What about nonlinear problems?

$$x(T) = M(x_0) \qquad x_0 \sim N(\mu, B)$$

$$y = Hx(T) + \eta, \qquad \eta \sim N(0, R)$$

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

$$= \frac{1}{2} r^T r$$

$$r = \begin{pmatrix} B^{-1/2}(x-\mu) \\ R^{-1/2}(HM(x)-y) \end{pmatrix}$$

Recall: GN → iterate solution of linear problem!

$$\hat{r} = \begin{pmatrix} B^{-1/2}(x-\mu) \\ R^{-1/2}(HMx-y) \end{pmatrix}$$

$$M(x) \approx Mx \qquad \text{linearization of the model.}$$

→ Solve lin. opt. problem

$$x = \mu + K(y - HM\mu)$$

→ repeat.

<u>More detail:</u> $\quad x_{k+1} = x_k - (J^TJ)^{-1} J^T r(x_k)$

$$r = \begin{pmatrix} B^{-1/2}(x-\mu) \\ R^{-1/2}(H M(x) - y)\end{pmatrix} \qquad J = \begin{pmatrix} B^{-1/2} \\ R^{-1/2} H M \end{pmatrix}$$

$\leftarrow \left.\frac{\partial M(x)}{\partial x}\right|_{x=x_k}$ n×n matrix, see below!

$$J^TJ = (B^{-1} + M^TH^TR^{-1}HM)$$

$$(J^TJ)^{-1} = (I - K\hat H)B \quad , \quad \hat H = HM, \quad K = B\hat H^T(\hat H B \hat H^T + R)^{-1}$$

$$J^T r = B^{-1}(x_k-\mu) + M^TH^TR^{-1}(HM(x_k)-y)$$

$$(J^TJ)^{-1}J^Tr = (I-K\hat H)B\left(B^{-1}(x_k-\mu) + M^TH^TR^{-1}(HM(x_k)-y)\right)$$

$M^TH^T = \hat H^T$

$$= (I-K\hat H)(x_k-\mu) + \underbrace{(I-K\hat H)B\hat H^T R^{-1}(HM(x_k)-y)}$$

$$(I-K\hat H)B\hat H^T R^{-1} = (B\hat H^TR^{-1} - K\hat H B \hat H^T R^{-1})$$

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

$$= B\hat H^T\left(I - (\hat H B \hat H^T + R)^{-1}\hat H B \hat H^T\right)R^{-1}$$

$$= B\hat H^T(\hat H B\hat H^T+R)^{-1}\underbrace{\left((\cancel{\hat HB\hat H^T}+R) - \cancel{\hat H B\hat H^T}\right)}_{=I}R^{-1}$$

$$= B\hat H^T(\hat HB\hat H^T+R)^{-1} = K$$

$$\Rightarrow (J^TJ)^{-1}J^Tr = (I-K\hat H)(x_k-\mu) + K(HM(x_k)-y)$$

$$x_{k+1} = x_k - (J^T J)^{-1} J^T r(x_k)$$

$$= x_k - (I - K\hat{H})(x_k - \mu) - K(HM(x_k) - y)$$

$$= \cancel{x_k} - \cancel{x_k} + \mu + K\hat{H}x_k - K\hat{H}\mu - K(HM(x_k) - y)$$

$$= \mu + K\hat{H}(x_k - \mu) - K(HM(x_k) + y)$$

$$= \mu + K(\hat{H}(x_k - \mu) - HM(x_k) + y)$$

Similar to KF

How to linearize the model

discrete time model: $x_T = M(x_0)$

means: integrate from $t=0$ to $t=T$.

