Estimates of Earthquake With Markov Models in the East Anatolian Fault Zone

Adem DOĞANER and Sinan ÇALIK
Fırat University, Department of Statistics, Elazığ, Turkey
adoganer@firat.edu.tr
(Geliş/Received: 03.08.2012; Kabul/Accepted: 26.09.200*)

Abstract
Stochastic methods are often used for estimates of future for environmental events realized. Markov chains are one of these methods. Markov chains can applied to estimates of the earthquake but it provides only information about a single parameter. In this article, epicenter of may occurring earthquake were estimated on East Anatolian Fault Zone. Epicenter parameter and times of seismic inactivity were associated for estimates of earthquake. Probabilities were provided by Hidden Markov Model with forwards algorithm and these probabilities were converted to state sequence. Transition probabilities of this sequence were examined with Discrete Time Markov Chains. As a result new earthquake were expected in 4th state (Sincik- Lake Hazar ) with 0.73 probabilities. Results and method were discussed.

Key words: Hidden Markov model, Epicenter, East Anatolian Fault Zone

Markov Modelleri ile Doğu Anadolu Fay Zonu’ndaki Deprem Tahminleri

Özet

Anahtar kelimeler: Saklı Markov modeli, Dış merkez, Doğu Anadolu Fay Zonu

1. Introduction
Earthquakes are seismic activities that occurred in the region caused great destruction and loss. Estimates of earthquake are a great importance for to minimize losses, but estimates of earthquakes precisely is not possible. Some statistical methods can be used for estimate of earthquake. Stochastic processes are one of these methods and were studied on these methods. Poisson processes and point processes were used by Ogata [1] and by Ferraes [2] for estimates of earthquake occurrences. Markov Chain can be used such as Poisson process for estimate of earthquake. Application of Markov Chains to geologic processes were discussed by Harbaugh and Bonham-Carter [3]. Markov Chains were applied by Nava et al. [4] to Japan region for the evolution of seismic risk. Besides Heng Tsai [5] used Markov Chains for estimates of earthquake recurrence. But Markov Chains only gives transition probabilities of a single parameter. Hidden Markov model have two processes known as sequence of unobserved state and sequence of observation. Ebel et al.[6] used Hidden Markov Model for earthquake forecasting in California. East Anatolian Fault System, which produce earthquake is an active fault system. Risk assessment for this fault is important.

In this study, probabilities of earthquake occurrence in epicenters of East Anatolian Fault
Zone were estimated by using hidden Markov model and forward algorithm. Transition probabilities of epicenters were estimated by using discrete parameter Markov chains. Relationship was obtained between epicenter and times of seismic inactivity.

2. Material and Methods

2.1. Data

East Anatolian Fault Zone, known as Türkoğlu-Antakya segment, Gölbasi-Türkoğlu Segment, Çelikhan-Erkenek Segment, Lake Hazar-Sincik Segment, Palu-Lake Hazar Segment and Karløva-Bingöl Segment, is composed six segments. Occurred earthquakes in area radius of 20 km the East Anatolian Fault Zone were included in this study. Earthquake data were supplied from Boğaziçi University Kandilli Observatory and Earthquake Research Institute National Earthquake Monitoring Centre. Data were included for $M_L \geq 4.0$ earthquakes occurred between 1975 and 2011.

In this study, data were included about known as one of earthquake parameters epicenter and times of seismic inactivity. The parameter is epicenter. Epicenters were categorized according to segments of East Anatolian Fault Systems. Epicenters were represented as $S_1 = \text{Türkoğlu - Antakya segment}$, $S_2 = \text{Gölbasi - Türkoğlu segment}$, $S_3 = \text{Çelikhan - Erkenek segment}$, $S_4 = \text{Lake Hazar - Sincik segment}$, $S_5 = \text{Palu - Lake Hazar segment}$ and $S_6 = \text{Karløva - Bingöl segment}$. In this study, times of inactivity were also considered. Times of inactivity were categorized as $O_1(\text{Time} = 0 - 90 \text{ days})$, $O_2(\text{Time} = 91 - 180 \text{ days})$, $O_3(\text{Time} = 181 - 270 \text{ days})$, $O_4(\text{Time} = 271 - 360 \text{ days})$, $O_5(\text{Time} \geq 361 \text{ days})$, respectively.

