\psi}\mathbf{X}^{-1} \tag{24}$$

By comparison of (22) and (24) conclude that

$$ln\,\mathbf{P}(t) = t\,ln\,\mathbf{P}(1) \tag{25}$$

or,

$$\mathbf{P}(t) = \mathbf{P}^t(1) \tag{26}$$

Given (15), equation (26) is the same as (16) Q.E.D.

<u>Result 1.</u> Matrix $\mathbf{P}^t$ fulfils definition 1. Thus, $\mathbf{P}^t \subseteq \mathbf{Q}$. This comes from the following remarks.

<u>Remark 3</u>. Sum of each row of $\mathbf{P}^t$ is equal to one. Since $\mathbf{P}^t$ satisfies Markovian properties (theorem 1).

<u>Remark 4</u>. Sum of imaginary parts of each row is equal to zero. This immediately comes from remark 3.

<u>Remark 5</u>. If q_{ij} denotes the ijth element of $\mathbf{P}^t$ for $t \geq 0$, then $\left|q_{ij}\right| \leq 1$ for all i and j belonging to S. This remark can be concluded form theorem 1.

<u>Remark 6</u>. If $\mathbf{Q} = \mathbf{P}^t, t \geq 0$ equals to the complex matrix defined by (13), then $\left|V_{jk}\right| \leq 1 \quad \forall j,k \in S$. Since,

$$1 \geq \left| q_{jk} \right| = \left| u_{jk} + iv_{jk} \right| \Rightarrow 1 \geq \sqrt{u^2_{jk} + iv^2_{jk}}$$
$$1 \geq u^2_{jk} + v^2_{jk} \Rightarrow 1 \geq \left| v_{jk} \right|.$$

<u>Remark 7</u>. Given **Q** as in remark 6, then $u_{jk} \in [0,1]$. This also comes immediately from theorem 1.

3. Discussion on broken times

The broken time discrete Markov chain is not always a complex probability matrix defined by definition 1. Matrix $\mathbf{P}^t$ has different properties with respect to t and eigenvalues. λ_i may be real (positive or negative) or complex depending on the characteristic polynomial of **P**.

Since **P** is a non–negative matrix, Forbenius theorem (Takayama (1974), Nikaido (1970)) assures that **P** has a positive dominant eigenvalue

$$\lambda_1 \succ 0 \quad \text{(Frobenius root)}$$
$$(27)$$
and
$$\left| \lambda_i \right| \leq \lambda_1 \qquad \forall i \in \{2, ..., N\} \tag{28}$$

Furthermore, if **P** is also a Markov matrix then its Frobenius root is equal to one, (Bellman (1970), Takayama (1974)). Therefore,

$$\lambda_1 = 1 \tag{29}$$
$$\left| \lambda_i \right| \leq 1 \qquad \forall i \in S \tag{30}$$

With the above information, consider the following discussions.

$$a) \; \lambda_i \in (0,1] \qquad \forall i \in S$$

In this case all $\lambda_i^t \geq 0$ for $t \geq 0$ and no imaginary part occures in matrix $\mathbf{P}^t$. λ_i are all positive for i belonging to S if we can decompose the matrix **P** to two positive semi-definite and positive definite matrices **B** and **C** of the same size (Mardia, Kent, Bibby (1982)) as

$$\mathbf{P} = \mathbf{C}^{-1} \mathbf{B}$$
$$b) \; \lambda_i \in [-1,1], \lambda_i \neq 0, \forall i \in S$$

λ_i^t, $t \geq 0$ belongs to sets of real and imaginary numbers based on the value of t. In this case $\mathbf{P}^t$ belongs to the class of generalized stochastic matrix **Q** of definition 1. For $\lambda_i \in R$, it is sufficient that **P** be positive definite.

$$c)\,\lambda_i \in C\,, \left|\lambda_i\right| \in (0,1]\ \forall\, i \in S$$

$\mathbf{P}^t$ in this case for $t \geq 0$ and $t \notin N$ belongs to the class of generalized Morkov matrices of definition 1.

$$d)\,t \in N\ \text{(Natural numbers)}$$

In all cases of a, b, and c we never coincide with complex probabilities. Since $\mathbf{P}^t$ can be drived by simply multiplying $\mathbf{P}$, t times.

$$e)\,t \in Z\ \text{(integer numbers)}$$

In this case, $\mathbf{P}^t$ is a real matrix but does not always satisfy condition 2 of definition 1.
$$f)\,t \in R^-$$

$\mathbf{P}^t$ is a complex matrix but does always satisfy conditions 2 and 3 of definition 1.

