APPPHYS217 Thursday 22 April 2010

Continuous-time Markov Chains

[See for example R. Durrett, *Essentials of Stochastic Processes* (Springer, 2004)]
We consider a random process $\{x_t, t \geq 0\}$ taking values in a discrete set of states $\{a_1, a_2, \ldots, a_K\}$ with Markovian transition probabilities

$$p_{ij}(h) \equiv \Pr\{x_{t+h} = a_j | x_t = a_i\},$$

satisfying (Chapman-Kolmogorov equation)

$$\sum_k p_{ik}(s) p_{kj}(t) = p_{ij}(s+t),$$

and

$$\begin{aligned} \lim_{h\to 0} p_{ij}(h) &= 1 - \nu_i h + o(h), \quad j = i, \\ &= \nu_{ij} h + o(h), \quad j \neq i, \end{aligned}$$

where the ν_{ij} are non-negative rate coefficients with

$$\nu_i \equiv \sum_{j \neq i} \nu_{ij}.$$

Such a process can be associated with random flights on a graph whose nodes correspond to the states a_i and whose directed edges correspond to the non-zero transition rates $\nu_{ij} > 0$.

Example [Durrett Ch. 4, Ex. 1.1]: Let the a_i be non-negative integers correspond to the number of events accumulated by a Poisson process with rate λ. Then

$$\nu_{ij} = \lambda \delta(j, i+1),$$

and x_t is a process that starts at zero and jumps in unit increments. The corresponding graph is linear with unidirectional edges.

Example [Durrett Ch. 4, Ex. 1.3]: The 'M/M/s queue' with s tellers, customer arrival rate λ, and service rate μ : the a_i are non-negative integers corresponding to the number of customers in line, and

$$\begin{aligned} \nu_{i(i+1)} &= \lambda, \\ \nu_{i(i-1)} &= i\mu, \quad 0 \leq i \leq s, \\ &= s\mu, \quad i > s. \end{aligned}$$

The corresponding graph is again linear but now with edges in both directions.

Example: Michaelis-Menten-like enzyme kinetics [H. Qian, Biophys. J. **95**, 10 (2008)].

It is common practice to study the 'dynamics' of a continuous-time Markov Chain (CTMC) using a probability vector on the state space,

$$\vec{p}(t) \equiv \begin{pmatrix} \Pr(x_t = a_1) \\ \Pr(x_t = a_2) \\ \vdots \\ \Pr(x_t = a_K) \end{pmatrix},$$

which evolves according to

$$\frac{d}{dt} p_i(t) = \sum_{j \neq i} \nu_{ji} p_j(t) - \sum_{j \neq i} \nu_{ij} p_i(t),$$

which we can rewrite as a simple matrix multiplication

$$\frac{d}{dt} p_i(t) = \sum_{j} \nu_{ji} p_j(t),$$

if we define

$$\nu_{ii} \equiv -\sum_{j \neq i} \nu_{ij}.$$

Note that this is a kind of master equation that evolves our state of knowledge about x_t if we know the transition rates but can't observe the actual jumps.

Example: Evolution of $\vec{p}(t)$ for the Poisson example.

Hidden Markov Models

A Hidden Markov Model (HMM) is basically a Markov Chain observed in noise. Working for the moment in discrete time we can think of an HMM as a pair of random processes $\{X_t, Y_t\}$ where X_t is an unobserved Markov Chain while, at each value of t, Y_t is an observed random variable whose statistics are determined by X_t. The usual problem formulation is, given observations of the Y_t and knowledge of the transition probabilities for X_t, try to infer the evolution of X_t. Note that we can pose this problem either as a recursive ('real-time') estimation problem as would be relevant for feedback control, in which you are required to estimate X_t at time t on the basis of the Y_s with $s \leq t$ only, or as a 'smoothing'/'interpolation' problem in which you can use the Y_s at all times s in order to estimate X_t (more like an offline data processing scenario).

In what follows we will consider the specific case of an observed signal with gaussian noise, where the mean of y_t (returning to continuous time) is determined by x_t but the variance of y_t is constant:

$$dy_t = x_t\, dt + \beta\, dW_t.$$

In this case (and in some simple generalizations) we can solve the recursive estimation problem using the Wonham filter.

The Wonham filter

[W. M. Wonham, "Some applications of stochastic differential equations to optimal nonlinear filtering," J. SIAM Control Ser. A **2**, 347 (1965)]

Assume the setup described above: the underlying CTMC $\{x_t, t \geq 0\}$ takes values in a finite set of real numbers $\{a_1, \ldots, a_K\}$ and has transition probabilities given by

$$\begin{aligned} \lim_{h \to 0} p_{ij}(h) &= 1 - \nu_i h + o(h), \quad j = i, \\ &= \nu_{ij} h + o(h), \quad j \neq i. \end{aligned}$$

The observed process is dy_t with

$$dy_t = x_t\, dt + \beta dW_t.$$

Our task is to compute the posterior (conditional) probabilities

$$p_j(t) = \Pr\left\{x_t = a_j \,|\, dy_s,\ 0 \leq s \leq t\right\}, \quad j = 1, \ldots, K.$$

Wonham uses stochastic calculus methods to show that these evolve according to

$$dp_j(t) = \left\{ -\nu_j p_j(t) + \sum_{\substack{i=1 \\ i \neq j}}^{K} \nu_{ij} p_i(t) \right\} dt - \beta^{-2} \langle x \rangle_t [a_j - \langle x \rangle_t] p_j(t) dt + \beta^{-2} [a_j - \langle x \rangle_t] p_j(t)\, dy_t,$$

where $\langle x \rangle_t$ is the conditional mean defined by

$$\langle x \rangle_t \equiv \sum_{i=1}^{K} p_i(t) a_i.$$

Wonham notes that the terms not in the curly braces correspond to the filter one would use for a constant signal in gaussian noise ($\nu_{ij} = 0$), and we have seen above that the term in curly braces is the master equation for the evolution of $\vec{p}$ without observations. The observation term can be rewritten

$$-\beta^{-2} \langle x \rangle_t [a_j - \langle x \rangle_t] p_j(t) dt + \beta^{-2} [a_j - \langle x \rangle_t] p_j(t)\, dy_t = -\beta^{-2} [a_j - \langle x \rangle_t] \{\langle x \rangle_t dt - dy_t\} p_j(t),$$

which we see has an ‘innovations’ structure. The ‘gain’ of the corrector term scales with β^{-2}, which is perhaps not surprising. If the observed signal dy_t is below the expected signal $\langle x \rangle_t dt,$ the conditional probability of any state a_j whose value is above the conditional mean $\langle x \rangle_t$ is decreased while that of any state whose value is above the conditional mean is increased. Likewise, if the observed signal is above the expected signal, updates are applied in the opposite directions.

Note that the Wonham filter is a nonlinear SDE since we have terms like $\langle x \rangle_t p_j(t)$.

We finish with an example from [“Continuous quantum error correction as classical hybrid control,” New J. Phys. **11**, 105044 (2009)].