Implicit sampling: drawing samples by solving eqns

(Non-Gaussian proposal)

Observation: we can draw samples by solving algebraic equations.

Example: Solve:

$$\frac{1}{2}(x-\mu)^T C^{-1}(x-\mu) = \frac{1}{2}\xi^T\xi \quad , \quad \xi \sim N(0,I)$$

for x. Solution:

$$x = \mu + C^{+1/2}\xi \quad \rightarrow x \sim N(\mu, C)$$

↳ works well for Gaussian, but how about other distributions?

Reference dist.
↓ solve eqns
Proposal dist
↓ weights
Target

Set up:

Target distribution: $p(x) \propto \exp(-F(x))$

Reference distribution: $g(x) \propto \exp(-G(x))$

Assume: $G(\xi) \geq 0$, e.g. $G(\xi) = \frac{1}{2}\xi^T C^{-1}\xi$.

$\xi \sim g(\cdot)$

Try to solve: $\underbrace{F(x) - \phi}_{\geq 0} = \underbrace{G(\xi)}_{\geq 0}$, $\phi = \min F(x)$

→ "Should have a solution, possibly many"

→ 1 eqn in n variables.

→ we need more eqns!

Try this: $G(x) = \frac{1}{2}x^T C^{-1} x$, $C = LL^T$, $\phi = \min F$, $\mu = \arg\min F$.

$$\left.\begin{aligned} F(x) - \phi &= \frac{1}{2}\xi^T C^{-1}\xi \\ x &= \mu + \lambda L\xi \end{aligned}\right\} \text{ n+1 eqns in n+1 variables!}$$

This set of eqns is "easy" to solve.

$$F(x+\lambda L\xi) - \phi = \tfrac{1}{2}\xi^T C^{-1}\xi \quad \rightarrow \text{ solve scalar eqn for scalar } \lambda.$$

<u>Question:</u> What is the distribution of samples generated by solving these eqns repeatedly?

$$q(x) \propto g(\xi) \left| \frac{dx}{d\xi} \right| \qquad \text{(Change of vars!)}$$

$$g(\xi) \propto \exp\left(-\tfrac{1}{2}\xi^T C^{-1}\xi\right)$$

$$\left| \frac{dx}{d\xi} \right| = ?$$

$$x = \mu + \lambda L \xi$$

a row vector!

$$\frac{dx}{d\xi} = \lambda L^T + (L^T\xi)\frac{d\lambda}{d\xi}$$

$$\frac{\partial\lambda}{\partial\xi} = \frac{\partial\lambda}{\partial\rho}\frac{\partial\rho}{\partial\xi}, \qquad \rho = \tfrac{1}{2}\xi^T C^{-1}\xi$$

$$\frac{\partial\rho}{\partial\xi} = C^{-1}\xi$$

Scalar!

$$\Rightarrow \frac{\partial\lambda}{\partial\xi} = \frac{\partial\lambda}{\partial\rho}\cdot\left(C^{-1}\xi\right)^T$$

$$\Rightarrow \frac{dx}{d\xi} = \lambda L^T + (L^T\xi)\frac{\partial\lambda}{\partial\rho}\left(C^{-1}\xi\right)^T$$

$$= L^T\left(\lambda I + \frac{\partial\lambda}{\partial\rho}\,\xi\left(C^{-1}\xi\right)^T\right)$$

$$\left|\frac{dx}{d\xi}\right| = \left|\det\left(L^T\left(\lambda I + \frac{\partial\lambda}{\partial\xi}\,\xi\,(C^{-1}\xi)^T\right)\right)\right|$$

$$= \det(L)\cdot\left|\det\left(\lambda I + \frac{\partial\lambda}{\partial\xi}\,\xi\,(C^{-1}\xi)^T\right)\right|$$

Recall: $\det(A+BC') = \det(A)\,\det(I + BA^{-1}C')$

$$= \det(L)\cdot\det(\lambda I)\,\det\Big(1 + \underbrace{(\xi^T C^{-1}\xi)}_{=2\varrho}\frac{\partial\lambda}{\partial\xi}\cdot\frac{1}{\lambda}\Big)$$

$$= \det(L)\left|\lambda^n\left(1 + 2\tfrac{\varrho}{\lambda}\frac{\partial\lambda}{\partial\xi}\right)\right|$$

$$= \underline{\det(L)\left|\lambda^{n-1}\left(\lambda + 2\varrho\frac{\partial\lambda}{\partial\xi}\right)\right|}$$

↳ Distribution of Samples:

$$q(x) \propto \exp\Big(\underbrace{-\tfrac{1}{2}\,\xi^T C^{-1}\xi}_{=F(x)-\phi}\Big)\left|\frac{dx}{d\xi}\right|$$

$$\underline{q(x) \propto \exp(-F(x))\left|\lambda^{n-1}\left(\lambda + 2\varrho\frac{\partial\lambda}{\partial\xi}\right)\right|}$$

- **Not** a Gaussian
- Samples are "easy" to obtain -> solve scalar eqn
- $q(x)$ is "easy to compute -> scalar derivative.

