Note Title 4/16/2006

Weighted Sampling

Weighted Sample

A set of weighted samples $\{(\underline{x}^{(j)}, \omega^{(j)}): j=1 \text{ to } m\}$ is called proper w.r.t. π if for any square integrable function $h(\cdot)$.

$$E[h(x^{(j)})\,\omega^{(j)}] = c\, E_\pi h(\underline{x}), \quad j = 1 \text{ to } m.$$

c is a normalization constant

joint distribution $g(\omega, x)$ for weights & samples.

Requires: $$E_g\left\{\frac{h(x)\,\omega}{E_g(\omega)}\right\} = E_\pi\{h(x)\}$$

which implies $$\frac{E_g(\omega|x)}{E_g(\omega)}\, g(x) = \pi(x) \quad (*)$$

$(*)$ is a N & S condition.

Importance Sampling is a special case where ω is a deterministic function of x (i.e. $\omega = \pi(x)/g(x)$).

Note: express $g(\omega, x) = g(\omega|x)\, g(x)$

Now, we show connection to rejection sampling.

(page 2) Verify that weighted sampling is unbiased (maybe require normalization of the weights).

Require:

$$\sum_{x,\omega} \omega h(x) g(x,\omega) = c \sum_{x} \pi(x) h(x)$$

where $g(x,\omega) = g(\omega|x) g(x)$.

If we want this to be true for any $h(x)$, then we need

$$\sum_{\omega} \omega g(x,\omega) = c\pi(x)$$

with normalization

$$\sum_{x,\omega} \frac{\omega h(x) g(x,\omega)}{\sum_{x,\omega} \omega g(x,\omega)} = \sum_{x} \pi(x) h(x)$$

Alternatively, need $\dfrac{\sum_{\omega} \omega g(x,\omega)}{\sum_{\omega,x} \omega g(x,\omega)} = \pi(x)$

Clearly, Importance sampling is a special case

set $g(\omega|x) = \delta\left(\omega - \dfrac{\pi(x)}{g(x)}\right)$

Rejection Sampling as Weighted Sampling

Recall Rejection Sampling.

$l(x) = c\pi(x)$ is known (c unknown)

Pick $g(x)$ & M

s.t. $Mg(x) \geq l(x)$

Sample in two stages:

(i) sample x from $g(x)$

(ii) accept sample with prob $r(x) = l(x)/Mg(x)$

In context of weighted sampling.

$$g(x, w) = g(w \mid x)\, g(x)$$

Assign sample x a weight w $\quad w = 1$ or $w = 0$

Define: $g(w=1 \mid x) = \dfrac{l(x)}{Mg(x)}$ $\quad$ $g(w=0 \mid x) = 1 - g(w=1 \mid x)$

$$\sum_{x,w} w\, h(x)\, g(x,w) = \sum_x h(x)\, g(w=1 \mid x)\, g(x)$$

As required

$$= \sum_x h(x) \frac{l(x)}{Mg(x)} g(x) = \frac{c}{M} \sum_x h(x)\pi(x)$$

Also $\displaystyle\sum_{x,w} w\, g(x,w) = \sum_x g(w=1 \mid x)\, g(x) = \frac{c}{M} \sum_x \pi(x) = \frac{c}{M}$

Hence: $\dfrac{1/m \sum_{i=1}^{m} w^i h(x^i)}{1/m \sum_{i=1}^{m} h(x^i)}$ is an estimator of $\sum_x \pi(x) h(x)$

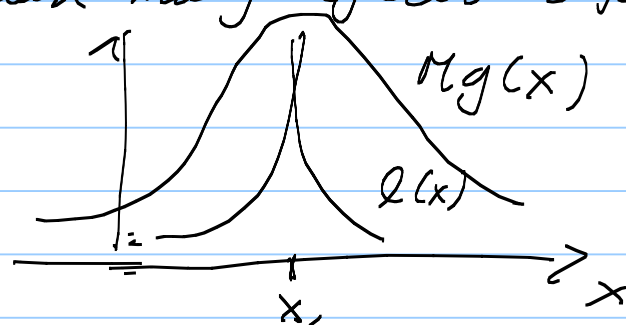
Rejection Sampling

Control Rejection Sampling.

