Ensemble of 4D-Var (EDA or RTO)

Consider (again) a linear Gaussian problem

$$x_1 = M x_0 \quad . \quad x_0 \sim N(\mu, B)$$
$$y_1 = H x_1 + \eta \quad . \quad \eta \sim N(0, R)$$

4D-Var:

Minimize $F(x_0) = \frac{1}{2}(x_0-\mu)B^{-1}(x_0-\mu) + \frac{1}{2}(HMx_0 - y)^T R^{-1}(HMx_0 - y)$

Solution: Solve $\nabla F = 0$

$$\hookrightarrow \quad x_0^* = \mu + K(y - HM\mu)$$
$$K = BM^TH^T(HMBM^TH^T + R)^{-1}$$

Now suppose $\tilde{\mu}$ and $\tilde{y}$ are random variables with $E(\tilde{\mu}) = \mu$, $E[\tilde{y}] = y$

Define $x_0^* = \tilde{\mu} + K(\tilde{y} - HM\tilde{\mu})$

$$E[x_0] = E[\tilde{\mu}] + K(E[\tilde{y}] - HM\,E(\tilde{\mu}))$$
$$= \mu + K(y - HM\mu) \quad \checkmark$$

$$\mathrm{Cov}(x_0) = \mathrm{Cov}((I - KHM)\tilde{\mu} + K\tilde{y})$$
$$\underset{\text{independence!}}{=} (I - KHM)P_{\tilde{\mu}}(I - KHM)^T + K P_{\tilde{y}} K^T$$
$$\overset{!}{=} (I - KHM)B$$
(should be!)

Try: $P_{\tilde{\mu}} = B$
$P_{\tilde{y}} = R$

Define $\hat{H} = HM$

$\mathrm{cov}(x_0) =$

$$(I - K\hat{H}) B (I - K\hat{H})^T + KRK^T$$

$$= (B - K\hat{H}B)(I - KH)^T + KRK^T$$

$$= \underset{\sim}{B} - B\hat{H}^T K^T - \underset{\sim}{K\hat{H}B} + K\hat{H}B\hat{H}^T K^T + KRK^T$$

$$= (I - K\hat{H})B + K(\hat{H}B\hat{H}^T + R)K^T - B\hat{H}^T K^T$$

$$= (I - K\hat{H})B + \underbrace{B\hat{H}^T \underbrace{(\hat{H}B\hat{H}^T + R)^{-1}(\hat{H}B\hat{H}^T + R)}_{= I}{}^{-1} K^T - B\hat{H}^T K^T}_{= 0}$$

$$= (I - K\hat{H})B \quad \checkmark$$

<u>Result</u> we can draw samples from

$$p(x_0|y) \propto p(x_0)\, p(y|x_0) \propto \exp(-F(x_0))$$

by solving the stochastic opt. problem

$$\operatorname{argmin} \ \tfrac{1}{2}(x_0 - \tilde{\mu})B^{-1}(x_0 - \tilde{\mu}) + \tfrac{1}{2}(HMx_0 - \tilde{y})^T R^{-1}(HMx_0 - \tilde{y})$$

$$\tilde{\mu} \sim N(\mu, B)$$

$$\tilde{y} \sim N(y, R)$$

↳ similar to perturbed obs. EnKF!

We can do this for nonlinear problems as well!

$$F(x_0) = \frac{1}{2}(x_0-\mu)B^{-1}(x_0-\mu) + \frac{1}{2}(HM(x_0)-y)^T R^{-1}(HM(x_0)-y)$$

Rewrite:

$$F(x_0) = \frac{1}{2}\begin{pmatrix} B^{-1/2}(x_0-\mu) \\ R^{-1/2}(HM(x_0)-y) \end{pmatrix}^T \begin{pmatrix} B^{-1/2}(x_0-\mu) \\ R^{-1/2}(HM(x_0)-y) \end{pmatrix}$$

$$= \frac{1}{2}(f(x)-z)^T(f(x)-z)$$

where $z = \begin{pmatrix} B^{-1/2}\mu \\ R^{-1/2} y \end{pmatrix}$ $\quad f(x) = \begin{pmatrix} B^{-1/2} x_0 \\ R^{-1/2} HM(x_0) \end{pmatrix}$

→ To see what ~~sampling from~~ this means, we consider

$$p(x|y) \propto \exp\left(-\frac{1}{2}\|f(x)-y\|^2\right)$$

2 norm. $\|u\|^2 = u^T u$.

Question: What distribution do we sample when we solve

$$\arg\min \frac{1}{2}\|f(x)-\tilde{z}\|^2$$

Solve the optimization problem:

$$\arg\min F(x) = \frac{1}{2}\|f(x)-\tilde{z}\|^2 = \frac{1}{2} r(x)^T r(x)$$

$$r(x) = f(x) - \tilde{z}$$

$$\nabla F = 0$$

$\Rightarrow$ Solve: $J^T(f(x)-\tilde{z}) = 0$

write $J = QR$, $Q^TQ = I$

R upper triangular

(not our usual R, sorry!)

Note: J is not a square matrix

$\Rightarrow$ Q, R not square matrices!

$$\Leftrightarrow \quad J^T(f(x)-\tilde{z}) = 0$$

$$R^TQ^T(f(x)-\tilde{z}) = 0$$

$\downarrow$ Q is invertible

$$\Leftrightarrow \quad Q^T(f(x)-\tilde{z}) = 0$$

$$Q^T f(x) = Q^T\tilde{z}$$

Define a r.v. by $w = Q^T(z+\eta)$, $\eta \sim N(0, I)$

$$\Rightarrow p(w) = N(Q^Tz, Q^TQ) = N(Q^Tz, I)$$

$$\Rightarrow p(w) \propto \exp\left(-\tfrac{1}{2}\|w - Q^Tz\|^2\right)$$

Define another r.v. by:

$$Q^T g(\theta) = w \qquad \Rightarrow \quad \theta = g^{-1}(Qw) \quad \text{(formally)}$$

Then:

$$p(\theta) = p(w)\left|\frac{dw}{d\theta}\right| \qquad \text{// recall change of vars.}$$

$$\frac{dw}{d\theta} = \frac{\partial}{\partial\theta}\left(Q^T g(\theta)\right) = Q^T J(\theta) \quad , \quad J(\theta) = \begin{pmatrix} B^{-1/2} \\ R^{-1/2} H M' \end{pmatrix}$$

$$\Rightarrow p(\theta) = |\det(Q^T J(\theta))| \exp\left(-\tfrac{1}{2}\|w - Q^T z\|^2\right)$$

$$= |\det(Q^T J(\theta))| \exp\left(-\tfrac{1}{2}\|Q^T g(\theta) - Q^T z\|^2\right)$$

$$= \underbrace{|\det(Q^T J(\theta))| \exp\left(-\tfrac{1}{2}\|Q^T (g(\theta) - z)\|^2\right)}_{\text{not a Gaussian!}}$$

~> If "almost" looks like $p(x_0|y) \propto \exp(-\tfrac{1}{2}\|g(x_0) - z\|^2)$

~> This perhaps gives good results!

~> "This" is called EDA or RTO.

B, μ, y

↓

Repeatedly solve opt. problem with perturbed obs

↓

Propagate forward + Localize + Inflate

EDA algorithm