HMM in crosses and small pedigrees

Lecture 8, Statistics 246,

February 17, 2004

Discrete-time Markov chains

Consider a sequence of random variables $X_1, X_2, X_3, \ldots$with common finite state space S. This sequence forms a Markov chain if for all t, $\{X_1, \ldots, X_{t-1}\}$ and $\{X_{t+1}, X_{t+2}, \ldots\}$ are conditionally independent given X_t, equivalently,

$$pr(X_t \,|\, X_{t-1}, X_{t-2}, \ldots) = pr(X_t \,|\, X_{t-1}).$$

The matrix $p(i,j;t) = pr(X_t = j \,|\, X_{t-1} = i)$ is the transition matrix at step t.

When $p(i,j;t) = p(i,j)$ for all i and j, independent of t, we say the Markov chain is time-homogeneous, or has stationary transition probabilities. Many of the chains we'll be meeting will be inhomogeneous, and t will be in space, not time.

There are plenty of good books on elementary Markov chain theory, Feller vol 1 being my favourite, but they mostly concentrate on asymptotic behaviour in the homogeneous case. For the time being we don't need this, or much else from the general theory, apart from the fact that multi-step transition matrices are products of 1-step transition matrices (**Exercise**).

Hidden Markov Models (HMM)

If (X_t) is a Markov chain, and f is an arbitrary function on the state space, then $(f(X_t))$ will not in general be a Markov chain.

Exercise: Construct an example to demonstrate the last assertion.

It is sometimes the case that associated with a Markov chain (X_t) is another process, (Y_t) say, whose terms are conditionally independent given the chain (X_t). This happens with so-called semi-Markov chains.

Both functions of Markov chains and this last situation are covered by the following useful definition, based on the work of L. E. Baum and colleagues around 1970. A bivariate Markov chain (X_t, Y_t) is called a Hidden Markov Model if a) (X_t) is a Markov chain, and b) the distribution of Y_t given X_t, X_{t-1}, X_{t-2} …. depends only on X_t and X_{t-1}. In many examples, this dependence is only on X_t, but in some, it can extend beyond X_{t-1}, and/or include Y_t. Once you see how the defining property is used in the calculations, you will get an idea of the possible extensions.

Exercise. Explain how functions of Markov chains are always HMM.