229 Seismic activity with $M_L \geq 4$ were occurred in fault zone between 1975 and 2011. Aftershocks occurs that depend on main shocks. Aftershocks were excluded in this study, therefore each random variables must be independent the other according to Markov property. 126 earthquakes were included in this study.

![Figure 1. Study Area (East Anatolian Fault Zone and segments)](image)

2.2. Parameter Relationship

In this study, between the parameters of earthquake statistical relationship was examined. Transitions of earthquake epicenter in East Anatolian Fault Zone were wanted to estimate. Outputs of earthquake were considered, since transitions of earthquake epicenter cannot directly observed. Time of seismic inactivity is an output of earthquake. The relationship between epicenter parameter and times of seismic inactivity were analyzed with Chi-Squared test for independence. Results of analysis were represented table 1. The relationship between epicenter parameter and times of seismic inactivity statistically were significant.

<table>
<thead>
<tr>
<th>Table 1. Parameter Relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X^2$</td>
</tr>
<tr>
<td>Epicenter*Times of seismic inactivity</td>
</tr>
</tbody>
</table>

2.3. Markov Chains and Hidden Markov Model

2.3.1. Theoretical Background

Markov Chains base on, when state $i$ at the time $t$ were known, state $j$ at the time $t + 1$ depend on only state $i$ at the time $t$. In other words, a stochastic process with the Markov property is expressed by:

$$P(X_{n+1} = j|X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) =$$
This equation often is called as memorylessness [7].

Hidden Markov model is Markov Chains having additional features. Theory of Hidden Markov Model was developed by Baum et al. [8,9,10,11,12]. Hidden Markov model has a two stochastic processes. Each Hidden Markov Model is defined by states, state probabilities, transition probabilities, emission probabilities and initial probabilities [13,15].

Hidden Markov model were characterized by the following [14],
1. The N states of the Model, defined by \( S = \{ S_1, S_2, ..., S_N \} \).
2. M observation symbols per state \( V = \{ v_1, v_2, ..., v_m \} \),
3. The State transition probability distribution \( A = \{ a_{ij} \} \), where \( a_{ij} \) is the probability that the state at the time \( t + 1 \) is \( S_j \) is given when the state at time \( t \) is \( S_i \).
4. The observation symbol probability distribution in each state, \( B = \{ b_j(k) \} \), where \( b_j(k) \) is the probability that symbol \( v_k \) is emitted in state \( S_j \).
   \[ b_j(k) = p(\alpha_t = v_k | q_t = j), \]
   \[ 1 \leq j \leq N, \quad 1 \leq k \leq M \]  

Where \( \alpha_t \) denotes the \( k^{th} \) observation symbol in the alphabet, and \( o_t \) the current parameter vector.

5. The Hidden Markov Model is the initial state distribution \( \pi = p(\pi_i) \), where \( \pi_i \) is the probability that the model is in state \( S_i \) at the time \( t = 0 \) with
   \[ \pi_i = p(\pi_i = i) \text{ and } 1 \leq i \leq N \]

Discrete parameter hidden Markov model generally is expressed as \( \lambda = (A, B, \pi) \) [13,14,15].

2.4. Application

In this study, 126 earthquakes and epicenter data and times of inactivity about this earthquake were examined in order to epicenter data and times of inactivity adjust to Hidden Markov Model.

First step, transition probabilities matrix of state, emission probabilities matrix and initial probabilities vector were provided about Hidden Markov model. Transition probabilities matrix of state, emission probabilities matrix and initial probabilities vector were showed in Table 2.

Transition probabilities from each state to others state were given in transition probabilities matrix. Example transition probability from State \( S_1 \) known as Türkoğlu-Antakya segment to State \( S_4 \) known as Hazar Lake-Sincik segment were determined as 0.80.

