Adaptive rejection Metropolis sampling

Dr. Jarad Niemi

STAT 615 - Iowa State University

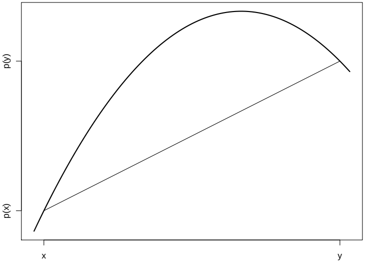
November 14, 2017

(Logarithmically) Concave Univariate Function

A function $p(\theta)$ is concave if

$$p((1-t)x + t\,y) \ge (1-t)p(x) + t\,p(y)$$

for any $0 \le t \le 1$.



If $p(x)$ is twice differentiabe, then $p(x)$ is concave if and only if $p''(x) \le 0$.
A function $p(x)$ is log-concave if $\log p(x)$ is concave.

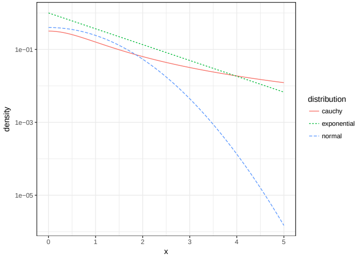
Examples

$X \sim N(0,1)$ has a log-concave density since

$$\frac{d^2}{dx^2}\log\, e^{-x^2/2} = \frac{d^2}{dx^2} - x^2/2 = \frac{d}{dx} - x = -1.$$

$X \sim Ca(0,1)$ has a non-log-concave density since

$$\frac{d^2}{dx^2}\log\frac{1}{1+x^2} = \frac{d}{dx}\frac{-2x}{1+x^2} = \frac{2(x^2-1)}{(1+x^2)^2}.$$



Log-concave distributions

- Log-concave distributions
  - normal
  - exponential
  - Uniform
  - Laplace
  - Gamma (shape parameter is ≥ 1)
  - Wishart ($n \geq p + 1$)
  - Dirichlet (all parameters ≥ 1)
- Non-log-concave distributions
  - Log-normal
  - Student t
  - F-distribution

Exponential distribution

An exponential distribution has pdf

$$p(\theta; b) = be^{-b\theta}$$

and thus has log-density

$$\log p(\theta; b) = \log(b) - b\theta$$

which is trivially log-concave since

$$\frac{d^2}{d\theta^2} \log(b) - b\theta = \frac{d}{d\theta} - b = 0 \leq 0.$$

The exponential distribution, or exponential function, is unique in that it matches the bound for the definition of log-concavity.

Prior-posterior example

The product of log-concave functions is also log-concave since

$$\log\left(\prod_{i=1}^n p_i(x)\right) = \sum_{i=1}^n \log p_i(x).$$

Assume

$$Y_i \stackrel{ind}{\sim} N(\theta,1) \quad \text{and} \quad \theta \sim La(0,1)$$

then the posterior

$$p(\theta|y) \propto \left[\prod_{i=1}^n N(y_i;\theta,1)\right] La(\theta;0,1)$$

is log-concave since - $N(y_i;\theta,1)$ is a log-concave function for θ for each y_i and - $La(\theta;0,1)$ is a log-concave distribution.

Rejection sampling

Suppose we are interested in sampling from a target distribution $p(\theta|y)$ using a proposal $q(\theta)$.
To use this algorithm, we must find

$$M \geq \frac{p(\theta|y)}{q(\theta)} \forall \theta$$

where the optimal M is $\sup_\theta p(\theta|y)/q(\theta)$.
Rejection sampling performs the following

1. Sample $\theta \sim q(\theta)$.
2. Accept θ as a draw from $p(\theta|y)$ with probability

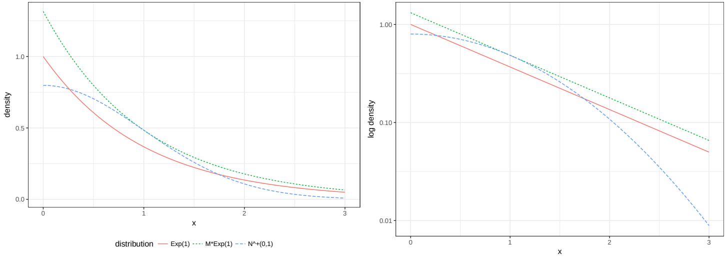
$$\frac{1}{M}\frac{p(\theta|y)}{q(\theta)}$$

   otherwise return to step 1.

