Lecture 8

Coupling From The Past (CFTP)

goal: sample from the stationary distn of a MC exactly!

how: assume MC started at $t = -\infty$ and try to understand its value at present time by looking at "properties of recent steps of the chain"

↳ we'll try to identify such properties.

Preliminaries: Random function representation of MC

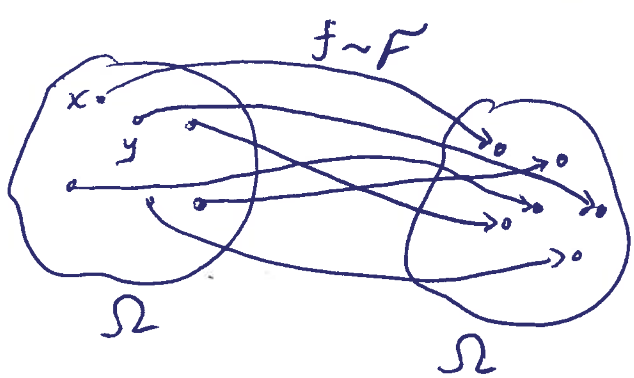
Def: Let P be an ergodic MC on Ω. A probability distribution $\mathcal{F}$ over functions from Ω to Ω is a random function representation of P if:

$$\forall x, y \in \Omega : \Pr_{f \sim \mathcal{F}}[f(x) = y] = P(x, y).$$

Remark: • Observe that $\mathcal{F}$ defines a coupling as follows

$$(X, Y) \to (f(X), f(Y))$$

• In fact, it is a complete coupling (or grand coupling) since it defines a step for all states in Ω simultaneously



- e.g. random walk on n-cycle: $\Omega = \{0, 1, \dots, n-1\}$

$$P(i,j) = \begin{cases} 1/2 & \text{, if } k = i+1 \pmod n \\ 1/2 & \text{, if } k = i-1 \pmod n \\ 0 & \text{, ow.} \end{cases}$$

Random mapping representation:

consider functions: $f(i) = i+1 \mod n$

$f'(i) = i-1 \mod n$

F choses f w.pr. $1/2$

f' w.pr. $1/2$

proposition: Every transition matrix P has a random mapping representation

Proof: ∘ Let $\Omega = \{x_1, x_2, \dots, x_n\}$ and for all j, k define $F_{j,k} = \sum_{i=1}^{k} P(x_j, x_i)$

∘ Random mapping representation:

- pick $r \in [0,1]$ uniformly at random
- define f as follows:

$$f(x_j) = x_k \quad \text{if} \quad F_{j,k-1} < r \le F_{j,k}.$$

• Clearly:

$$\Pr[f(x_j) = x_k] = \frac{F_{j,k} - F_{j,k-1}}{1} = P(x_j, x_k) \quad \boxtimes$$

Main: Coupling from the past

- Observe that an equivalent way to describe the evolution of the MC is to choose random functions $f_t \sim F$ i.i.d. for all $t = 0, 1, \dots$
- then:

$$X_t = f_{t-1} \circ f_{t-2} \circ \dots \circ f_0(X_0)$$

Suppose $(f_t)_{t=-\infty}^{+\infty}$ are i.i.d. samples from $\mathcal{F}$.

Define: $F_i^j = f_{j-1} \circ f_{j-2} \circ \dots \circ f_{i+1} \circ f_i$, for $j \geq i$,

F_0^t : "forwards" simulation of MC for t-steps

F_{-t}^0 : evolution of MC from time $-t$ to time 0
(i.e. from the past simulation)

Def: The coalescence time T_c is

$$T_c = \min \{ t : F_0^t \text{ is a constant function} \}$$

- Since $F_0^{T_c}$ is a constant function, in the grand coupling defined by $\mathcal{F}$, all paths of the MC starting at all points of Ω have met at time T_c.

- Reasonable, albeit false, intuition: $F_0^{T_c}(x)$ is distributed according to π, for all $x \in$.

- Remarkably, intuition is correct for the "from the past" simulation!!