The observation symbols probabilities for each state were given in emission probabilities matrix. Example State 1 known as Türkoğlu-Antakya segment produced 0.60 probability observation symbol \( O_1 \) and 0.40 probability observation symbol \( O_3 \).

Initial state probabilities equally for each state were distributed since initial state probabilities distribution were not known. Each initial state have 0.16666667 probability.

\[
\begin{array}{ccccccc}
\text{a. Transition Probabilities Matrix} \\
S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \\
\hline
S_1 & 0 & 0.20 & 0.80 & 0 & 0 & 0 \\
S_2 & 0 & 0.20 & 0 & 0.40 & 0 & 0.40 \\
S_3 & 0.07 & 0.07 & 0.33 & 0.20 & 0.13 & 0.20 \\
S_4 & 0.02 & 0 & 0.10 & 0.54 & 0.15 & 0.19 \\
S_5 & 0.08 & 0.04 & 0.08 & 0.16 & 0.52 & 0.12 \\
S_6 & 0.04 & 0.07 & 0.07 & 0.33 & 0.16 & 0.33 \\
\end{array}
\]

\[
\begin{array}{cccc}
\text{b. Emission Probabilities Matrix} \\
S_1 & S_2 & S_3 \\
\hline
S_1 & 0.60 & 0.40 \\
S_2 & 0.60 & 0.20 \\
S_3 & 0.54 & 0.20 \\
S_4 & 0.77 & 0.13 \\
S_5 & 0.54 & 0.42 \\
S_6 & 0.92 & 0.08 \\
\end{array}
\]

\[
\begin{array}{cccccc}
\text{c. Initial Probabilities Vector} \\
S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \\
\hline
\pi = [0.166 \quad 0.166 \quad 0.166 \quad 0.166 \quad 0.166 \quad 0.167] \\
\end{array}
\]
2.5. Forward Algorithm

2.5.1. Theoretical Background

The variable $\alpha_t(i)$ was defined as probability of state $S_t$ is situated as a result of partial observation until at the time $t$ [14].

Equation is defined as:

$$\alpha_t(i) = P(O_1, O_2, ..., O_t, q_t = S_t | \lambda)$$  \hspace{1cm} (3)

Forward algorithm contain three parts [14].

Initialization:

$$\alpha_1(i) = \pi_i b_i(O_1) \hspace{1cm} 1 \leq i \leq N \hspace{1cm} (4)$$

Induction:

$$\alpha_{t+1}(j) = \sum_{i=1}^{N} \alpha_t(i) a_{ij} b_j(O_{t+1}) \hspace{1cm} (5)$$

Termination:

$$P(O | \lambda) = \sum_{i=1}^{N} \alpha_T(i) \hspace{1cm} (6)$$

The purpose of use of forward algorithm in this study, realization probabilities of state sequence in accordance with hidden Markov model were wanted to estimate by observation sequence. Result of forward algorithm showed in Table 3.

Table 3. State Probability for earthquake sequence and result of Forward algorithm

<table>
<thead>
<tr>
<th>Sequence</th>
<th>States</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>124</th>
<th>125</th>
</tr>
</thead>
<tbody>
<tr>
<td>S_1</td>
<td>0.1000</td>
<td>0.0133</td>
<td>0.0101</td>
<td>0.0076</td>
<td>0.0000</td>
<td>.</td>
<td>.</td>
<td>4.77e-49</td>
</tr>
<tr>
<td>S_2</td>
<td>0.1000</td>
<td>0.0243</td>
<td>0.0116</td>
<td>0.0078</td>
<td>0.0018</td>
<td>.</td>
<td>.</td>
<td>4.80e-49</td>
</tr>
<tr>
<td>S_3</td>
<td>0.0900</td>
<td>0.0434</td>
<td>0.0278</td>
<td>0.0199</td>
<td>0.0053</td>
<td>.</td>
<td>.</td>
<td>1.23e-48</td>
</tr>
<tr>
<td>S_4</td>
<td>0.1283</td>
<td>0.2096</td>
<td>0.1500</td>
<td>0.1076</td>
<td>0.0130</td>
<td>.</td>
<td>.</td>
<td>6.61e-48</td>
</tr>
<tr>
<td>S_5</td>
<td>0.0900</td>
<td>0.0552</td>
<td>0.0469</td>
<td>0.0359</td>
<td>0.0205</td>
<td>.</td>
<td>.</td>
<td>2.26e-48</td>
</tr>
<tr>
<td>S_6</td>
<td>0.1533</td>
<td>0.1322</td>
<td>0.0998</td>
<td>0.0711</td>
<td>0.0044</td>
<td>.</td>
<td>.</td>
<td>4.34e-48</td>
</tr>
</tbody>
</table>

