Robust Hidden Markov Models with Systematic Local Maxima Avoidance for Pattern Analysis in Seismicity Data

We employed robust hidden Markov models (HMMs) to perform statistical pattern analysis of seismic events and crustal deformation. These models allowed us to classify different kinds of events or modes of deformation, and furthermore gave us a statistical basis for understanding relationships between different classes, including a potential basis for prediction.

The hidden Markov model (HMM) approach assumes that the observed data has been generated by an unobservable dynamical statistical process. The process is of a particular form such that each observation is coincident with the system being in a particular discrete state. The dynamics are the model are constructed so that the next state is directly dependent only on the current state; it is a first order Markov process. The model is completely described by a set of model parameters: the initial state probabilities, the first order Markov chain state-to-state transition probabilities, and the probability distribution of observable outputs associated with each state.

Application of the model to data involves optimizing these model parameters with respect to some function of the observations, typically the likelihood of the observations given the model. Our work focused on the fact that this objective function typically has a number of local maxima that is exponential in the model size (the number of states). This means that not only is it very difficult to discover the global maximum, but also that results can vary widely between applications of the model. For some domains, such as speech processing, sufficient a priori information about the system is available such that this problem can be avoided. However, for the geophysical data sets under consideration, this a priori information is not available.

Our approach involves analytical location of sub-optimal local maxima; once the locations of these maxima have been found, then we can employ a modified optimization procedure based on the application of statistical priors that is designed to avoid these sub-optimal points in the parameter space. The end result is a robust technique for estimating the optimal parameters of an HMM and thereby the statistical properties of the data. We compare this method to the method of deterministic annealing as applied to hidden Markov models, and discuss a combined algorithm that employs both techniques simultaneously to the advantage of each.

We present results of this method as applied to two particular geophysical data sets of general interest. The first is a catalog of seismic events in the Southern California region. Each earthquake was treated as an observation in six dimensions: latitude, longitude, depth, magnitude, time since the previous event, and time until the next event. Our preliminary investigations using yielded results that included models identifying significant classes of earthquakes, such as aftershock sequences and earthquake swarms. The second data set, consisting of Global Positioning System (GPS) measurements collected by the Southern California Integrated GPS Network (SCIGN), yielded equally promising models that clearly identified certain modes of crustal deformation caused by various physical phenomena such as earthquakes and ground water pumping. In both cases, our robust technique allowed us to generate high quality models of the data that were consistent across many experiments.