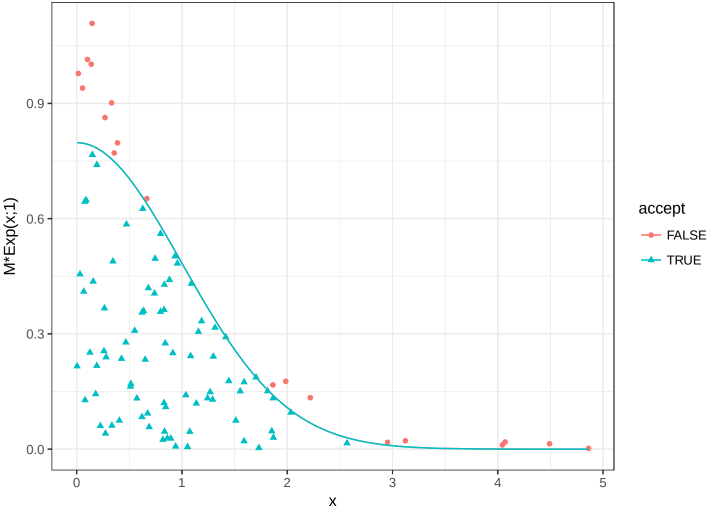
Rejection sampling envelope

Target $N^+(0,1)$ and proposal $Exp(1)$.
Then

$$\frac{\sqrt{2/\pi}e^{-\theta^2/2}}{e^{-\theta}} \le 1.315489 = M$$

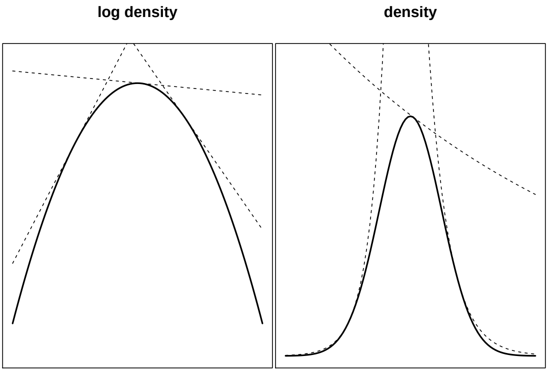


Rejection sampling example



Adaptive rejection sampling

Idea: build a piece-wise linear envelope to the log-density as a proposal distribution

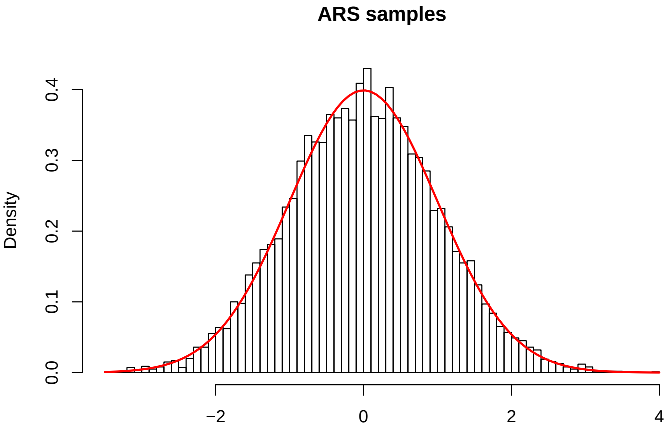


Pseudo-algorithm for adaptive rejection sampling

1. Choose starting locations θ, call the set Θ
2. Construct piece-wise linear envelope $\log q(\theta)$ to the log-density
    a. Calculate $\log f(\theta|y)$ and $(\log f(\theta|y))'$.
    b. Find line intersections
3. Sample a proposed value θ^* from the envelope $q(\theta)$
    a. Sample an interval
    b. Sample a truncated (and possibly negative of an) exponential r.v.
4. Perform rejection sampling
    a. Sample $u \sim Unif(0,1)$
    b. Accept θ^* if $u \leq f(\theta^*|y)/q(\theta^*)$.
5. If rejected, add θ^* to Θ and return to 2.

Adaptive rejection sampling (ARS) in R

```
library(ars)
f = function(x) -x^2/2 # log of standard normal density
fp = function(x) -x    # derivative of log of standard normal density
x = ars(1e4, f, fp)
```



ARS in R - non-log-concave density

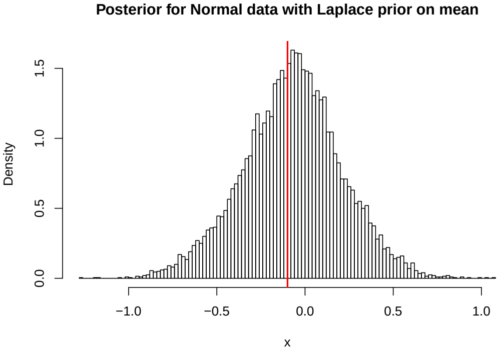
```
f = function(x) log(1/(1+x^2)) # log of standard cauchy density
fp = function(x) -2*x/(1+x^2)  # derivative of log of cauchy density
x = ars(1e4, f, fp)
```

```
##
## Error in sobroutine initial_...
## ifault= 5
```

