Building an HMM for Earthquake Prediction

Understanding HMM States

What Are the States?

In HMMs, “states” are hidden conditions that affect what you observe. In earthquake prediction, states might represent different seismic modes or regimes. Most studies use only a few states (usually between 2 and 5):

- **Two-State Model:**
  One state represents a quiet period with low activity, and the other represents an active period where the chance of a big quake increases.
- **Multiple States:**
  We could use three to five states to capture gradual changes. For instance, we could have states for low, medium, and high activity. Additional states help capture the gradual escalation before a large event.
- **Incorporating Precursors:**
  Some researchers combine earthquake counts with extra signals such as GPS measurements, radon levels, or electromagnetic readings. In these models, the states are defined not only by past earthquake activity but also by the presence of unusual precursor signals.
- **Regional Differences:**
  In some cases, the states reflect differences across geographic regions. For example, one state might describe activity in one part of Taiwan while another state describes a different fault zone.

Researchers typically keep the number of states small and choose the best model using criteria such as AIC or BIC. (Rewards accuracy, but penalizes complexity)

Transition Probabilities

How Do States Change?

Transition probabilities indicate how likely it is for the system to move from one state to another. In a basic HMM, these probabilities are fixed and estimated using algorithms like Baum–Welch.

Here are some ways to enhance this:

- **Poisson HMMs:**
  When modeling earthquake counts, We can assume that each state follows a Poisson process with its own rate and then estimate transitions between these rates.
- **Hidden Semi-Markov Models:**
  In a regular HMM, the time spent in each state follows a memoryless (geometric/exponential) distribution. A hidden semi-Markov model lets us specify a more realistic duration for each state.
- **Time-Varying Transitions:**
  Instead of fixed probabilities, allow the probabilities to change over time or depend on external factors. For instance, the chance of moving to the active state may increase as more time passes since the last big quake.
- **Bayesian Estimation:**
  Using Bayesian methods, treat the transition probabilities as uncertain parameters with a prior distribution. After observing data, we obtain a full distribution (with credible intervals) for each transition probability. This approach reveals the uncertainty in the estimates.

Adding More Data: Geophysical Indicators

Beyond Earthquake Counts

To improve predictions we should also bring in additional data sources. The following were used in other studies:

- **GPS/Strain Data:**
  Continuous GPS measurements show ground deformation. A buildup in strain may signal an increased chance of a large quake.
- **Radon and EM Signals:**
  Monitoring radon levels in the soil or electromagnetic signals in the atmosphere can provide clues. Sudden changes in these readings may indicate that a significant event is coming.
- **Hybrid Models:**
  We can use machine learning to detect anomalies in these precursor signals. A neural network could spot unusual patterns, which then inform the HMM about a potential state change.

Using Bayesian Methods for Predictions

Getting a Forecast with Confidence

Rather than just output a single probability, a Bayesian approach produces a forecast with an uncertainty range:

- **Bayesian HMM:**
  In a Bayesian setup, we assign prior distributions to the model parameters (such as transition probabilities and emission rates) and update them with data. The result is a probability distribution that lets us say, for example, "There is a 10% chance of a big quake in the next month, with a 90% credible interval ranging from 5% to 18%."
- **Real-Time Updating:**
  As new data comes in, whether more earthquakes or fresh precursor information we can update the probabilities. This keeps the forecast current and shows how much confidence we have in the prediction at any moment.

Putting It All Together

Here's a blueprint for an improved earthquake forecasting model using an HMM:

1. **Design the States:**
   Start with two or three states, such as “quiet,” “moderate activity,” and “active” (where the active state is tied to precursor anomalies).
2. **Include Extra Data:**
   In addition to earthquake counts, incorporate GPS data, radon levels, and electromagnetic signals. A multivariate approach can capture the complex interactions between these data streams.
3. **Flexible Transitions:**
   Use models that allow transition probabilities to change over time or depend on the current duration in a state. Bayesian methods can help quantify the uncertainty in these probabilities.
4. **Bayesian Inference:**
   Run a Bayesian inference process (using MCMC or variational methods) to obtain posterior distributions over the model parameters. This approach produces not only probability forecasts but also credible intervals, making the predictions more informative.
5. **Validation:**
   Test the model using both global earthquake catalogs and Taiwan-specific data. Verify if the forecast probabilities align with observed events, and adjust if necessary.
   PGA