4. Complex probability justification

Interpretation of the "Complex probability" as defined by definition 1 is not very simple and needs more elaborations. The interesting problem is that, it exists in operational works of statistics as the example 1 discussed. Many similar examples like the cited may be gathered.

With this definition of probability, the moments of a real random variable are complex. Although the t–step distribution π_t of initial distribution π_0 with respect to $\mathbf{P}^t$ may be complex, they have the same total as π_0. That is, if

$$\boldsymbol{\pi}_0 = (\pi_{01},...,\pi_{0N}) \tag{32}$$

Then,

$$\boldsymbol{\pi}_t = \boldsymbol{\pi}_o\, \mathbf{P}^t = \boldsymbol{\pi}_o\, \mathbf{Q} = \boldsymbol{\pi}_o\, \mathbf{U} + i\, \boldsymbol{\pi}_o\, \mathbf{V} \tag{33}$$

And we have the following remark accordingly,

Remark 8. Sum of t-step distribution is equal to sum of initial distribution. That is,

$$\sum_{j=1}^{N} \pi_{oj} = \sum_{j=1}^{N} \pi_{tj} \tag{34}$$

This can be derived based on (32) and (33) as

$$(\pi_{t1},...,\pi_{tN}) = \left(\sum_{j=1}^{N} \pi_{oj} U_{j1},..., \sum_{j=1}^{N} \pi_{oj} U_{jn}\right) + i\left(\sum_{j=1}^{N} \pi_{oj} v_{ji},..., \sum_{j=1}^{N} \pi_{oj} v_{jN}\right) \tag{35}$$

And, sum of t–step distribution is

$$\sum_{j=1}^{N} \pi_{tj} = \sum_{j=1}^{N} \pi_{oj} (u_{j1} + ... + u_{jN}) + i \sum_{j=1}^{N} \pi_{oj} (v_{j1},...,v_{jn}) \tag{36}$$

The two parentheses in (36) are one and zero respectively based on conditions 4 and 5 of definition 1. Thus, (36) and (34) are the same.

The above remark 8 states that though there exist imaginary transition probabilities to move from state j to k, the total sum of "imaginary transitions" is equal to zero. On the other hand, after t^{th} step transition, the total distribution has no imaginary part.

5. Summary

By summarizing the discrete and continous times Markov stochastic processes a class of real world problems was introduced which can not be solved by each of the procedures. The solutions of these problems coincide with "Complex probabilities" of transitions that are inherent in mathematical formulation of the model. Complex probability is defined and some of its properties with respect to the cited class are examined. Justification of the idea of complex probability needs more work and is left for further research.

6. Acknowledgements

7. Refrences

❖ R. Bellman (1970), Introduction to matrix analysis, McGraw–Hill.
❖ P.J. Dhrymes (1978), Mathematics for econometrics. Springer-Verlag.
❖ W. Feller (1970, 1971), An Introduction to probability theory and its applications, vols. 1,2, Wiley, New York.
❖ P.G. Hoel, S.C. Port, C. J. Stone (1972), Introduction to stochastic processes. Houghton Mifflin, New York.
❖ S. Karlin, H.M.Taylor (1975), A first course in stochastic processes. Academic Press.
❖ E. Klein (1973), Mathematical methods in theoretical economics, topological and vector space foundations of equilibrium analysis, Academic Press.
❖ K.V. Mardia, J.T. Kent, J.M. Bibby (1982), Multivariate analysis, Academic Press.
❖ H. Nikaido (1970), Introduction to sets and mapping in modern economics. North Holland Publishing Co.
❖ S.S. Searle (1982), Matrix algebra useful for statistics. Wiley.
❖ A.Takayama (1974) Mathematical economics. Dyrden Press, Illinois.
❖ E. Wentzel, L. Ovcharov (1986) Applied problems in probability theory. Mir, Moscow.