2.6. Model Evolution

Hidden Markov model and forward algorithm were applied to Earthquake data until this stage of study.

At the stage of model evolution, probabilities of state by forward algorithm were examined. In a result of forward algorithm, a state having highest probability for each earthquake in sequence were included as a new
Estimates of Earthquake With Markov Models in the East Anatolian Fault Zone

epicenter of earthquake to our model. Example, State $S_6$ has the highest probability for first earthquake in table 3. So a new state sequence were obtained. Obtaining this new state sequence were shown figure 3.

Table 3. New State Sequence

<table>
<thead>
<tr>
<th>S.</th>
<th>State</th>
<th>S.</th>
<th>State</th>
<th>S.</th>
<th>State</th>
<th>S.</th>
<th>State</th>
<th>S.</th>
<th>State</th>
<th>S.</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$S_1$</td>
<td>21</td>
<td>$S_2$</td>
<td>41</td>
<td>$S_3$</td>
<td>61</td>
<td>$S_4$</td>
<td>81</td>
<td>$S_5$</td>
<td>101</td>
<td>$S_6$</td>
</tr>
<tr>
<td>2</td>
<td>$S_1$</td>
<td>22</td>
<td>$S_2$</td>
<td>42</td>
<td>$S_3$</td>
<td>62</td>
<td>$S_4$</td>
<td>82</td>
<td>$S_5$</td>
<td>102</td>
<td>$S_6$</td>
</tr>
<tr>
<td>3</td>
<td>$S_1$</td>
<td>23</td>
<td>$S_2$</td>
<td>43</td>
<td>$S_3$</td>
<td>63</td>
<td>$S_4$</td>
<td>83</td>
<td>$S_5$</td>
<td>103</td>
<td>$S_6$</td>
</tr>
<tr>
<td>4</td>
<td>$S_1$</td>
<td>24</td>
<td>$S_2$</td>
<td>44</td>
<td>$S_3$</td>
<td>64</td>
<td>$S_4$</td>
<td>84</td>
<td>$S_5$</td>
<td>104</td>
<td>$S_6$</td>
</tr>
<tr>
<td>5</td>
<td>$S_1$</td>
<td>25</td>
<td>$S_2$</td>
<td>45</td>
<td>$S_3$</td>
<td>65</td>
<td>$S_4$</td>
<td>85</td>
<td>$S_5$</td>
<td>105</td>
<td>$S_6$</td>
</tr>
<tr>
<td>6</td>
<td>$S_1$</td>
<td>26</td>
<td>$S_2$</td>
<td>46</td>
<td>$S_3$</td>
<td>66</td>
<td>$S_4$</td>
<td>86</td>
<td>$S_5$</td>
<td>106</td>
<td>$S_6$</td>
</tr>
<tr>
<td>7</td>
<td>$S_1$</td>
<td>27</td>
<td>$S_2$</td>
<td>47</td>
<td>$S_3$</td>
<td>67</td>
<td>$S_4$</td>
<td>87</td>
<td>$S_5$</td>
<td>107</td>
<td>$S_6$</td>
</tr>
<tr>
<td>8</td>
<td>$S_1$</td>
<td>28</td>
<td>$S_2$</td>
<td>48</td>
<td>$S_3$</td>
<td>68</td>
<td>$S_4$</td>