ARS in R - prior-posterior example

$$Y_i \stackrel{ind}{\sim} N(\theta, 1) \quad \text{and} \quad \theta \sim La(0,1)$$

```
y = rnorm(10)
f = Vectorize(function(theta) sum(-(y-theta)^2/2) - abs(theta))
fp = Vectorize(function(theta) sum((y-theta)) - (theta>0) + (theta<0))
x = ars(1e4, f, fp)
```



Comments on ARS

- Derivative free ARS
- Checking for log-concavity
  - Decreasing derivatives
- Initial points for unbounded support:
  - initial derivative must be positive
  - final derivative must be negative
- Lower bound for multiple samples
  - Connect points
- Probability of acceptance increases at subsequent steps

Adaptive rejection Metropolis sampling (ARMS)

Adaptive rejection sampling is only suitable for log-concave densities. For non-log-concave densities adaptive rejection Metropolis sampling can be used

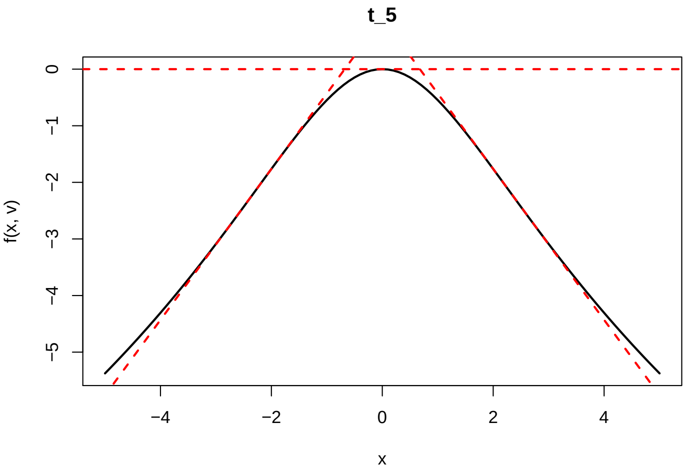
ARMS algorithm

1. Choose starting locations for θ, call the set Θ.
2. Construct piece-wise linear pseudo-envelope $\log q(\theta)$ to $\log p(\theta|y)$.
3. Sample $\theta^* \sim q(\theta)$ and $U \sim Unif(0,1)$.
   a. If $U \leq p(\theta^*|y)/q(\theta^*)$, proceed to Step 4.
   b. Otherwise, return to 3 add θ^* to Θ.
4. Perform Metropolis step: Set $\theta^{(i)} = \theta^*$ with probability

$$\min\left\{1, \frac{p(\theta^*|y)}{p(\theta^{(i)}|y)} \frac{\min\{p(\theta^{(i-1)}|y), q(\theta^{(i-1)})\}}{\min\{p(\theta^*|y), q(\theta^*)\}}\right\}$$

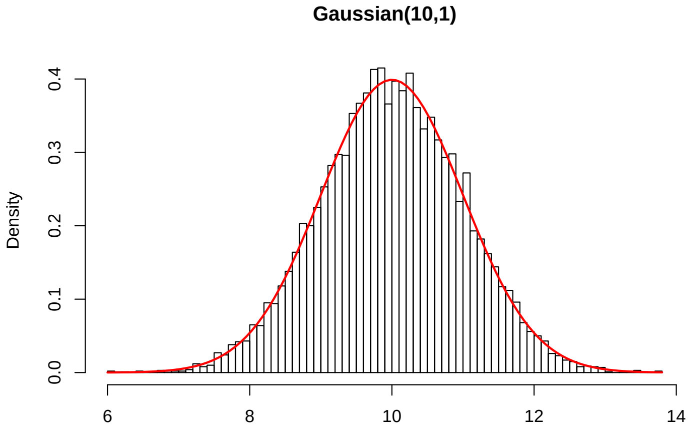
otherwise set $\theta^{(i)} = \theta^{(i-1)}$.

ARMS pseudo-envelope



ARMS in R

```
f = function(x,mean) -(x-mean)^2/2
x = dlm::arms(runif(1,3,17), f, function(x,mean) ((x-mean)>-7)*((x-mean)<7),
        1e4, mean=10)
hist(x,101,prob=TRUE,main="Gaussian(10,1)")
curve(dnorm(x,10), add=TRUE, lwd=2, col='red')
```



Theoretical consideration of ARMS

- ARMS is an independent Metropolis-Hastings algorithm
    - Proposal changes, due to updating q, i.e. adding more points in to Θ, thus inhomogenous.
    - We need to stop updating q at some point to enforce homogeneity.