MATH 56A: STOCHASTIC PROCESSES
CHAPTER 7

7. Reversal

This chapter talks about time reversal. A Markov process is a state X_t which changes with time. If we run time backwards what does it look like?

7.1. **Basic equation.**

There is one point which is obvious. As time progresses, we know that a Markov process will converge to equilibrium. If we reverse time then it will tend to go away from the equilibrium (contrary to what we expect) unless we start in equilibrium. If a process is in equilibrium, it will stay in equilibrium (fluctuating between the various individual states which make up the equilibrium). When we run the film backwards, it will fluctuate between the same states. So, we get a theorem:

Theorem 7.1. *A Markov process with equilibrium distribution π remains a Markov process (with the same equilibrium) when time is reversed provided that*

1. *left limits are replaced by right limits,*
2. *the process is irreducible*
3. *and nonexplosive.*

The time revered process has a different transition matrix

$$\widehat{P} = \Pi^{-1} P^t \Pi$$

where $P = (p(x, y))$ is the transition matrix for the original process and

$$\Pi = \begin{pmatrix} \pi(1) & 0 & \cdots \\ 0 & \pi(2) & \cdots \\ 0 & 0 & \cdots \end{pmatrix}$$

is the diagonal matrix with diagonal entries $\pi(1), \pi(2), \cdots$ given by the equilibrium distribution. In other words,

$$\hat{p}(x, y) = \pi(x)^{-1} p(y, x) \pi(y)$$

or

$$\pi(x) \hat{p}(x, y) = \pi(y) p(y, x)$$

Date: December 14, 2006.

This makes sense because $\pi(y)p(y,x)$ is the equilibrium probability that a random particle will start at y and then go to x. When we run the film backwards we will see that particle starting at x and moving to y. So, the probability of that is $\pi(x)\hat{p}(x,y)$

$$x\bullet \xleftarrow{p(y,x)} \bullet_{\pi(y)} y$$

$$x\bullet_{\pi(x)} \xrightarrow{\hat{p}(x,y)} \bullet y$$

7.1.1. *Example 1.* Take the continuous Markov chain with infinitesmal generator

$$A = \begin{pmatrix} -2 & 2 & 0 \\ 0 & -4 & 4 \\ 1 & 1 & -2 \end{pmatrix}$$

The row are required to have sum zero and terms off the diagonal must be nonnegative. The equilibrium distribution (satisfying $\pi A = 0$) is

$$\pi = (1/4, 1/4, 1/2)$$

So, the time reversed process is

$$\widehat{A} = \Pi^{-1} A^t \Pi = \begin{pmatrix} 4 & & \\ & 4 & \\ & & 2 \end{pmatrix} \begin{pmatrix} -2 & 0 & 1 \\ 2 & -4 & 1 \\ 0 & 4 & -2 \end{pmatrix} \begin{pmatrix} 1/4 & & \\ & 1/4 & \\ & & 1/2 \end{pmatrix}$$

$$\widehat{A} = \begin{pmatrix} -2 & 0 & 2 \\ 2 & -4 & 2 \\ 0 & 2 & -2 \end{pmatrix}$$

7.2. **Reversible process.**

Definition 7.2. *A Markov process is called* reversible *if* $\widehat{P} = P$. *This is the same as* $\widehat{A} = A$. *We say it is* reversible with respect to a measure π *if*

$$\pi(x)p(x,y) = \pi(y)p(y,x)$$

Example 1 is not a reversible process because $\hat{A} \neq A$.

Theorem 7.3. *If a Markov chain is reversible wrt a measure* π *then*

(1) *If* $\sum \pi(k) < \infty$ *then*

$$\lambda(j) = \frac{\pi(j)}{\sum \pi(k)}$$

is the (unique) invariant probability distribution.

(2) *If* $\sum \pi(k) = \infty$ *then the process is not positive recurrent.*

7.2.1. *example 2.* Take the random walk on $S = \{0, 1, 2, \cdots\}$ where the probability of going right is p. I.e., $p(k, k+1) = p, p(k+1, k) = 1 - p$.

(1) Show that this is a reversible process.
(2) Find the measure π
(3) What is the invariant distribution λ?

To answer the first two questions we have to solve the equation:

$$\pi(k)p(k, k+1) = \pi(k+1)p(k+1, k)$$

or:

$$\pi(k+1) = \frac{p}{1-p}\pi(k)$$

This has an obvious solution:

$$\pi(k) = \left(\frac{p}{1-p}\right)^k$$

Therefore, the random walk is reversible.

Now we want to find the invariant distribution λ.

If $p < 1/2$ then

$$\sum_{k=0}^{\infty} \pi(k) = \frac{1-p}{1-2p}$$

So, the equilibrium distribution is

$$\lambda(k) = \frac{p^k(1-2p}{(1-p)^{k+1}}$$

If $p \geq 1/2$ then

$$\sum_{k=0}^{\infty} \pi(k) = \infty$$

since the terms don't go to zero. So the process is not positive recurrent and there is no equilibrium.

