- 1st example sheet will be on Moodle tomorrow.
- Details of dates & times for examples classes will be announced this week.

Stochastic Sampling

model params

model space, e.g. $\mathbb{R}^d$

Let $x \in \mathcal{X}$

r.v.

sample space

Bayesian posterior

P be dist on $\mathcal{X}$ (target dist)

vector in $\mathbb{R}^d$

$x_i \overset{iid}{\sim} P$ for $i = 1, 2, \dots, n$ $\ast$.

Goal of stochastic sampling is to produce x_i.

Given that we can evaluate $P(x)$ .... or $f(x) \propto P(x)$

samples alone don't ~~he~~ find normalising evidence Z.

suppose we have $\phi(x)$

$$E[\phi] = \int dx\ P(x)\ \phi(x)$$

this solved with samples

$$S_n = \frac{1}{n} \sum_{i=1}^{n} \phi(x_i)$$

$S_n \to E[\phi]$ as $n \to \infty$. error scales as $\sim \frac{1}{\sqrt{n}}$. for any d.

most deterministic methods converge faster, but $n \sim \exp(d)$.

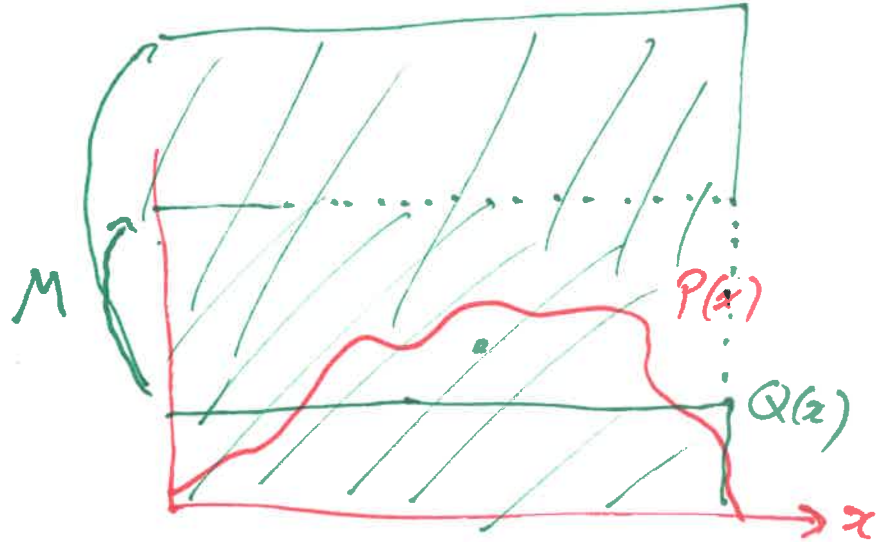
Why is this hard?

we can evaluate $f(x) \propto P(x)$.

Basic methods for stochastic sampling.

- rejection sampling. — can be v. hard to find a good proposal Q in high dim.
- transform (inverse CDF) sampling — CDF only in 1d; in general can't find ϕ.
- direct enumeration. — scale badly/exponentially with d.

Rejection sampling



1: $x \sim Q$
- $u \sim U(0,1)$
- (x, Mu)

if $Mu < P(x)/Q(x)$

then x is our sample

else go back to 1

Transform Sampling

$x \sim Q(x)$

$y = \phi(x)$ design ϕ such that $y \sim P(\cdot)$.

$u \sim U(0,1)$

CDF $F(x) = \int^x P(x)\,dx$

<u>exercise</u> show that $x = F^{-1}(u)$

$x \sim P(x)$

Markov Chains

Def **Markov Chain** a sequence of points $x_0, x_1, x_2, \ldots x_i \ldots$

$$P(x_i \mid \underbrace{x_0, x_1 \ldots,}\, x_{i-1}) = \underline{P(x_i \mid x_{i-1})}$$

Transition Probabilities

Time-homogeneous Markov chains $\quad P(x_{i+1} \mid x_i) = P(x_1 \mid x_0)$

transition prob doesn't depend on i.

$$P(x' \mid x) \equiv \underline{p(x', x)}$$

We care about the long-term behaviour.

start at x_0

after 1 iteration $\quad P(x_1 | x_0) = \rho(x_1, x_0)$

after 2 iterations $\quad P(x_2 | x_0) = \int dx_1 \, P(x_2 | x_1, x_0) \, P(x_1 | x_0)$

$$= \int dx_1 \, P(x_2 | x_1) \, P(x_1 | x_0)$$

$$= \int dx_1 \, \rho(x_2, x_1) \, \rho(x_1, x_0)$$

by induction

$$P(x_i | x_0) = \int dx_{i-1} \int dx_{i-2} \cdots \int dx_2 \int dx_1 \; \rho(x_i, x_{i-1}) \, \rho(x_{i-1}, x_{i-2}) \cdots \rho(x_2, x_1) \, \rho(x_1, x_0)$$

irreducibility

if for any x_0 for any $A \subset \mathcal{X}_0$ there exist ~~a~~ number of step n such that

$$\int_A dx_n \, P(x_n \mid x_0) > 0$$

"go any where" property.