Solution: REINFORCE alg.

(4) (cont'd)

$$\nabla_\psi \mathop{E}_{s\sim q} f(s)$$

$$= \nabla_\psi \sum_s q(s) f(s)$$

} q, not f, depends on ψ.

$$= \sum_s \nabla_\psi [q(s) f(s)]$$

$$= \sum_s \left([\nabla_\psi q(s)] f(s) + q(s) \nabla_\psi f(s) \right)$$

$$= \sum_s [q(s) \nabla \log q(s)] f(s) + q(s) \nabla f(s)$$

$$= \sum_s q(s) [f(s) \nabla \log q(s)] + \nabla f(s)$$

$$= \mathop{E}_{s\sim q} [f(s) \cdot \nabla \log q(s)] + \nabla f(s)$$

Note: $\nabla q(s) \cdot \nabla \log q(s) = \dfrac{\nabla q(s)}{q(s)}$

$$\approx \frac{1}{L} \sum_{\ell=1}^{L} \left[f(s_\ell^\ell) \nabla \log q(s^\ell) + \nabla f(s^\ell) \right]$$

where $s^\ell \sim q$

~~$\frac{1}{L} \sum_{\ell=1}^{L} \nabla f(s^\ell)$ is fine.~~

but $\frac{1}{L} \sum_{\ell=1}^{L} f(s^\ell) \nabla \log q(s^\ell)$

has high variance.

ie, its value varies widely depending on the samples s^ℓ.

Why?

Example

why?

suppose $\alpha = 1$, $s \in \{0, 1\}$

$$q(s=1) = \alpha$$

$$q(s=0) = 1 - \alpha$$

$$\therefore \nabla_\alpha \log q(s=1) = \frac{1}{\alpha} > 0$$

$$\nabla_\alpha \log q(s=0) = \frac{-1}{1-\alpha} < 0$$

if $\alpha = 1$, then

$f(s) \nabla \log q(s)$ has 2 values

$f(1) / \alpha$

$-f(0) / (1-\alpha)$

if $f(1)$ + $f(0)$ are both large positive numbers, then $f(1)/2 >> 0$

& $-f(0)/(1-\alpha) << 0$

ie, the value of $f(s) \nabla \log q(s)$ varies widely, even if $f(s)$ does not.

More Generally

$$\frac{1}{L}\sum_{\ell=1}^{L} f(s^\ell)\,\nabla \log q(s^\ell)$$

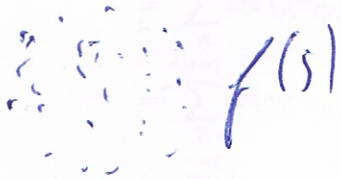
(s^ℓ: discrete vector.)

has high variance because $f(s^\ell)$ is large (& negative) for all s^ℓ, because it is a likelihood, whereas $\nabla \log q(s^\ell)$ is +ve for some s^ℓ & −ve for others, since q is a distribution, so an increase in q for some s^ℓ must result in a decrease for other s^ℓ.