7.3. **Symmetric process.**

Definition 7.4. *A Markov chain is called* symmetric *if* $p(x, y) = p(y, x)$. *This implies reversible with respect to the uniform measure:* $\pi(x) = 1$ *for all* x *and the process is positive recurrent if and only if there are finitely many states.*

I talked about one example which is related to the final exam. It is example 3 on page 160: Here S is the set of all N-tuples $(a_1, a_2, \cdots, a_N)$ where $a_i = 0, 1$ and the infinitesimal generator is

$$\alpha(a, b) = \begin{cases} 1 & \text{if } a, b \text{ differ in exactly one coordinate} \\ 0 & \text{otherwise} \end{cases}$$

This is symmetric: $\alpha(a,b) = \alpha(b,a)$.

We want to find the second largest eigenvalue
$_2$ of A. (The largest eigenvalue is $\lambda_1 = 0$. The second largest is negative with minimal absolute value.) The eigenvectors of A are also eigenvectors of $P = e^A$ with eigenvalue e^λ, the largest being $e^0 = 1$ and the second largest being $e^{\lambda_2} < 1$.

The first thing I said was that these eigenvectors are π-orthogonal.

Definition 7.5.

$$\langle v, w \rangle_\pi := \sum_{x \in S} v(x)w(x)\pi(x)$$

When $\pi(x) = 1$ (as is the case in this example) this is just the dot product. v, w are called *π-orthogonal* if

$$\langle v, w \rangle_\pi = 0$$

According to the book the eigenvalues of A are

$$\lambda_j = -2j/N$$

for $j = 0, 1, 2, \cdots, N$. This implies that the distance from X_t to the equilibrium distribution decreases at the rate of $-2/N$ on the average:

$$\mathbb{E}(||X_t - \pi||) \leq e^{-2t/N} ||X_0 - \pi||$$

7.4. **Statistical mechanics.** I was trying to explain the Gibbs potential in class and I gave you a crash course in statistical mechanics.

The fundamental assumption is that *All states are equally likely.* Suppose that we have two systems A, B with energy E_1, E_2. Suppose that

$$\Omega_A(E_1) = \#\text{states of } A \text{ with energy } E_1$$

$$\Omega_B(E_2) = \#\text{states of } B \text{ with energy } E_2$$

Then

$$\Omega_A(E_1)\Omega_B(E_2) = \#\text{states of } (A, B) \text{ with energy } E_1 \text{ for } A,\ E_2 \text{ for } B$$

Suppose that the two systems can exchange energy. Then they will exchange energy until the number of states is maximal. This is the same as when the log of the number of states is maximal:

$$\ln \Omega_A(E_1 + \Delta E) + \ln \Omega_B(E_2 - \Delta E) = 0$$

or:

$$\frac{\partial}{\partial E_1} \ln \Omega_A(E_1) = \frac{\partial}{\partial E_2} \ln \Omega_B(E_2) = \beta \quad \text{(constant)}$$

Define the *entropy* of the system A to be $S(E) = \ln \Omega_A(E)$. In equilibrium we have to have

$$\frac{\partial}{\partial E} S(E) = \beta$$

We think of B as an infinite reservoir whose temperature will not change if we take energy out.

Every state has equal probability. But, a state x of A with energy $E(x)$ cannot exist without taking $E(x)$ out of the environment B. Then the number of states of the environment decreases by a factor of $e^{-\beta E(x)}$. Therefore, the probability of the state is proportional to $e^{-\beta E(x)}$. So, the probability of state x is

$$\mathbb{P}(x) = \frac{e^{-\beta E(x)}}{\sum_{y \in S} e^{-\beta E(y)}}$$

The denominator is the *partition function*

$$Z(\beta) = \sum_{y \in S} e^{-\beta E(y)}$$

We looked at the *Ising model* in which there are points in a lattice and a state x is given by putting a sign $\epsilon_i(x) = \pm 1$ at each lattice point i and the energy of the state x is given by

$$E(x) = \sum_{i-j} |\epsilon_i(x) - \epsilon_j(x)| \cdot H$$

(This is $2H$ times the number of adjacent lattice point i, j so that the signs $\epsilon_i(x), \epsilon_j(x)$ are different.) Then I tried to explain the *Gibbs sampler* which is the Markov process which selects a lattice site i at random (with probability 1/#lattice points) and then changes $\epsilon_i(x)$ according to the probability of the new state y. So,

$$p(x, y) = \frac{1}{\#\text{lattice points}} \frac{\mathbb{P}(y)}{\mathbb{P}(y) + \mathbb{P}(y')}$$

if x, y differ at only one possible location i and y' is the other possible state which might differ from x at location i. (So, $x = y$ or $x = y'$.)

The Gibbs sampler has the effect of slowly pushing every state towards equilibrium.