5.6 Hybridní algoritmy

IMPORTANCE SAMPLING

- to get normalizing constant
- $f(\theta) \propto g(\theta)$, i.e. $f(\theta) = c\, g(\theta)$, $c > 0$ not known

Observation:

$$1 = \int_{\Theta} f(\theta)\, d\lambda(\theta) = \int_{\Theta} c \cdot g(\theta)\, d\lambda(\theta)$$

Let $g_u(\theta)$ be density of any distribution with support Θ (or larger).

$$\Rightarrow 1 = \int_{\Theta} \frac{c \cdot g(\theta)}{c\, g_u(\theta)} \cdot c\, g_u(\theta)\, d\lambda(\theta) = c \int_{\Theta} \frac{g(\theta)}{g_u(\theta)} g_u(\theta)\, d\lambda(\theta)$$

$$= c \cdot E_{g_u}\left(\frac{g(\theta)}{g_u(\theta)} \right)$$

Hence

$$c = \frac{1}{\underbrace{E_{g_u}\left(\frac{g(\theta)}{g_u(\theta)} \right)}}$$

→ can be estimated by MC(MC) method by sampling from g_u

Note: small MC error in estimation of c if $\frac{g(\theta)}{g_u(\theta)} \sim 1 \quad \forall\, \theta \in \Theta$

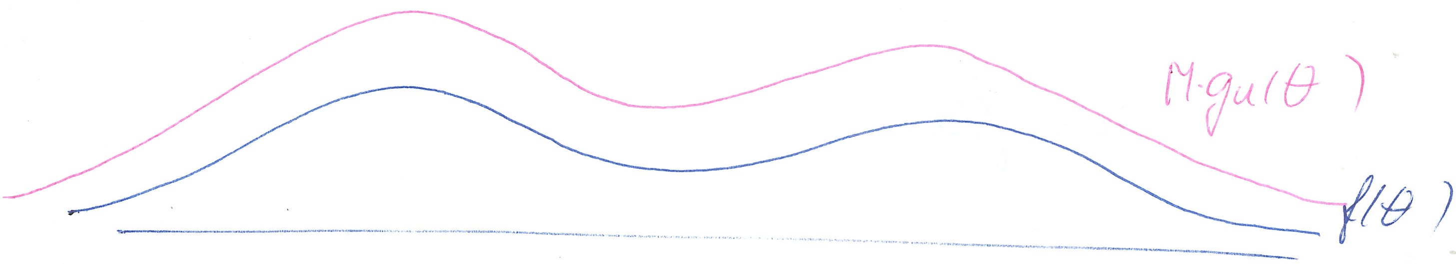
REJECTION SAMPLING

- Want to sample from $f(\theta) = c\, g(\theta)$
  - c can be obtained by importance sampl.
- Let $g_u(\theta)$ be a density of any distribution with support Θ (or larger) such that for some M

$$\underbrace{f(\theta)}_{c \cdot g(\theta)} \leq \underbrace{M \cdot g_u(\theta)}_{\text{envelop}} \quad \forall \theta \in \Theta \text{ (alm. everywh.)}$$

$$1 \leq M < \infty$$

Note: g_u should have heavier tails than f to be able to find M



ALGORITHM:

(a) generate $\theta \sim g_u(\theta)$

(b) accept θ with probability $\dfrac{f(\theta)}{M \cdot g_u(\theta)}$

= we generate (θ, V) uniformly from a subgraph of $M g_u(\theta)$

- $V = U \cdot M g_u(\theta)$, $U \sim U(0,1)$
- θ accepted if $U \leq \dfrac{f(\theta)}{M \cdot g_u(\theta)}$

$\rightarrow$ accepted pairs are sampled uniformly from a subgraph of $f(\theta)$

Probability of acceptance

$$P(\text{accept}) = P\left(U \le \frac{f(\theta)}{M g_u(\theta)}\right) =$$

$$= \mathbb{E}\left(P\left(U \le \frac{f(\theta)}{M g_u(\theta)} \,\middle|\, \theta\right)\right) =$$

θ with a density $g_u(\theta)$

$$= \mathbb{E}_{g_u} \frac{f(\theta)}{M g_u(\theta)} = \int \frac{f(\theta)}{M g_u(\theta)} g_u(\theta)\, d\lambda(\theta) = \frac{1}{M}$$

$\Rightarrow$ small M, larger acceptance (better)

NOTE: one set of sampled values from g_u can be used 1) to calculate c (importance sample.)

2) to decide which values accepted by rejection sampling

ADAPTIVE REJECTION SAMPLING

- attempt to get $g_u(\theta) \approx f(\theta) = c\,g(\theta)$

ASSUMPTIONS: $f(\theta)$ (and $g(\theta)$) log-concave,
i.e. $\log f(\theta)$ (and $\log g(\theta)$) concave

MAKE $g_u(\theta)$ = piecewise exponential distrib.

$\log g_u(\theta)$ = piecewise linear
obtained by using tangents
to $\log f(\theta)$ (or $\log g(\theta)$)