

Derivative Based Adaptive Rejection Sampling (a review)

Erwann Rogard

2022-04-06

Abstract

This review was first published on 2007-11-19 as companion to a Mathematica package¹. The present version is just a revamp².

The purpose of the algorithm is to sample from a univariate log-concave (target) density, known up to a constant of proportionality, as commonly arising in Gibbs sampling applications[1][2]. The algorithm is derived from rejection sampling but instead of being pre-specified, the covering (proposal) density is constructed adaptively using only evaluations of the un-normalized density and its derivative. Geometrically, the log of the former is the upper envelope of the convex hull delimited by the tangents to the log un-normalized target at the evaluation points. This modification to the original algorithm requires a new rejection rule which depends on the distances between the true density and the upper hull, evaluated at the candidate draw, and that of the latter to a lower hull, formed by joining the evaluation points. Consider the definitions:

¹<https://library.wolfram.com/infocenter/ID/7071/>

²Revisions:https://github.com/erwannr/statistics/commits/main/deriv_ars

Definitions		
D		domain (connected)
D_{\pm}		lower/upper bound
$f(\cdot)$		log-concave density
$g(\cdot)$	$f(\cdot) \propto g(\cdot)$	un-normalized density
$h(\cdot)$	$\log g(\cdot)$	
$h'(\cdot)$	$h'(x) = \partial h(\cdot)/\partial x$	
x_k	$x_k \in D$	k th abscissae
$x_{(k)}$	$x_{(k)} \leq x_{(k+1)}$	k th ordered abscissae
$t_k^u(\cdot)$		tangent to $h(\cdot)$ at $x_{(k)}$, $k = 1, \dots, K$
$t_k^l(\cdot)$		line through $\{(x_{(j)}, h(x_{(j)})) : j = k, k+1\}$, $k = 1, \dots, K-1$
X	$\{x_k : k = 1, \dots, K\}$	set of $K \geq 2$ points such that $h'(x_{(1)}) > 0$ and $h'(x_{(K)}) < 0$, if $D_- = -\infty$ and $D_+ = +\infty$, respectively
$z_{(k)}$	$t_k^u(z_{(k)}) = t_{k+1}^u(z_{(k)})$	tangent intersection abscissae, $k = 1, \dots, K-1$
$z_{(0)}$	D_-	
$z_{(K)}$	D_+	
$u(\cdot)$		upper hull envelope of $h(\cdot)$
$s(\cdot)$	$s(x) \propto \exp(u(x))$	proposal density
c	$\int_D \exp(u(x)) dx$	normalizing constant
$S(\cdot)$	$\int_{z_{(0)}}^{(\cdot)} s(x) dx$	cumulative density
$l(\cdot)$		lower hull envelope

In particular,

$$t_k^u(x) = h(x_{(k)}) + (x - x_{(k)})h'(x_{(k)}) \quad (1)$$

$$t_k^l(x) = h(x_{(k)}) + (x - x_{(k)}) \frac{h(x_{(k+1)}) - h(x_{(k)})}{x_{(k+1)} - x_{(k)}} \quad (2)$$

$$z_{(k)} = -(t_{k+1}^u(0) - t_k^u(0)) / (h'(x_{(k+1)}) - h'(x_{(k)})) \quad (3)$$

$$u(x) = \sum_{k=1}^K 1\{z_{(k-1)} \leq x < z_{(k)}\} t_k^u(x) \quad (4)$$

$$S_{\text{un}}^u(x) = \sum_{k=1}^K \int_{x \wedge z_{(k-1)}}^{x \wedge z_{(k)}} \exp(t_k^u(y)) dy \quad (5)$$

$$= \sum_{k=1}^K \frac{1}{h'(x_{(k)})} (\exp(t_k^u(x \wedge z_{(k)})) - \exp(t_k^u(x \wedge z_{(k-1)}))) \quad (6)$$

$$S_u^{-1}(u) = \inf\{x : S_u^{\text{un}}(x) \geq u\} \quad (7)$$

$$= x_{(k_*)} + \frac{1}{h'(x_{(k_*)})} \left(\log \left(e^{t_{k_*}^u(z_{(k_*-1)})} + h'(x_{(k_*)})(cu - S_u^{\text{un}}(z_{(k_*-1)})) \right) - h(x_{(k_*)}) \right) \quad (8)$$

$$l(x) = \sum_{k=1}^{K-1} 1\{x_{(k)} \leq x < x_{(k+1)}\} t_k^l(x) - \infty(1 - 1\{x_{(1)} \leq x < x_{(K)}\}) \quad (9)$$

where $k_* = \min\{k : S^u(z_{(k)}) \geq u\}$.

The exponential terms in the un-normalized cumulative density are susceptible to overflows. To avoid this we do

$$h(\cdot) \leftarrow h(\cdot) - \max_k h(z_{(k)}) \quad (10)$$

which is permissible because $h(\cdot)$ is the log of an un-normalized version of $f(\cdot)$.

The algorithm below, initialized with some suitable X , generates one draw from f :

Algorithm 1 (ARS).

1. $\omega_l \stackrel{\text{iid}}{\sim} \text{Unif}(0, 1), l \in \{0, 1, 2\}; x^* \leftarrow S_u^{-1}(\omega_0)$
2. if($\omega_1 \leq \exp(l(x^*) - u(x^*))$) { return x^* } else {
if($\omega_2 \leq \exp(h(x^*) - u(x^*))$) {
(a) $X \leftarrow (X, x^*), K \leftarrow K + 1,$

```

    (b) update derived quantities;
    (c) return  $x^*$ 
} else{ goto 1.}
}

```

Bibliography

- [1] Walter R. Gilks and Pascal Wild. “Adaptive Rejection Sampling for Gibbs Sampling”. In: *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 41.2 (1992), pp. 337–348. URL: <http://www.jstor.org/stable/2347565>.
- [2] Wild. Pascal and Walter R. Gilks. “Algorithm AS 287: Adaptive Rejection Sampling from Log-Concave Density Functions”. In: *Applied Statistics* 42.4 (1993), pp. 701–709.