Lecture 7. Weighted Sampling

Note Title 4/16/2006

Weighted Sample

A set of weighted samples $\{(\underline{x}^{(j)}, \omega^{(j)}) : j = 1 \text{ to } m\}$ is called <u>proper</u> 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: $\dfrac{E_g\{h(x)\omega\}}{E_g(\omega)} = E_{\pi}\{h(x)\}$

(see next page for details) which implies $\dfrac{E_g(\omega \mid x)}{E_g(\omega)}\, g(x) = \pi(x)$ $(*)$

$(*)$ necessary and sufficient condition.

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

<u>Note</u>: express $g(\omega, x) = g(\omega \mid x)\, g(x)$

Now, we show connection to rejection sampling.

(page 2)

Proof of statements from first page.

<u>Require:</u>

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

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

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

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

<u>Clearly</u>, Important sampling is a special case

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

Dirac delta function

i.e. the weight ω is a function $\dfrac{\pi(x)}{g(x)}$ of the samples x (no randomness).

Rejection Sampling as Weighted Sampling

Recall Rejection Sampling.

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

Pick $g(x)$ & M

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

Sample in two stages:

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

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

In context of weighted sampling.

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

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

Define: $g(w=1 \mid x) = \dfrac{\ell(x)}{Mg(x)}$ $\quad \Big| \quad g(\omega=0 \mid x) = 1 - g(\omega=1 \mid x)$

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

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

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

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

Justifies rejection sampling.

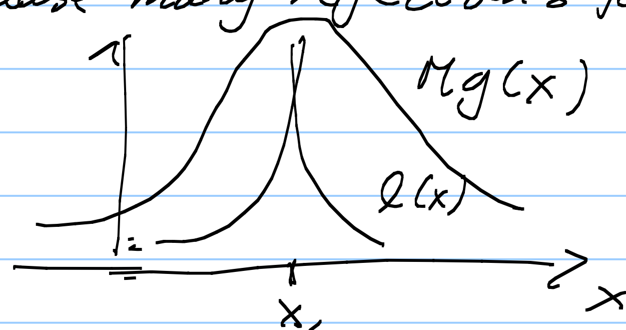
(page 4)

Rejection Control.

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

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

Satisfying condition at x_c — 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 $h(x)$)

(Page 5)

Rejection Control

Sample x from $g(x)$.

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

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

Define:

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

$\delta(\,)$ delta-function.

i.e. ω can take value 0 (reject)

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

q_c - constant.

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

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

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

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

$$= q_c/c.$$

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

(Page 6)

Rejection Control.

<u>Hence</u>:
$$\frac{\sum_{x,\omega} h(x)\,\omega\, g(\omega|x)\, g(x)}{\sum_{x,\omega} \omega\, g(\omega|x)\, g(x)} = \sum_{x} \pi(x) h(x)$$

So sample (ω, x) from
$$g(\omega, x) = \underset{\substack{\uparrow \\ \text{defined on previous page.}}}{g(\omega|x)}\, g(x)$$

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) **Rejection Control.** (Note: error in lecture)

Define $g^*(x) = \frac{r(x)g(x)}{q_c}$ with $q_c = \sum_x r(x)g(x)$.

Then $g(w|x)g(x) = q_c\,\delta\left(w - \frac{\pi(x)}{c\,g^*(x)}\right) g^*(x)$

$+ \delta(w)\{1 - r(x)\}\,g(x)$

So Rejection Control (RC) is like importance sampling from $g^*(x)$, but with the possibily of assigning zero weight (second term)

Note: 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 (RS)

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 is better than sampling from $g()$.