Implicit Sampling:

Suppose target distribution can be written as

$$p(x) \propto \exp(-F(x))$$

p is Gaussian if F is quadratic, and then F can be written as

$$F(x) = \phi + \frac{1}{2}(x-\mu)^T H(x-\mu)$$

$$\phi = \min F(x)$$

$$H = \nabla^2 F$$

For non-Gaussian p, expand F around its minimizer

$$F(x) = \underbrace{\phi + \tfrac{1}{2}(x-\mu) H (x-\mu)}_{\substack{\text{quadratic approximation} \\ \text{of } F. \\ \text{Call this } Q(x)}} + \mathcal{O}(\varepsilon^{1/2})$$

Idea: use Gaussian proposal

$$q(x) = \exp(-Q(x))$$

Then weights are:

$$w(x) \propto \frac{\exp(-F(x))}{\exp(-Q(x))} \propto \exp(-(F(x) - Q(x)))$$

$$\propto \exp(-\mathcal{O}(\varepsilon^{1/2}))$$

And a long calculation shows: $\rho = \dfrac{E(w^2)}{E(w)^2} = 1 + \mathcal{O}(\varepsilon).$

$\Rightarrow$ Implicit sampling is "good" for near Gaussian problems.

Algorithm:

(1) Solve optimization problem: $\min F(x)$

Result: $\mu = \arg\min F(x)$

$\phi = \min F(x)$

$H = \nabla^2 F$

(2) Sample proposal: $q(x) = \mathcal{N}(\mu, H^{-1})$

(3) Compute weights: $w(x) \propto \exp(-(F(x) - Q(x)))$

↳ weighted ensemble $\{x_i, w_i\}$ has distribution p.

↳ "has distribution p" mean that

$$\sum f(x_i) w_i \to E_p[f(x)]$$

Alternative algorithm:

$q(x) \sim$ Multivariate t-distribution with parameters μ, H^{-1}.

→ This can be a bit better and more stable (Owen)

For the Gaussian Case: how do we get

$$\rho = 1 + O(\varepsilon) \ ?$$

Variance Lemma

Definitions: $$u(x) = 1 + \varepsilon^r u_1(x) + \varepsilon^{2r} u_2(x) + O(\varepsilon^{3r})$$

$$Q = \frac{E[u^2(x)]}{E[u(x)]} - 1$$

Then: $$Q = \varepsilon^{2r} \operatorname{var}(u_1) + O(\varepsilon^{2r})$$

Proof:

$$E[u(x)] = 1 + \varepsilon^r E[u_1] + \varepsilon^{2r} E[u_2] + O(\varepsilon^{3r})$$

$$\begin{aligned} E[u(x)]^2 = {} & 1 + \varepsilon^r E[u_1] + \varepsilon^{2r} E[u_2] \\ & + \varepsilon^r E[u_1] + \varepsilon^{2r} E[u_1]^2 \\ & + \varepsilon^{2r} E[u_2] + O(\varepsilon^{3r}) \end{aligned}$$

$$= 1 + \varepsilon^r 2E[u_1] + \varepsilon^{2r} (E[u_1]^2 + 2E[u_2]) + O(\varepsilon^{3r})$$

$$\begin{aligned} u(x)^2 = {} & 1 + \varepsilon^r u_1(x) + \varepsilon^{2r} u_2(x) \\ & + \varepsilon^r u_1(x) + \varepsilon^{2r} u_1(x)^2 \\ & + \varepsilon^{2r} u_2(x) + O(\varepsilon^{3r}) \end{aligned}$$

$$= 1 + \varepsilon^r 2u_1(x) + \varepsilon^{2r} (u_1(x)^2 + 2u_2(x)) + O(\varepsilon^{3r})$$

$$E[u(x)^2] = 1 + 2\varepsilon^r E[u_1] + \varepsilon^{2r} (E[u_1^2] + 2E[u_2]) + O(\varepsilon^{3r})$$

$$Q = \frac{1 + 2\varepsilon^r E[u_1] + \varepsilon^{2r}\left(E[u_1^2] + 2E[u_2]\right) + \mathcal{O}(\varepsilon^{3r})}{1 + 2\varepsilon^r E[u_1] + \varepsilon^{2r}\left(E[u_1]^2 + 2E[u_2]\right) + \mathcal{O}(\varepsilon^{3r})} - 1$$