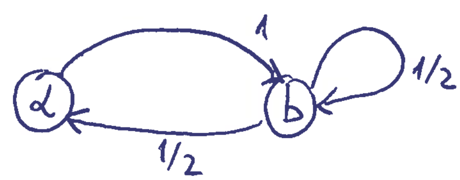
Formally, define the stopping time:

$$T = \min\{t : F_{-t}^{0} \text{ is a constant function}\}.$$

Theorem [Propp & Wilson '96]: Suppose T is finite with probability 1. Then $Z_{-\infty}^{0} := F_{-T}^{0}(x)$ is distributed according to π.

Gaining Intuition: Forward vs from the past simulation.

consider the following MC:

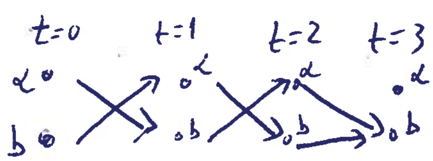


stationary distn:
$\pi(\alpha) = 1/3$
$\pi(b) = 2/3$

random function representation:

 w pr. $1/2$  w pr. $1/2$

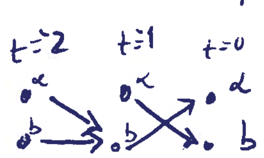
→ example forward simulation



$T_c = 3$

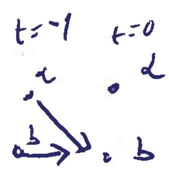
claim: $F_0^{T_c}(x) = b$

→ examples from the past simulations:



$T = 2$

$F_{-T}^{0}(x) = \alpha$



$T = 1$

$F_{-T}^{0}(x) = b$

- Crucial difference between Forward & From the past simulation.

$$F^0_{-t}(x) = F^0_{-T}(x), \ \forall \ t > T$$

while $F^t_0(x)$ & $F^{T_c}_0(x)$ are not necessarily the same constant functions.

- claim: Show that T & T_c have the same distribution.
Ex 1pt

Proof of theorem: - Since T is finite w/ prob 1, $Z^0_{-\infty}$ is well-defined w/ probability 1. Similarly, define the stopping time:

$$T' = \min\{t : F^1_{-t} \text{ is a constant fn}\}$$

and $Z^1_{-\infty} := F^1_{-T'}(x)$.

- Claim 1: If π_0 is the distribution of $Z^0_{-\infty}$ and π_1 is —//— $Z^1_{-\infty}$, then

$$\pi_0 \equiv \pi_1.$$

Proof: (intuition: Shift time to the right by 1)

Both $Z^0_{-\infty}$ and $Z^1_{-\infty}$ is the constant value of the function obtained via ~~from~~ the past simulation from some past time $-t$ to some fixed time 0 or 1 respectively. □

- Can couple the sampling of $Z^0_{-\infty}$ and $Z^1_{-\infty}$ by coupling F^0_{-t} & F^1_{-t} to choose the same $f_{t'}$ at all $t' < 0$.

- It follows that i. $T' \leq T$ (since if F^0_{-T} is constant, then F^1_{-T} is constan

ii. $Z^1_{-\infty} = f(Z^0_{-\infty})^{(*)}$, where $f \sim F$ is the random function chosen at time 0 from F^1

(*) and Claim 1 imply that $\pi_0 = \pi_1 = \pi$ (the stationary distribution of the chain)

since this is the unique distribution that is fixed under the application of a random $f \sim F$ ⊠

An exact Sampling Algorithm

$t = 0$; $F_t^0 \leftarrow$ identity function;
repeat
  sample random $f_t \sim F$
  $F_{t-1}^0 \leftarrow F_t^0 \circ f_t$
  $t \leftarrow t-1$
until F_t^0 is a constant function.
return the unique state in the range of $F_t^0(\cdot)$.

Claim: Suppose F guarantees coalescence time is finite w/prob. 1. Then the above procedure terminates w/ prob. 1 & the output is distributed according to π.

Proof: follows immediately from our theorem

Implementation Issues: to check if F_t^0 is constant in the end of each iteration need to compute the value of the function for all $x \in \Omega$, but Ω could be huge.

However, it can be implemented efficiently in many settings where ~~this~~ can be done implicitly.

Next time: apply coupling from the past to monotone settings.