Example: 3 time steps with Euler discretization

$$x_1 = x_0 + \Delta t\, f(x_0)$$

$$x_2 = x_1 + \Delta t\, f(x_1)$$

$$x_3 = x_2 + \Delta t\, f(x_2)$$

$$\frac{dx_3}{dx_0} = \frac{dx_2}{dx_0} + \Delta t\, f'(x_2) \cdot \frac{\partial x_2}{\partial x_0}$$

$$= \frac{dx_1}{dx_0} + \Delta t\, f'(x_1) \frac{\partial x_1}{\partial x_0}$$

$$= I + \Delta t\, f'(x_0)$$

→ Integrate linearized eqn:

To get $\frac{dx_3}{dx_0}$:

$$\frac{dx_1}{dx_0} = I + \Delta t\, f'(x_0)$$ (note, this is a matrix)

$$\frac{dx_2}{dx_0} = \frac{dx_1}{dx_0} + \Delta t\, f'(x_1) \frac{\partial x_1}{\partial x_0}$$

$$\frac{dx_3}{dx_0} = \frac{dx_2}{dx_0} + \Delta t\, f'(x_2) \frac{\partial x_2}{\partial x_0}$$ ← this is M, the linearized model.

Cycling 4D-Var

$$x_T = M(x_0) \quad , \; x_0 \sim N(\mu, B)$$
$$y = Hx_T + \eta \; , \quad \eta \sim N(0, R)$$

4D-Var:
$$\min_{x_0} F(x_0)$$

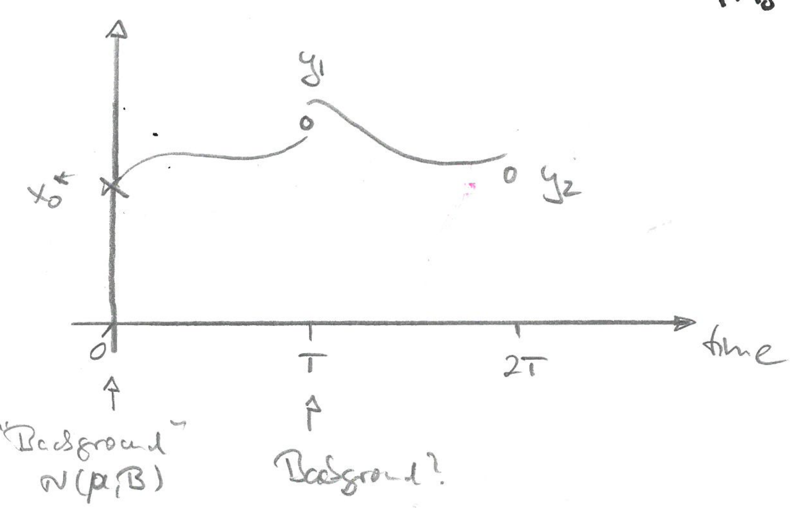
$$F(x_0) = \frac{1}{2}(x_0-\mu)^T B(x_0-\mu) + \frac{1}{2}(HM(x_0)-y)^T R(HM(x_0)-y)$$

$$= -\log p(x_0|y)$$

→ We can approximate $p(x_0|y)$ by the Gaussian:

$$N(x_0^*, H^{-1})$$

$$x_0^* = \arg\min F(x_0)$$
$$\underbrace{H = \nabla^2 F|_{x_0^*}}_{\text{You can also use approximate GN Hessian } H \approx 2J^TJ|_{x_0^*}} \quad , \text{Hessian at } x_0^*.$$



- What if a new observation comes in?
- How do we cycle this?

$p(x_0|y_1) \approx \mathcal{N}(x_0^*, H^{-1})$

$$\underbrace{p(x_1|y_1) \approx \mathcal{N}(Mx_0^*, MH^{-1}M^T)}$$

this is the "new" Background for obs y_2

↳ $\mu = Mx_0^*$
$B = MH^{-1}M^T$

↳ Now we can cycle this.

<u>Algorithm:</u>

Given: μ, B, y, R

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

$\rightarrow x_0^*, H (\approx 2J^TJ)$

↳ update background: $\mu \leftarrow M(x_0^*)$
$B \leftarrow MH^{-1}M^T$

↳ repeat.

Issues: * 2 Approximations

↳ $p(x_0|y_1)$ is <u>not</u> Gaussian

↳ $p(x_1|y_1)$ is not Gaussian.

→ Difficult to justify.

→ "often" works, sometimes we need more tricks.