Preliminary HMM Model for Earthquake Prediction

Model Summary: Assumptions, Limitations, and Future Improvements

---

Assumptions

- **Two-State Framework (We can easily make it 2-5):**
  The model assumes that seismic activity can be characterized by just two hidden states:
  - **State 0 ("Quiet")** representing low seismic activity.
  - **State 1 ("Active")** representing high seismic activity (with precursor anomalies).
- **Emission Distribution:**
  Observations (earthquake counts, precursor signals) are assumed to follow a multivariate Gaussian distribution conditioned on the state:

$$\mathbf{x}_t \mid S_t = k \sim \mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k), \quad k \in \{0, 1\}.$$

- **Markov Property:**
  The state sequence follows a first-order Markov chain. This means that the probability of transitioning to a new state depends only on the current state:

$$P(S_t = j \mid S_{t-1} = i).$$

- **Bayesian Priors:**
  In the Bayesian extension, the model assumes reasonable prior distributions. (With large data the choice of prior has minimal effect):
  - Dirichlet priors for the transition probabilities.
  - Normal-Inverse-Wishart priors for the emission parameters (mean and covariance).
- **Data Representation:**
  It is assumed that the selected data (seismic records and precursor indicators) sufficiently capture the underlying processes that govern large earthquakes.

Potential Limitations

- **Gaussian Emission Assumption:**
  The assumption that the observation data follow a Gaussian distribution may not hold for all types of seismic or precursor data, which could be skewed or have heavy tails.
- **Data Quality and Availability:**
  The model's performance depends heavily on the quality, resolution, and completeness of the seismic and precursor datasets. Limited or noisy data can reduce predictive accuracy.
- **Static vs. Dynamic Transitions:**
  Even with planned improvements like time-varying or semi-Markov transitions, the current model might not fully capture the complex, dynamic nature of earthquake triggering.
- **Sensitivity to Priors:**
  In the Bayesian framework, the choice of priors can significantly influence the results,

particularly when data is sparse or the signal-to-noise ratio is low.

Areas for Potential Improvement

- **Increase State Complexity:**
  Consider exploring models with more than two states or using continuous latent variables to better capture the gradation in seismic activity.
- **Alternative Emission Models:**
  Investigate non-Gaussian or mixture models for emissions to more accurately reflect the distribution of earthquake and precursor data.
- **Incorporate Additional Data Sources:**
  Enhance the model by integrating more types of data (e.g., satellite imagery, higher resolution GPS data, ground deformation measurements) to improve the detection of precursor signals.
- **Dynamic Transition Modeling:**
  Develop methods to allow transition probabilities to vary over time or be influenced by external covariates, thus better reflecting the underlying physics of earthquake processes.
- **Hybrid Modeling Approaches:**
  Combine the HMM with machine learning techniques, such as deep learning for feature extraction or ensemble methods, to boost overall forecasting performance.
- **Real-Time Updating:**
  Implement an online Bayesian updating mechanism to continuously refine forecasts as new data becomes available, enhancing the model's responsiveness and reliability.
- **Robust Validation:**
  Test and calibrate the model across diverse datasets (both global and region-specific, like Taiwan) to ensure the model generalizes well and truly improves upon previous studies.

---

1. Hidden States and Their Dynamics

Let $S_t \in \{0, 1\}$ denote the hidden state at time t, where:

- **State 0 ("Quiet")**: Low seismic activity.
- **State 1 ("Active")**: High seismic activity (with precursor anomalies).

The states evolve according to a Markov chain with transition probabilities given by a 2×2 matrix:

$$A_{ij} = P(S_t = j \mid S_{t-1} = i), \quad i, j \in \{0, 1\}.$$

The initial state distribution is:

$$\pi_i = P(S_1 = i), \quad i \in \{0, 1\}.$$

---

2. Emission (Observation) Model

At each time t, we observe a vector $\mathbf{x}_t \in \mathbb{R}^d$ (a combination of earthquake counts and precursor signals). Assume that, given the state $S_t = k$, the observation is drawn from a multivariate Gaussian distribution:

$$\mathbf{x}_t \mid S_t = k \sim \mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k), \quad k \in \{0, 1\}.$$

In other words,

$$p(\mathbf{x}_t \mid S_t = k) = \frac{1}{(2\pi)^{d/2} |\Sigma_k|^{1/2}} \exp\left\{ -\frac{1}{2} (\mathbf{x}_t - \boldsymbol{\mu}_k)^\top \Sigma_k^{-1} (\mathbf{x}_t - \boldsymbol{\mu}_k) \right\}.$$