More on computing the derivative:

$\frac{\partial \lambda}{\partial \rho} = ?$ Try implicit differentiation.

$$F(x) - \phi = \frac{1}{2}\xi^T C^{-1}\xi$$

$$F(\mu + \lambda L\xi) - \phi = \rho$$

Take derivatives:

$$\nabla F \cdot \underbrace{\frac{\partial}{\partial \rho}(\mu + \lambda L\xi)}_{\frac{\partial \lambda}{\partial \rho} L\xi} = 1$$

$$\Rightarrow \quad \frac{\partial \lambda}{\partial \rho} = \frac{1}{\nabla F \cdot L\xi}$$

More on geometry

$F(x) - \phi = \frac{1}{2}\xi^T C^{-1}\xi$

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

→ choosing a ξ fixes a level set of F

→ we look for an x such that $F(x) = \phi + \frac{1}{2}\xi^T C^{-1}\xi$

→ we look in direction $L\xi$

↳ $x = \mu + \lambda L\xi$

→ all we need is be able to compute ∇F, which we do already when minimizing F.

Implicit sampling algorithm.

(1) Solve optimization problem $\phi = \min F$ $\quad \mu = \arg\min F$

(2). Solve $\quad F(x) - \phi = \frac{1}{2}\xi^T C^{-1}\xi \quad , \quad \xi \sim N(0, C)$

$$x = \mu + \lambda L\xi$$

(3) proposal distribution: $q(x) \propto \exp(-F(x)) \left| \lambda^{n-1}\left(\lambda + 2\rho \frac{\partial \lambda}{\partial \rho}\right) \right|$

(4) weights: $\quad w \propto \frac{p(x)}{q(x)} \propto \left| \lambda^{n-1}\left(\lambda + 2\rho\frac{\partial \lambda}{\partial \rho}\right) \right|$

How to choose the matrix L?

Try Gaussian case: $F(x) = \frac{1}{2}(x-\mu) H (x-\mu)$, $\phi = 0$

$$F(x) - \phi = \frac{1}{2}\xi^T C \xi$$

$$x = \mu + \lambda L \xi$$

$$\frac{1}{2}\lambda^2 \xi^T L^T H L \xi = \frac{1}{2}\xi^T C^{-1} \xi$$

pick: $C = I$

$H = L^{-T} L^{-1}$ $\quad \left(H^{-1} = L L^T \right)$

$$\Rightarrow \frac{1}{2}\lambda^2 \xi^T \underbrace{L^T L^{-T} L L^{-1}}_{=I} \xi = \frac{1}{2}\xi^T \xi$$

$$\Rightarrow \underline{\lambda = \pm 1}$$

→ For non-Gaussian: $H = \left.\frac{\partial^2 F}{\partial x^2}\right|_{x=\mu}$.

Notes:

- Implicit sampling with random maps:

$$x = \mu + \lambda L \xi, \quad \xi \sim N(0, I)$$

$H^{-1} = L L^T$

λ solves eqs

$$w(x) \propto \left|\lambda^{n-1}\left(\lambda + 2\xi \frac{\partial \lambda}{\partial \xi}\right)\right|$$

- Gaussian / quadratic approximation

$$x = \mu + L\xi$$

$$w(x) \propto \exp(-(F(x) - Q(x)))$$

→ Similar! Both algorithms also similar to 4D-Var / optimization.

One can show:

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

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

$$\rho = 1 + \varepsilon \; E[\overset{\mathrm{Var}(C_3)}{\cancel{C_3^2(x)}}] \frac{(1+d)^2}{(2+d)(4+d)}$$

→ does not give much of an edge of Gaussian proposals if ε is small (problem is nearly Gaussian).

One can also symmetrize this algorithm to get

$$\rho = 1 + \varepsilon^2 \cdot (\dots)$$