<td>88</td>
<td>$S_5$</td>
<td>108</td>
<td>$S_6$</td>
</tr>
<tr>
<td>9</td>
<td>$S_1$</td>
<td>29</td>
<td>$S_2$</td>
<td>49</td>
<td>$S_3$</td>
<td>69</td>
<td>$S_4$</td>
<td>89</td>
<td>$S_5$</td>
<td>109</td>
<td>$S_6$</td>
</tr>
<tr>
<td>10</td>
<td>$S_1$</td>
<td>30</td>
<td>$S_2$</td>
<td>50</td>
<td>$S_3$</td>
<td>70</td>
<td>$S_4$</td>
<td>90</td>
<td>$S_5$</td>
<td>110</td>
<td>$S_6$</td>
</tr>
<tr>
<td>11</td>
<td>$S_1$</td>
<td>31</td>
<td>$S_2$</td>
<td>51</td>
<td>$S_3$</td>
<td>71</td>
<td>$S_4$</td>
<td>91</td>
<td>$S_5$</td>
<td>111</td>
<td>$S_6$</td>
</tr>
<tr>
<td>12</td>
<td>$S_1$</td>
<td>32</td>
<td>$S_2$</td>
<td>52</td>
<td>$S_3$</td>
<td>72</td>
<td>$S_4$</td>
<td>92</td>
<td>$S_5$</td>
<td>112</td>
<td>$S_6$</td>
</tr>
<tr>
<td>13</td>
<td>$S_1$</td>
<td>33</td>
<td>$S_2$</td>
<td>53</td>
<td>$S_3$</td>
<td>73</td>
<td>$S_4$</td>
<td>93</td>
<td>$S_5$</td>
<td>113</td>
<td>$S_6$</td>
</tr>
<tr>
<td>14</td>
<td>$S_1$</td>
<td>34</td>
<td>$S_2$</td>
<td>54</td>
<td>$S_3$</td>
<td>74</td>
<td>$S_4$</td>
<td>94</td>
<td>$S_5$</td>
<td>114</td>
<td>$S_6$</td>
</tr>
<tr>
<td>15</td>
<td>$S_1$</td>
<td>35</td>
<td>$S_2$</td>
<td>55</td>
<td>$S_3$</td>
<td>75</td>
<td>$S_4$</td>
<td>95</td>
<td>$S_5$</td>
<td>115</td>
<td>$S_6$</td>
</tr>
<tr>
<td>16</td>
<td>$S_1$</td>
<td>36</td>
<td>$S_2$</td>
<td>56</td>
<td>$S_3$</td>
<td>76</td>
<td>$S_4$</td>
<td>96</td>
<td>$S_5$</td>
<td>116</td>
<td>$S_6$</td>
</tr>
<tr>
<td>17</td>
<td>$S_1$</td>
<td>37</td>
<td>$S_2$</td>
<td>57</td>
<td>$S_3$</td>
<td>77</td>
<td>$S_4$</td>
<td>97</td>
<td>$S_5$</td>
<td>117</td>
<td>$S_6$</td>
</tr>
<tr>
<td>18</td>
<td>$S_1$</td>
<td>38</td>
<td>$S_2$</td>
<td>58</td>
<td>$S_3$</td>
<td>78</td>
<td>$S_4$</td>
<td>98</td>
<td>$S_5$</td>
<td>118</td>
<td>$S_6$</td>
</tr>
<tr>
<td>19</td>
<td>$S_1$</td>
<td>39</td>
<td>$S_2$</td>
<td>59</td>
<td>$S_3$</td>
<td>79</td>
<td>$S_4$</td>
<td>99</td>
<td>$S_5$</td>
<td>119</td>
<td>$S_6$</td>
</tr>
<tr>
<td>20</td>
<td>$S_1$</td>
<td>40</td>
<td>$S_2$</td>
<td>60</td>
<td>$S_3$</td>
<td>80</td>
<td>$S_4$</td>
<td>100</td>
<td>$S_5$</td>
<td>120</td>
<td>$S_6$</td>
</tr>
</tbody>
</table>