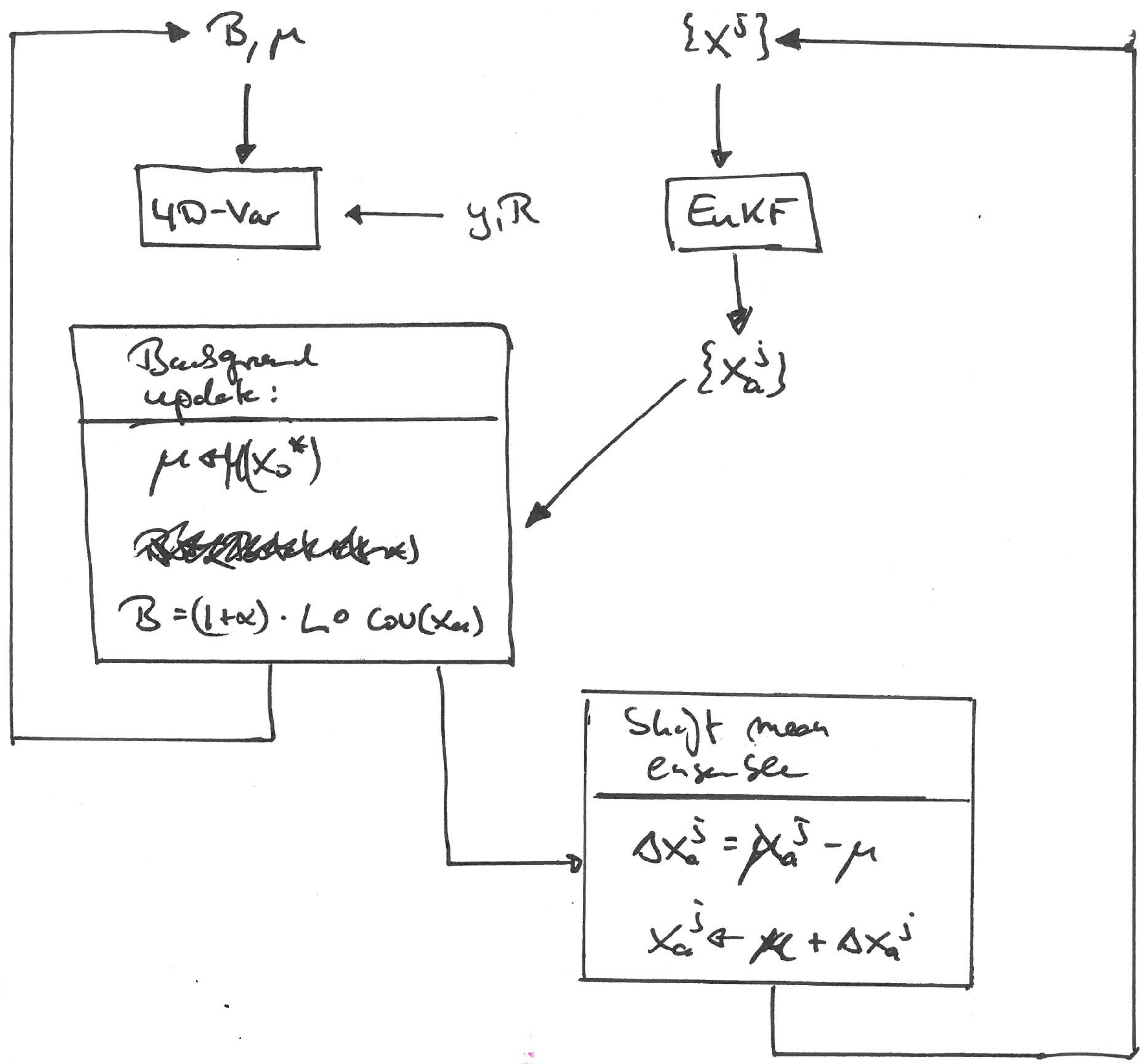
e.g. $B = \alpha B_{stat} + (1-\alpha) MH^{-1}M^T$

tune α.

Another "mix": Combine EnKF and 4D-Var = E4D-Var.

<u>Idea:</u> · use 4D-Var for state estimation

· use EnKF to update background covariances



· requires tuning → localization & inflation

· there is no guarantee that this works in nonlinear problems (where is the theory)

· Motivation: this is "ok" for ~~non~~linear problems
→ it should be "ok" for nonlinear problems.