Day 12 in 2021s

SPR Day 21

Markov Chain Monte Carlo & Metropolis Hastings Algorithm

Reading: Bishop PRML Sec 11.2.1 Markov chains

Sec 11.2.2 Metropolis Hastings

Topics:

1) MCMC with Propose-Accept transitions
2) Random Walk proposals: intuition
3) Detailed Balance, proof that RW works!
4) Metropolis + Metropolis-Hastings Algorithms

Goal: Sample from target distribution with PDF $p^*(z)$ over real $z \in \Omega \subseteq \mathbb{R}$

Challenge: Do NOT know how to evaluate $p^*(z)$, only known up to multiplicative constant

$\tilde{p}(z)$ is evaluate-able

$p^*(z) = c\,\tilde{p}(z)$ is NOT

because $c = \frac{1}{\int \tilde{p}(z)\,dz}$ is intractable (cannot do the integral)

Key idea: ratio of PDFs of two possible z values can be evaluated exactly

$$\frac{p^*(z^1)}{p^*(z^2)} = \frac{\cancel{c}\,\tilde{p}(z^1)}{\cancel{c}\,\tilde{p}(z^2)} = \frac{\tilde{p}(z^1)}{\tilde{p}(z^2)}$$ easy to compute!

Want to design Markov proposal that somehow uses ratios

Markov Transition Design: Propose then accept/reject

Transition distribution $\mathcal{T}$ takes current "state" value z_t, and produces new "state" value z_{t+1}

Idea: Propose new value $z' \sim Q(z' | z_t)$ then decide to accept w. prob $A(z', z_t)$

$$z_{t+1} = \begin{cases} z' & \text{if } u < A(z', z_t) \\ z_t & \text{otherwise} \end{cases}$$

where $z' \sim Q(z' | z_t)$
$u \sim \text{Unif}([0,1])$

$\boxed{\text{Example}}$

Start: $z_1 = 0.3$

$\mathcal{T}$

1) Sample $z' \sim Q(\cdot | z_1)$. $z' = -1.8$
2) Sample $u \sim \text{Unif}([0,1])$. $u = 0.3$
3) Eval $A(z', z_1) = 0.8$. $u < A$ so ACCEPT

$z_2 = -1.8$

$\mathcal{T}$

1) Sample $z' \sim Q(\cdot | z_2)$. $z' = 2.7$
2) Sample $u \sim \text{Unif}([0,1])$. $u = 0.91$
3) Eval $A(z', z_2) = 0.01$. $u > A$ so REJECT

$z_3 = -1.8$

(and keep going to get $z_1, z_2 \cdots z_S$

Remember, A is an accept probability <u>threshold</u>

We know $0 \le A \le +\infty$

$A = 0.0$ means no chance of accept
ALWAYS REJECT

$A \in (0, 0.5)$ means can accept, but likely reject

$A \in (0.5, 1]$ means can reject but likely accept

$A \ge 1$ means ALWAYS ACCEPT

So sometimes we write

$$p(\text{accept} \mid z', z_t) = \min\left(A(z', z_t), 1\right)$$

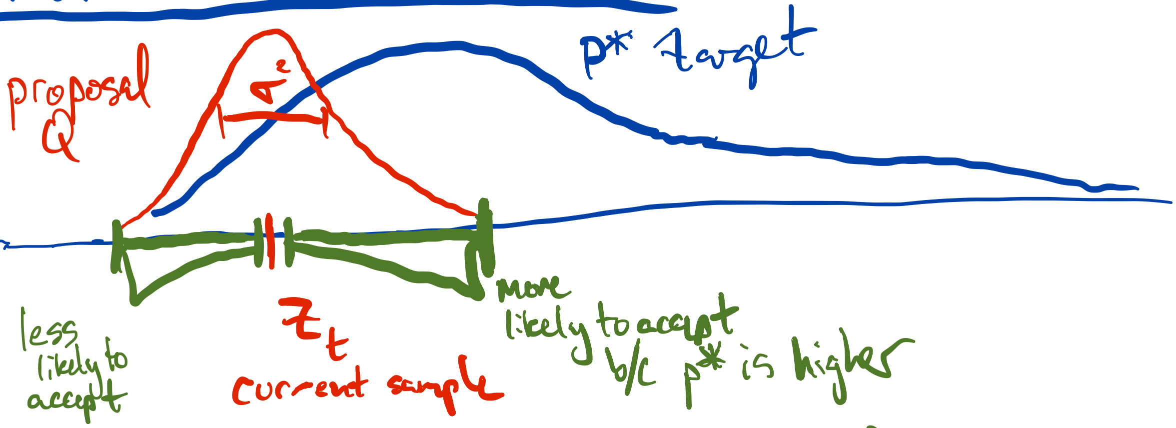
any value ≥ 1 is just set to 1

Proposal Idea: Random Walk

Random Walk $Q(z' \mid z_t) = \mathrm{Normal}(z' \mid \overset{\text{mean}}{z_t}, \overset{\text{var}}{\sigma^2})$ you pick $\sigma^2 > 0$ hyperparam.

why: easy to sample
likely to propose "nearby"

Intuition for Random Walk



let's try

$$A(z', z_t) = \frac{\tilde{p}(z')}{\tilde{p}(z_t)} = \frac{p^*(z')}{p^*(z_t)}$$

remember ratios of $\tilde{p}^*$ can be computed exactl

Given current z_t, and proposed value z',

$p^*(z')$ ___ than $p^*(z_t)$	ratio $\frac{p^*(z')}{p^*(z_t)}$	proposal will be
much smaller	$\ll 0.5$	likely rejected
a bit smaller	in $(0.5, 1)$	likely accepted
same or larger	≥ 1	certainly accepted

Intuitively, random walk proposal Q plus p^* ratio A makes sense, but how to show this will create Markov chain w/ stationary distr. p^*?