Want to get the best aspects of rejection and importance sampling.

Problem with Rejection Sampling — enforcing condition $\frac{l(x)}{Mg(x)} \leq 1$ is problematic. Strict enforcement may cause many rejections for some values of x. e.g.

Satisfying condition at x_L means high rejection rate elsewhere.



Problem with Importance Sampling

Samples with small weights contribute little to the estimate, but require evaluating $h(x)$ — which can be computationally expensive

Combine Rejection & Importance Sampling

- Relax the requirement $\frac{l(x)}{Mg(x)} \leq 1$.
- Reject samples with small weights. (don't waste time evaluating $h(x)$)

Control Rejection Sampling.

Sample x from $g(x)$.

Define $r(x) = \min\left\{1, \frac{\pi(x)}{c\,g(x)}\right\}$

"c" can be anything (i.e. not just the normalization constant of $\pi(x)$).

Define:

$$g(\omega \mid x) = \delta\left(\omega - \frac{\pi(x)}{c\,g(x)} \cdot \frac{q_c}{r(x)}\right) r(x) + \delta(\omega)\{1 - r(x)\}.$$

Delta Function.
Indicator

i.e. ω can take value 0 (reject)

or $\frac{\pi(x)}{c\,g(x)} \cdot \frac{q_c}{r(x)}$ (importance weight)

q_c - constant.

$$\sum_{x,\omega} h(x)\,\omega\, g(\omega \mid x)\, g(x)$$

$$= \sum_x h(x) \frac{\pi(x)}{c\,\cancel{g(x)}} \cdot \frac{q_c}{\cancel{r(x)}} \cancel{r(x)}\, \cancel{g(x)}$$

$$= \frac{q_c}{c} \sum_x h(x)\pi(x).$$

Also $\sum_{x,\omega} \omega\, g(\omega \mid x)\, g(x) = \sum_x \frac{\pi(x)}{c\,\cancel{g(x)}} \cdot \frac{q_c}{\cancel{r(x)}} \cancel{r(x)}\, \cancel{g(x)}$

$$= q_c / c.$$

weight $\frac{\pi(x)}{c\,g(x)} \cdot \frac{q_c}{r(x)}$.

(Page 6)

Control Rejection Sampling

Here:

$$\frac{\sum_{x,\omega} h(x)\,\omega\, g(\omega|x)\, g(x)}{\sum_{x,\omega} \omega\, g(\omega|x)\, g(x)} = \sum_x \pi(x) h(x)$$

So sample (ω, x) from

$$g(\omega, x) = g(\omega|x)\, g(x)$$

$g(\omega|x)$ defined on previous page.

Then

$$\frac{1/m \sum_{i=1}^{m} h(x^i)\,\omega^i}{1/m \sum_{i=1}^{m} \omega^i}$$

will converge to $\sum_x \pi(x) h(x)$ as $m \to \infty$.

(Page 7)

Control Rejection Sampling. (CRS)

An alternative way to think of this as a form of importance sampling.

Define $g^*(x) = \frac{r(x)g(x)}{q_c} = \frac{1}{q_c} \min\{g(x), \frac{\pi(x)}{c}\}$

with $q_c = \sum_x r(x)g(x)$.

Claim: Control Rejection Sampling is importance sampling with distribution $g^*(x)$ and weights $\frac{\pi(x)}{c\, g^*(x)}$.

Check: sampling from $\frac{r(x)g(x)}{q_c}$ is like sampling from $g(x)$ and then accepting the sample with probability $r(x)$. (as in CRS)

the weight given $\frac{\pi(x)}{c\, g^*(x)} = \frac{\pi(x)}{c\, g(x)} \cdot \frac{q_c}{r(x)}$.

same weight as (CRS).

Theorem (page 45 Liu)

$$\mathrm{var}_{g^*}\left\{\frac{\pi(x)}{g^*(x)}\right\} \leq \mathrm{var}_g\left\{\frac{\pi(x)}{g(x)}\right\}$$

So rejection control reduces the distance from sampling dist to target.