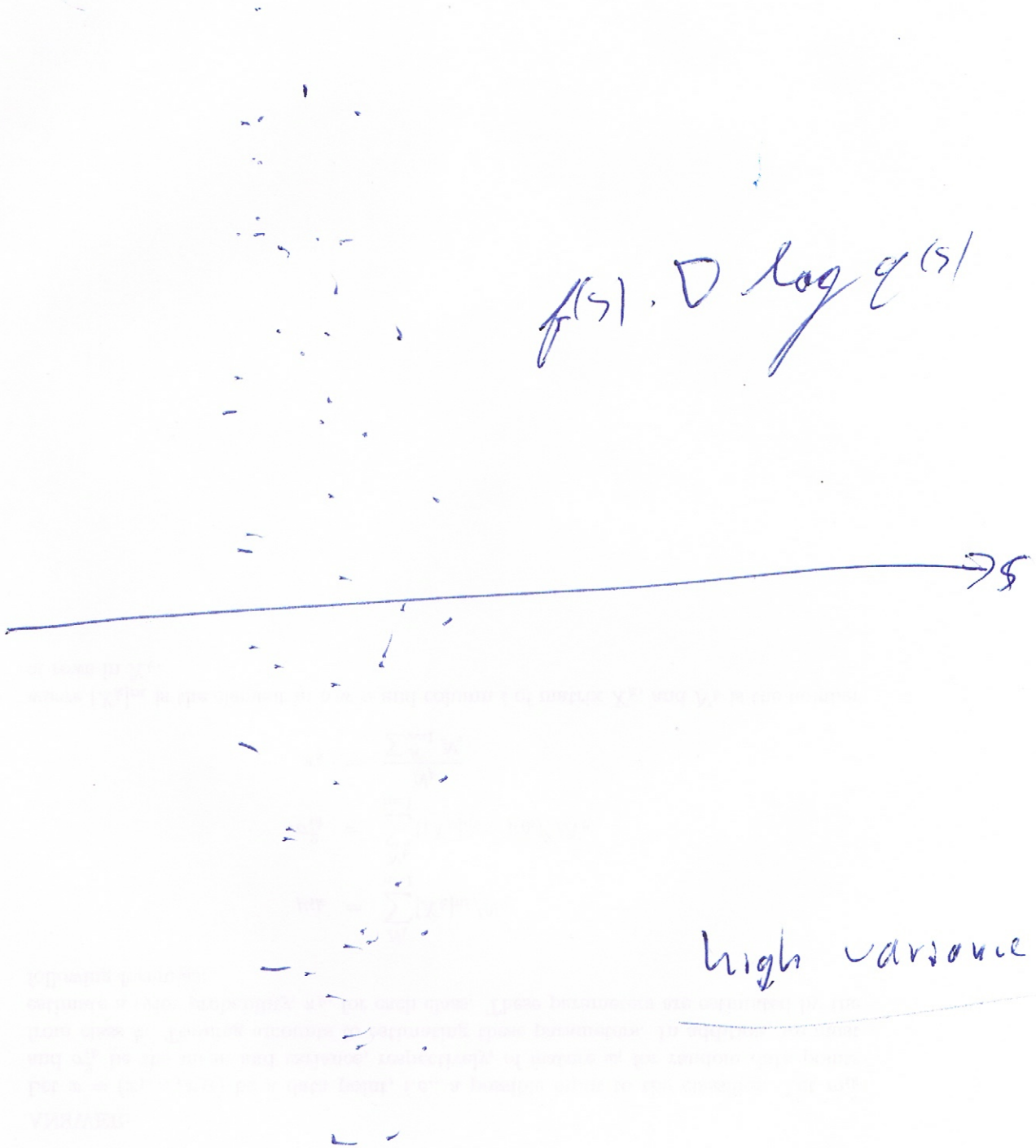
So, $f(s^\ell)\cdot \nabla \log q(s^\ell)$ is large & pos. for some s^ℓ, & large & neg for other s^ℓ.

ie, it has high variance.

In practice



22.3



Centering

We avoid this by centering f.

i.e., replace $f(s)$ by $f(s) - b$

for some constant, b.