Recall, p^* is <u>stationary</u> for transition $\mathcal{T}$ if:

for all $z_{t+1} \in \Omega$: $$p^*(z_{t+1}) = \int_{z_t} p^*(z_t)\,\mathcal{T}(z_{t+1} \mid z_t)\,dz_t$$

Can we show that when $\mathcal{T}$ uses rand walk Q and accept proba A, that this is satisfied?

First, define PDF: $$\mathcal{T}(z_{t+1} \mid z_t) = \min\left(1, \frac{p^*(z_{t+1})}{p^*(z_t)}\right) Q(z_{t+1} \mid z_t)$$ when $z_{t+1} \neq z_t$

Next, show for all $z_a \neq z_b$

$$p^*(z_a)\,\mathcal{T}(z_b \mid z_a) = p^*(z_b)\,\mathcal{T}(z_a \mid z_b)$$

$$\cancel{\min(p^*(z_a), p^*(z_b))}\, Q(z_b \mid z_a) = \cancel{\min(p^*(z_b), p^*(z_a))}\, Q(z_a \mid z_b)$$

and we know $Q(z_b \mid z_a) = Q(z_a \mid z_b)$ because $\underset{\text{Norm PDF}}{\mathcal{N}}(z_a \mid z_b, \sigma^2) = \frac{1}{\sigma}\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\boxed{(z_a - z_b)^2}} = \mathcal{N}(z_b \mid z_a, \sigma^2)$

Symmetric

Remember, for any $a>0$, $b>0$

$$a \cdot \min\left(1, \frac{b}{a}\right)$$

$$= \begin{cases} a \cdot 1 & \text{if } b > a \\ a \cdot \frac{b}{a} & \text{if } b \leq a \end{cases} = \min(a, b)$$

Now, for

$z_t = z_{t+1}$
$= \tilde{z}$

$$\cancel{p^*(z_t)}\,\tau(z_{t+1} \mid z_t) = \cancel{p^*(z_{t+1})}\,\tau(z_t \mid z_{t+1})$$

$z_t = z_{t+1}$ $z_t = z_{t+1}$

$$\tau(\tilde{z} \mid \tilde{z}) = \tau(\tilde{z} \mid \tilde{z})$$

true by symmetry

So we have shown for all $z_a \in \Omega$, $z_b \in \Omega$, that random walk transition τ satisfies:

$$p^*(z_a)\,\tau(z_b \mid z_a) = p^*(z_b)\,\tau(z_a \mid z_b)$$

We call this the DETAILED BALANCE condition

$z_a \sim p^*$ has same $z_b \sim p^*$
$z_b \sim \tau(\cdot \mid z_a)$ joint proba as $z_a \sim \tau(\cdot \mid z_b)$

but, how does this help show p^* is stationary of τ?

Proving stationary if τ satisfies DETAILED BALANCE

well,

$$p^*(z_{t+1}) = \int p^*(z_t)\,\tau(z_{t+1}|z_t)\,dz_t$$

$$= \int p^*(z_{t+1})\,\tau(z_t|z_{t+1})\,dz_t$$

by detailed balance

$$= p^*(z_{t+1}) \int \tau(z_t|z_{t+1})\,dz_t$$

bringing out term const wrt z_t

$$= p^*(z_{t+1})$$

because τ is a PDF and must integrate to 1

Thus, p^* is stationary distribution, because it meets required definition.

METROPOLIS MCMC Algorithm

Initialize $z_1 \in \Omega$

for t in $1, 2, \dots S-1$

1) $z' \sim Q(\cdot \mid z_t)$
2) $u \sim \mathrm{Unif}([0,1])$
3) $z_{t+1} = \begin{cases} z' & \text{if } u \leq \dfrac{\tilde{p}(z')}{\tilde{p}(z_t)} = A(z', z_t) \\ z_t & \text{otherwise} \end{cases}$

return $[z_1, z_2, \dots z_S]$

Assumes that Q valid PDF over Ω
easy to sample
easy to evaluate PDF
must be symmetric

$$Q(z_a \mid z_b) = Q(z_b \mid z_a)$$

for all $z_a \in \Omega$, $z_b \in \Omega$

if S large enough, can consider $z_{S-B}, \dots z_S$ as samples of p^*

METROPOLIS-HASTINGS MCMC Algorithm

Initialize $z_1 \in \Omega$

for t in $1, 2, \dots S-1$

1) $z' \sim Q(\cdot \mid z_t)$
2) $u \sim \mathrm{Unif}([0,1])$
3) $z_{t+1} = \begin{cases} z' & \text{if } u < \dfrac{\tilde{p}(z')}{\tilde{p}(z_t)} \dfrac{Q(z_t \mid z')}{Q(z' \mid z_t)} \\ z_t & \text{otherwise} \end{cases}$

return $[z_1, z_2, \dots z_S]$

Assumes only that
Q is valid PDF over Ω
easy to sample
& easy to evaluate PDF

can show this A also satisfies DETAILED BALANCE

Sanity check: What if we use $Q = p^*$ in Met-Hastings

Accept ratio $A = \frac{\tilde{p}(z')}{\tilde{p}(z_t)} \frac{p^*(z_t)}{p^*(z')} = \frac{\tilde{p}(z')\tilde{p}(z_t)c}{\tilde{p}(z_t)\tilde{p}(z')c} = 1$

So we'd always accept! Makes sense.