Grad Seminar

Liam Cregg

October 28, 2022

1 "Hidden" Markov Models

- Let (x_n) and (y_n) be discrete-time process taking values in Polish spaces $\mathbb{X}$ and $\mathbb{Y}$.
- (x_n) is called the state process and is Markov (i.e. characterized by its initial distribution μ and its transition kernel $T(x, dx')$).
- (y_n) is the observation process and is parameterized by the current value of (x_n) (i.e. $Q^{x'}(y, y')$), with $Q^{x'}(y, B) = \int_B r(x', y, y')\eta(dy')$. An example of this form is if (y_n) is updated as $y_{n+1} = h(y_n, x_{n+1}, w_{n+1})$ where w_{n+1} identically distributed and independent of X^{n+1} and Y^n.

Lemma 1.1. *(x_n, y_n) is a Markov process.*

- We denote our "belief" or the "filter" on the process (x_n) as $\pi_n^\rho(A) := P(x_n \in A|Y^n)$, where ρ is the initial measure on $\mathbb{X}, \mathbb{Y}$. Then we can write $\rho(dx, dy) = p_\rho(y, dx)\rho(\mathbb{X}, dy)$, and we can say that $\pi_0^\rho(A) = p_\rho(y_0, A)$.

  Note that we can describe the evolution of the filter in a Bayesian update fashion:

$$M(y, y', \pi)(A) = \frac{\int_A r(x', y, y') \int_{\mathbb{X}} T(x, dx')\pi(dx)}{\int_{\mathbb{X}} r(x', y, y') \int_{\mathbb{X}} T(x, dx')\pi(dx)}$$

  Then:

$$\pi_n(A) = M(y_{n-1}, y_n, \pi_{n-1}^\rho)(A)$$

Lemma 1.2. *(π_n^ρ, y_n) is a Markov process.*

- We are interested in the existence of unique invariant measures for this process. There was a paper by Kunita that "proved" that unique ergodicity of the state process was equivalent, but a counterexample was later presented:

Lemma 1.3. *Unique ergodicity for (x_n) is not sufficient for unique ergodicity of (π_n^ρ, y_n).*

Proof. Let $\mathbb{X} = \{1, 2, 3, 4\}$, $\mathbb{Y} = \{0, 1\}$, and let (x_n) have transition matrix:

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}$$

Note that (x_n) is an irreducible and aperiodic finite-state Markov chain, so it has a unique invariant measure, and that $T(x, \{1, 3\}) = \frac{1}{2}\ \forall\, x$. Now let $y_n = 1_{\{1,3\}}(x_n)$. Then (y_n) is an i.i.d. sequence with probability $\frac{1}{2}$.

Then consider the following distributions:

$$e_1 = \begin{pmatrix} \alpha \\ 0 \\ 1-\alpha \\ 0 \end{pmatrix} e_2 = \begin{pmatrix} 0 \\ \alpha \\ 0 \\ 1-\alpha \end{pmatrix} e_3 = \begin{pmatrix} 1-\alpha \\ 0 \\ \alpha \\ 0 \end{pmatrix} e_4 = \begin{pmatrix} 0 \\ 1-\alpha \\ 0 \\ \alpha \end{pmatrix}$$

Say (π_n^ρ) starts from e_1. Then it will cyclically move through the e's with probability $\frac{1}{2} \Rightarrow$ the uniform measure over $\{e_i\}_{i=1}^4$ is invariant. But by starting from a different distribution (by changing the value of alpha s.t. $\beta \neq \alpha$ and $\beta \neq 1-\alpha$), we obtain another invariant measure (still uniform, but with different support). $\square$

Comment on different reasons for this (not ergodic "enough", poor observation)

- We define an *approximate filter* $\pi_n^{\rho\rho'}$ as (the same as above but initial π_0 given by ρ'). The above properties still hold. We also note that $(x_n, y_n, \pi_n^{\rho\rho'})$ is Markov.
- We call the approximate filtering process *weakly stable in expectation* if $\forall \rho_1, \rho_2$ and $\forall f \in C_b(\mathbb{X})$, we have $\lim_{n\to\infty} E_{xy}[|\pi_n^{\rho\rho_1}(f) - \pi_n^{\rho\rho_2}(f)|] = 0$

Theorem 1.4. *If (x_n, y_n) has a unique invariant measure $\zeta(dx, dy)$ and the approximate filter process $(\pi_n^{\rho\rho'})$ is weakly stable in expectation, then there exists at most one unique invariant measure for $(x_n, y_n, \pi_n^{\rho\rho'})$.*

Proof. Assume that $m_1, m_2 \in \mathcal{P}(\mathbb{X} \times \mathbb{Y} \times \mathcal{P}(\mathbb{X}))$ are two invariant measures for the joint process $(x_n, y_n, \pi_n^{\rho\rho'})$. Then their projections on $\mathbb{X} \times \mathbb{Y}$ are invariant for (x_n, y_n). Then, by unique invariance of $\zeta(dx, dy)$ we have

$$m_i(dx, dy, d\nu) = P_{m_i}(d\nu|x, y)\zeta(dx, dy)$$

Then we show that $m_1(F) = m_2(F)$ for each F on a set of measure-determining functions [1], namely those s.t. $F(x, y, \nu) = \phi(x, y)H(\nu(\phi_1), \dots, \nu(\phi_l))$, where $\phi \in C(\mathbb{X} \times \mathbb{Y}), \phi_1, \dots, \phi_l \in C(\mathbb{X}), H$ is bounded and Lipschitz continuous with constant L_H, and $l \in \mathbb{N}$.

Let S be the transition operator associated with the process $\{x_n, y_n, \pi_n^{\rho\rho'}\}$. Then by invariance we have for $i = 1, 2$

$$m_i(F) = \int_{\mathbb{X}\times\mathbb{Y}\times\mathcal{P}(\mathbb{X})} \frac{1}{n} \sum_{j=0}^{n-1} S^j F(x, y, \nu) P_{m_i}(d\nu|x, y)\zeta(dx, dy)$$

And thus,

$$\begin{aligned}
&|m_1(F) - m_2(F)| \\
\leq &\int_{\mathbb{X}\times\mathbb{Y}\times\mathcal{P}(\mathbb{X})\times\mathcal{P}(\mathbb{X})} \frac{1}{n} \sum_{j=0}^{n-1} |S^j F(x, y, \nu_1) - S^j F(x, y, \nu_2)| P_{m_1}(x, y, \nu_1) P_{m_2}(x, y, \nu_2)\zeta(dx, dy) \\
\leq &L_H||\phi|| \int_{\mathbb{X}\times\mathbb{Y}\times\mathcal{P}(\mathbb{X})\times\mathcal{P}(\mathbb{X})} \frac{1}{n} \sum_{j=0}^{n-1} E_\rho[\sum_{i=1}^{l} |\pi_j^{\rho\rho_1}(\phi_i) - \pi_j^{\rho\rho_2}(\phi_i)|] P_{m_1}(x, y, \nu_1) P_{m_2}(x, y, \nu_2)\zeta(dx, dy)
\end{aligned}$$

Since the filter is weakly stable in expectation, and by the dominated convergence theorem, the last line converges to zero as $n \to \infty$. □