The latest stage of study, transition probabilities were estimated for new state sequence. Transition probabilities matrix were included by discrete parameter Markov Chains method. Obtaining probabilities by discrete parameter Markov Chains give to transition probabilities of between the epicenter. According to probabilities in this matrix, Earthquake may occur with 0.80 probability in state $S_4$ known as Hazar Lake-Sincik segment, after earthquake occurred in state $S_1$ known as Antakya-Türkoglu segment. Initial probabilities were estimated for new state sequence. Initial probabilities gives distribution of epicenter in new state sequence.
The product of initial probabilities vector and transition probabilities matrix gives probability distribution vector of epicenter for the next earthquake. Transition probabilities matrix, initial probabilities vector and probability distribution vector of epicenter for the next earthquake were represented \( P_0 \), \( \pi_0 \) and \( P_1 \) respectively. \( P_0 \), \( \pi_0 \) and \( P_1 \) were shown in Table 4.

**Table 4. Result Table**

<table>
<thead>
<tr>
<th>Transition Probabilities Matrix.</th>
<th>( S_1 )</th>
<th>( S_4 )</th>
<th>( S_5 )</th>
<th>( S_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_1 )</td>
<td>0.00</td>
<td>0.80</td>
<td>0.20</td>
<td>0.00</td>
</tr>
<tr>
<td>( S_4 )</td>
<td>0.03</td>
<td>0.76</td>
<td>0.18</td>
<td>0.03</td>
</tr>
<tr>
<td>( S_5 )</td>
<td>0.09</td>
<td>0.56</td>
<td>0.35</td>
<td>0.00</td>
</tr>
<tr>
<td>( S_6 )</td>
<td>0.00</td>
<td>0.80</td>
<td>0.00</td>
<td>0.20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial Probabilities Vector.</th>
<th>( \pi_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_0 )</td>
<td>[0.04 0.73 0.19 0.04]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Next step transition matrix.</th>
<th>( \pi_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_1 )</td>
<td>[0.04 0.73 0.20 0.03]</td>
</tr>
</tbody>
</table>

According to results, next earthquake was estimated to occur with 0.73 probability in State 4 known as Hazar Lake-Sincik segment.

### 3. Results

In study, Hidden Markov Model and forward algorithm were applied earthquake data of East Anatolian Fault Zone between 1975 and 2011. Transition probabilities matrix, initial probabilities vector, emission probabilities matrix and new state sequence were obtained. Discrete parameter Markov chains method was applied to new state sequence. Transition probabilities from each epicenter to others epicenter were calculated. Epicenter of next earthquake was estimated as a probabilistic.

According to obtained result, Earthquake may occur with 0.80 probability in State 4 after an earthquake occurred in State 1, with 0.76 probability in State 4 after a earthquake occurred in State 4, with 0.57 probability in State 4 after a earthquake occurred in State 5, with 0.80 probability in State 4 after a earthquake occurred in State 6.

As a result next earthquake were estimated to occur with 0.73 probability in State 4. Emission probabilities matrix show probabilities of times of inactivity for each epicenter. According to this matrix, earthquake will occur with 0.54 probabilities within between 1 and 90 days after an earthquake occurred in State 5. The latest earthquake in East Anatolian Fault Zone occurred in State 5.

In this study, Hidden Markov model methods were shown to be available for epicenter estimates.

### 4. Discussion

In seismic surveys, estimates of earthquake epicenter were been point of interest. Estimates of Earthquake can be used different methods. Some statistical methods such as Gumbell distribution can be used estimates of magnitude and determination of earthquake risk.

Discrete parameter Markov Chain is a using methods in estimates of earthquake epicenter as well as geological approaches. Discrete parameter Markov Chain is successful to examine as a probabilistic transition of earthquake epicenter. This method only accept as a sequence this transitions and it gives a probability distribution depending on process. Hidden Markov model makes more significant this model. Because Hidden Markov Model consider times of seismic inactivity as well as earthquake epicenter. Times of inactivity depend on earthquake epicenter.

The new state sequence was obtained by forward algorithm and this sequence was converted to transition probabilities matrix by discrete parameter Markov Chains.

### 5. References