---

3. Joint Likelihood of the Observations and States

The joint probability of a sequence of observations $\mathbf{x}_{1:T}$ and hidden states $S_{1:T}$ is:

$$p(\mathbf{x}_{1:T}, S_{1:T}) = \pi_{S_1} \, p(\mathbf{x}_1 \mid S_1) \prod_{t=2}^{T} A_{S_{t-1} S_t} \, p(\mathbf{x}_t \mid S_t).$$

---

4. Forecasting and State Inference

Using the Forward-Backward algorithm, you can compute the filtered state probabilities:

$$P(S_t = k \mid \mathbf{x}_{1:t}) \quad \text{for } k = 0, 1.$$

For example, the forward recursion is:

$$\alpha_t(k) = p(\mathbf{x}_t \mid S_t = k) \sum_{i=0}^{1} \alpha_{t-1}(i) A_{ik},$$

where $\alpha_t(k)$ is the forward probability at time t.
To forecast, if state 1 represents high risk for a large earthquake, then the forecast probability is approximated by:

$$P(\text{Large Quake at } t \mid \mathbf{x}_{1:t}) \approx P(S_t = 1 \mid \mathbf{x}_{1:t}).$$

---

5. Bayesian Extension for Uncertainty

To account for uncertainty in the parameter estimates we can incorporate a Bayesian approach:

- **Transition Probabilities:**
  Assign a Dirichlet prior to each row of the transition matrix:

$$A_{i,\cdot} \sim \text{Dirichlet}(\alpha_{i,0}, \alpha_{i,1}), \quad i \in \{0, 1\}.$$

- **Emission Parameters:**
  Use a Normal-Inverse-Wishart prior for the mean and covariance of each state:

$$\boldsymbol{\mu}_k \mid \Sigma_k \sim \mathcal{N}(m_0, \Sigma_k/\kappa_0), \quad \Sigma_k \sim \text{Inverse - Wishart}(\Psi_0, \nu_0), \quad k \in \{0, 1\}.$$

After observing data, these priors are updated via MCMC or variational inference to obtain a posterior distribution over the parameters. This gives a forecast probability and credible intervals:

$$P(S_t = 1 \mid \mathbf{x}_{1:t}, \text{data}) \quad \text{with credible intervals.}$$

---

6. How This Model Differs from Other Studies and Improvements Made

Several modifications have been made to improve the predictive capabilities of this model compared to earlier studies:

1. **Integration of Multivariate Emissions:**
   While many previous papers use simple univariate observations (e.g., just earthquake counts), this model incorporates multiple data sources (such as precursor signals from GPS, radon levels, or electromagnetic observations) into a multivariate Gaussian framework. This allows for a more comprehensive representation of the system.
2. **Bayesian Parameter Estimation:**
   Unlike traditional models that provide point estimates for the transition matrix and emission parameters, this approach uses Bayesian methods to quantify uncertainty. By placing Dirichlet and Normal-Inverse-Wishart priors on the parameters, we obtain posterior distributions and credible intervals, offering more informative forecasts.

3. **Flexible Transition Dynamics:**
   Traditional HMM approaches often assume fixed, memoryless transitions. This model suggests future improvements by integrating time-varying or semi-Markov transitions, where the probability of state changes can depend on the duration in the current state or external covariates, thereby better capturing the physics of earthquake preparation.
4. **Hybrid Data Sources:**
   This approach goes beyond using just seismic data by combining it with geophysical precursor data. Other papers may focus solely on seismic catalogs, but by including precursor signals, the model aims to detect early warning signs that a large earthquake may be imminent.

Summary

1. **Hidden States:**
   $S_t \in \{0,1\}$ following a Markov chain with transition matrix A and initial distribution π.
2. **Emissions:**

$$\mathbf{x}_t \mid S_t = k \sim \mathcal{N}(\boldsymbol{\mu}_k, \Sigma_k), \quad k = 0, 1.$$

3. **Joint Probability:**

$$p(\mathbf{x}_{1:T}, S_{1:T}) = \pi_{S_1}\, p(\mathbf{x}_1 \mid S_1) \prod_{t=2}^{T} A_{S_{t-1}S_t}\, p(\mathbf{x}_t \mid S_t).$$

4. **Forecasting:**
   Use the Forward-Backward algorithm to compute $P(S_t = 1 \mid \mathbf{x}_{1:t})$ as a proxy for the probability of a large earthquake.
5. **Bayesian Framework:**
   Place priors on A, $\boldsymbol{\mu}_k$, and Σ_k to obtain a posterior distribution and credible intervals for forecasts.