<u>Recall:</u> $\frac{1}{1+x} \approx 1 - x + x^2$

$$\rightarrow \frac{1}{1 + \underbrace{2\varepsilon^r E[u_1] + \varepsilon^{2r}\left(E[u_1]^2 + 2E[u_2]\right) + \mathcal{O}(\varepsilon^{2r})}_{x}}$$

$$\approx 1 - 2\varepsilon^r E[u_1] - \varepsilon^{2r}\left(E[u_1]^2 + 2E[u_2]\right) + \varepsilon^{2r}\, 4E[u_1]^2 + \mathcal{O}(\varepsilon^{3r})$$

$$= 1 - 2\varepsilon^r E[u_1] + \varepsilon^{2r}\left(3E[u_1]^2 - 2E[u_2]\right) + \mathcal{O}(\varepsilon^{3r})$$

$$\Rightarrow Q = \left(1 + \varepsilon^r 2E[u_1] + \varepsilon^{2r}\left(E[u_1^2] + 2E[u_2]\right) + \mathcal{O}(\varepsilon^{3r})\right)$$

$$\times \left(1 - \varepsilon^r 2E[u_1] + \varepsilon^{2r}\left(3E[u_1]^2 - 2E[u_2]\right) + \mathcal{O}(\varepsilon^{3r})\right) - 1$$

$$= \cancel{1 - \varepsilon^r 2E[u_1]} + \varepsilon^{2r}\left(3E[u_1]^2 \cancel{- 2E[u_2]}\right) \cancel{+ \varepsilon^r 2E[u_1]} - \varepsilon^{2r}\, 4E[u_1]^2 + \varepsilon^{2r}\left(E[u_1^2] \cancel{+ 2E[u_2]}\right) + \mathcal{O}(\varepsilon^{3r}) \cancel{-1}$$

$$= \varepsilon^{2r}\left(E[u_1^2] - E[u_1]^2\right) + \mathcal{O}(\varepsilon^{3r})$$

$$= \varepsilon^{2r}\, \mathrm{var}(u_1) + \mathcal{O}(\varepsilon^{3r})$$

<u>Note:</u> distribution we use to take averages irrelevant

Use this result on the sampling method:

$$q(x) \propto \exp(-\tfrac{1}{2}(\underline{x}-\mu)H(x-\mu)) \propto \exp(-Q(x))$$

$$p(x) \propto \exp(-F(x))$$

$$F(x) = \phi + \tfrac{1}{2}(\underline{x}-\mu)H(x-\mu) + \varepsilon^{1/2} \underbrace{c_3 + \varepsilon\, c_4}_{\text{Coeffs of Taylor series!}} + O(\varepsilon^{3/2})$$

$$w(x) \underset{\propto \frac{p(x)}{q(x)}}{\propto} \exp(-(F(x) - Q(x))) \propto \exp(-\varepsilon^{1/2}c_3 - \varepsilon\, c_4 + O(\varepsilon^{3/2}))$$

Recall: $\exp(-x) \approx 1 - x + x^2/2$

$$w(x) \approx 1 - \varepsilon^{1/2} c_3 - \varepsilon\, c_4 + \tfrac{1}{2}\varepsilon\, c_3^2 + O(\varepsilon^{3/2})$$

$$= \underbrace{1 - \varepsilon^{1/2} c_3 + \varepsilon(\tfrac{1}{2}c_3^2 - c_4) + O(\varepsilon^{3/2})}_{\text{use this as "u(x)" in Variance lemma!}}$$

$$\hookrightarrow \underline{Q = \varepsilon \operatorname{var}(c_3) + O(\varepsilon^{3/2})}$$

$$\rho = 1 + Q \quad \Rightarrow \quad \underline{\rho = 1 + \varepsilon \operatorname{var}(c_3) + O(\varepsilon^{3/2})}$$