note: $\nabla \underset{s \sim q}{E} [f(s) - b]$

$= \nabla \left(\left[\underset{s \sim q}{E} f(s) \right] - b \right)$

$= \nabla \underset{s \sim q}{E} f(s)$

Choose b so that $f(s) - b$ does not have large values of same sign as ∇

$f(1) \cdot \frac{f(1) - b}{2} = - \frac{f(0) - b}{1 - 2}$

$(f(1) - b) \nabla \log q(1)$

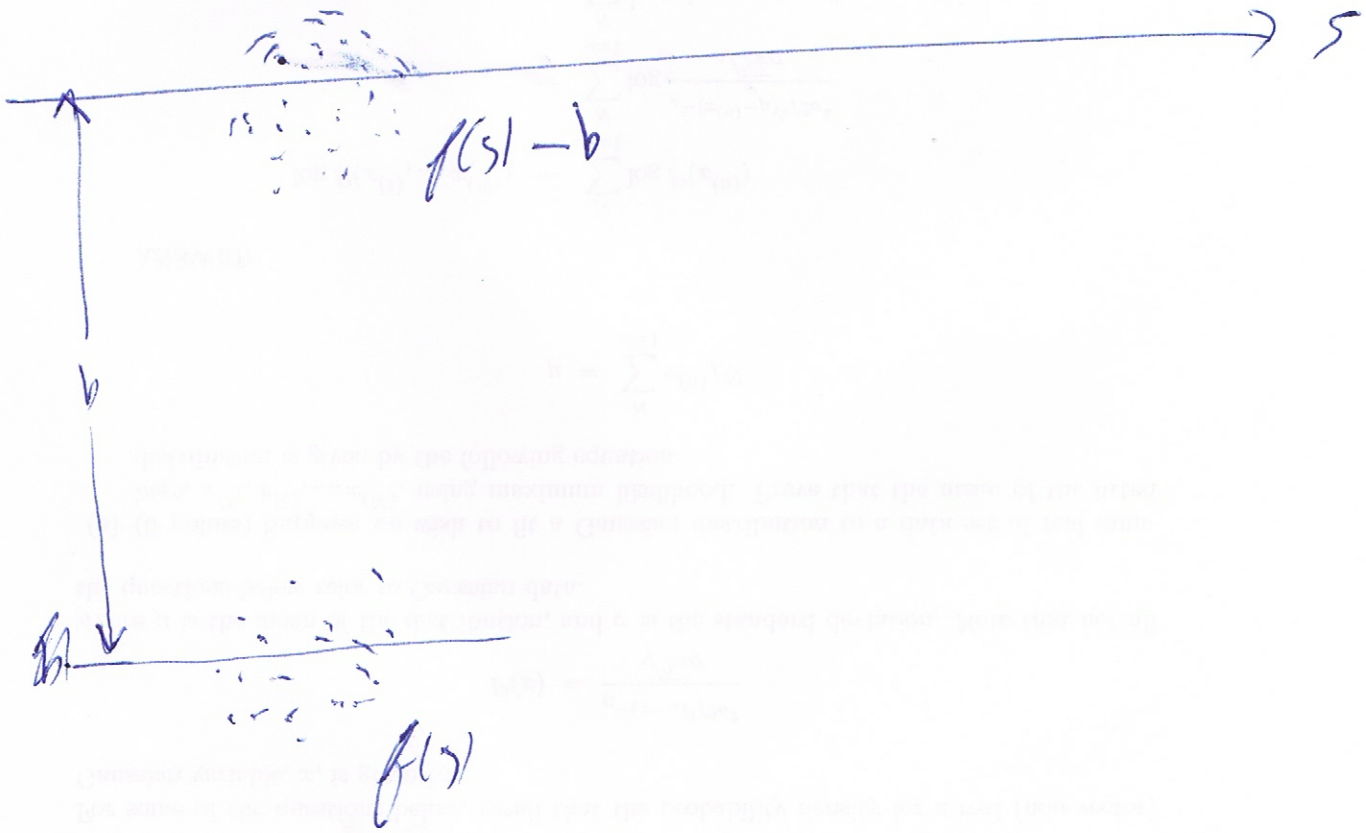
$= (f(0) - b) \nabla \log q(0)$

(exercise) (exercise)

We choose b so that $f(s) - b$ has values near 0. This gives

$(f(s) - b) \nabla \log q(s)$ low variance.

eg. choose $b = \frac{1}{L} \sum_{i=1}^{L} f(s^i)$.



(But what if $L = 1$?)

~~In our case,~~

This is an example of a variance reduction technique.

Statistics has many.

(eg. Control variates.)

Neural Variational Inference

Recall:

$$\mathcal{L} = \sum_n \mathop{E}_{s \sim q(s|x^n)} f_\phi(s, x^n)$$

$$\nabla_\phi \mathcal{L} = \sum_n \nabla_\phi \mathop{E}_{s \sim q(s|x^n)} f_\phi(s, x^n)$$

f & q depend on ϕ

$$= \sum_n \mathop{E}_{s \sim q(s|x^n)} \left[f_\phi(s, x^n) \cdot \nabla_\phi \log q(s|x^n) \right]$$

can't sample

$$= \sum_n \mathop{E}_{s \sim q} \left[f(s, x^n) - b_n \right] \cdot \nabla \log q(s|x^n)$$

b_n depends on x^n, just since f + q do, ie, $b_n = b(x^n)$

Use a neural net with input x to learn a good value of b. (See paper)

b is called a baseline

of $\nabla \mathcal{L}$

The variance is high if $f(s, x^u)$ tends to be large & of the same sign for most values of s.

So, NVI tries to shrink the average value of $[f(s, x^u) - b(x^u)]^2$, ~~averaged over s~~.

ie, to make $b(x^u) \approx \underset{s}{E} f(s, x^u)$.

Note: $f(s) \cdot \nabla \log q(s)$

can be computed even if f is non-differentiable.

If, we can estimate

$\nabla_\psi \underset{s \sim q}{E} f(s) \cdot \log q(s)$

without having to backpropagate through f.