Symmetrization

Algorithm idea: $\cancel{x} \sim N(\mu, H^{-1})$

then getting $x^+ = \mu + H^{-\frac{1}{2}}\xi$, $\xi \sim N(0, I)$

or $x^- = \mu - H^{-\frac{1}{2}}\xi$

is equally likely.

To make life easy, we switch coordinates : $\tilde{x} = H^{\frac{1}{2}}(x - \mu)$

$\Rightarrow x \sim N(0, I)$

Then : proposal distribution : $q(x) = N(0, I)$

weight : $w(x) = \frac{p(x)}{q(x)}$

draw $x \sim q$

Pick x with probability $p^+ = \frac{w(x)}{w(x) + w(-x)}$

$-x$ with probability $p^- = \frac{w(-x)}{w(x) + w(-x)}$

What is distribution, $q_s(x)$, of samples generated in this way?

↳ There are two ways of getting x:

propose x, chose x & propose $-x$, chose $-(-x)$

$$\Rightarrow q_s(x) = q(x)\frac{w(x)}{w(x)+w(-x)} + \underbrace{q(-x)}_{=q(x)}\underbrace{\frac{w(-(-x))}{w(x)+w(-x)}}_{=\frac{w(x)}{w(x)+w(-x)}}$$

$$q_s(x) = 2q(x)\frac{w(x)}{w(x)+w(-x)}$$

The weights thus are:

$$w_s(x) = \frac{p(x)}{q_s(x)} = \frac{1}{2}\frac{\left(w(x)+w(-x)\right)}{w(x)}\underbrace{\frac{p(x)}{q(x)}}_{w(x)} = \frac{1}{2}\left(w(x)+w(-x)\right)$$

What is $\rho_s(x) = \dfrac{E[\omega_s(x)^2]}{E[\omega_s(x)]^2}$?

Assuming, as before that $p(x) \propto \exp(-F(x))$

$$F(x) = \phi + \tfrac{1}{2}(x-\mu)H(x-\mu) + \varepsilon^{1/2} C_3(x) + \varepsilon C_4(x) + \mathcal{O}(\varepsilon^{3/2})$$

and that $q(x) = \exp(-Q(x))$

$$Q(x) = \phi + \tfrac{1}{2}(x-\mu)H(x-\mu)$$

We know that:

$$\omega(x) = 1 - \varepsilon^{1/2} C_3(x) + \varepsilon\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right)$$

$$\omega(-x) = 1 - \varepsilon^{1/2} C_3(-x) + \varepsilon\left(\tfrac{1}{2}C_3^2(-x) - C_4(-x)\right)$$

We also know that $C_3(x) = -C_3(-x)$

$C_4(x) = C_4(-x)$

Thus:

$$\omega_s(x) = \tfrac{1}{2}\left(\omega(x) + \omega(-x)\right)$$

$$= \tfrac{1}{2}\Big(1 - \cancel{\varepsilon^{1/2} C_3(x)} + \varepsilon\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right)$$

$$+ 1 + \cancel{\varepsilon^{1/2} C_3(x)} + \varepsilon\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right) + \mathcal{O}(\varepsilon^{3/2})\Big)$$

$$= 1 + \varepsilon\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right) + \mathcal{O}(\varepsilon^{3/2})$$

By variance lemma:

$$Q = \varepsilon^2 \operatorname{var}\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right) + \mathcal{O}(\varepsilon^{5/2})$$

$$\Rightarrow \rho_s = 1 + \varepsilon^2 \operatorname{var}\left(\tfrac{1}{2}C_3^2(x) - C_4(x)\right) + \mathcal{O}(\varepsilon^{5/2})$$