

### Equations de l'interpolation statistique

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

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

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$$

**Matrice de gain (ou poids statistiques):**  $\mathbf{K}$   
**Innovations (ou écarts aux observation):**  $\mathbf{y} - \mathbf{H}\mathbf{x}_b$

### Interpolation optimale

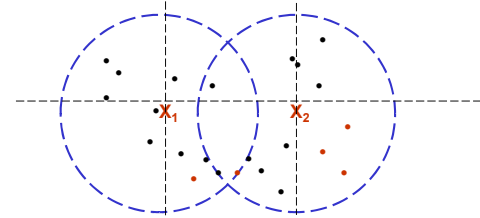
- Résolution directe du problème

$$\mathbf{x}_a(p_i) = \mathbf{x}_b(p_i) + \pi_i \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

$$= \mathbf{x}_b(p_i) + \pi_i \mathbf{B}\mathbf{H}^T \mathbf{w}$$

$$= \mathbf{x}_b(p_i) + \sum_{k=1}^K W_{ik} w_k$$

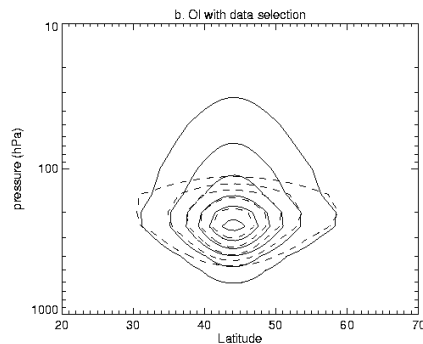
- \* Introduction de simplifications: sélection de données
- \* Conduit à des problèmes de petite dimension qui peuvent être résolus explicitement
- \* Champ analysé à différents points de grille n'utilise pas nécessairement les mêmes données



### Impact de la sélection de données

- Influence d'une observation

- \* Trait plein: 3D-Var
- \* Tireté: interpolation optimale
- \* Fig. 9 de Gauthier *et al.* (1999)



### Formulation variationnelle de l'interpolation statistique (Lewis et al., 2006: sect. 20.2 et ch. 10 à 12))

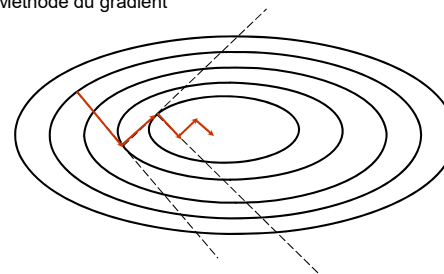
- Résolution de systèmes d'équations linéaires de grande taille

$$\mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\min J(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A}\mathbf{x} - \mathbf{x}^T \mathbf{b}$$

$$\nabla J(\mathbf{x}) = \mathbf{A}\mathbf{x} - \mathbf{b}$$

- \* Méthode du gradient



### Méthode du gradient

- En partant du point  $\mathbf{x}^{(k)}$ , trouver un point  $\mathbf{x}^{(k+1)}$  tel que  $J(\mathbf{x}^{(k+1)}) < J(\mathbf{x}^{(k)})$

$$J(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{Q} \mathbf{x} - \mathbf{x}^T \mathbf{b} + C$$

$\mathbf{Q}$  est définie positive ce qui signifie que pour tout vecteur  $\mathbf{x}$ ,  $\mathbf{x}^T \mathbf{Q} \mathbf{x} > 0$

Ceci nous assure qu'un minimum unique existe

- Calculer  $\nabla J(\mathbf{x}^{(k)})$   $\nabla J(\mathbf{x}^{(k)}) = \mathbf{Q} \mathbf{x}^{(k)} - \mathbf{b}$
- Direction de recherche:  $\mathbf{g}_k = -\nabla J(\mathbf{x}^{(k)}) = -(\mathbf{Q} \mathbf{x}^{(k)} - \mathbf{b})$
- Fouille linéaire  $f(\alpha) = J(\mathbf{x}^{(k)} + \alpha \mathbf{g}_k) = J(\mathbf{x}(\alpha))$   
 $\frac{df}{d\alpha}(\alpha) = \left[ \frac{d\mathbf{x}(\alpha)}{d\alpha} \right]^T \nabla J(\mathbf{x}^{(k)} + \alpha \mathbf{g}_k)$   
 $= \mathbf{g}_k^T [\mathbf{Q} \mathbf{x}^{(k)} - \mathbf{b} + \alpha \mathbf{Q} \mathbf{g}_k]$   
 $= \alpha \mathbf{g}_k^T \mathbf{Q} \mathbf{g}_k - \mathbf{g}_k^T \mathbf{b}$   
 $\alpha^* = \frac{\mathbf{g}_k^T \mathbf{b}}{\mathbf{g}_k^T \mathbf{Q} \mathbf{g}_k}$
- Nouvel itéré:  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^* \mathbf{g}_k$

### Le 3D-Var

- Fonction objective**

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{H} \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y})$$

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y})$$

- Au minimum de J,  $\mathbf{x} = \mathbf{x}_a$  et  $\nabla J(\mathbf{x}_a) = \mathbf{0}$ :**

$$\begin{aligned} \mathbf{B}^{-1} (\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x}_a - \mathbf{y}) &= \mathbf{B}^{-1} (\mathbf{x}_a - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} (\mathbf{x}_a - \mathbf{x}_b) - (\mathbf{y} - \mathbf{H} \mathbf{x}_b)) \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) (\mathbf{x}_a - \mathbf{x}_b) - \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}_b) \\ &= \mathbf{0} \end{aligned}$$

- Incrément d'analyse ( $\mathbf{x}_a - \mathbf{x}_b$ ):**

$$\begin{aligned} (\mathbf{x}_a - \mathbf{x}_b) &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}_b) \\ &\stackrel{?}{=} \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}_b) \end{aligned}$$

- Montrer que:**

$$(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \stackrel{?}{=} \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1}$$

### Démonstration:

$$(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1}$$

- Multiplication, de part et d'autre, à gauche par  $(\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})$  et à droite par  $(\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)$**

$$\begin{aligned} \mathbf{H}^T \mathbf{R}^{-1} &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{B} \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T)^{-1} \\ \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{R} + \mathbf{H} \mathbf{B} \mathbf{H}^T) &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{B} \mathbf{H}^T \\ \mathbf{H}^T + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T &= \mathbf{H}^T + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T \end{aligned}$$

### Matrice hessienne et covariance d'erreur d'analyse

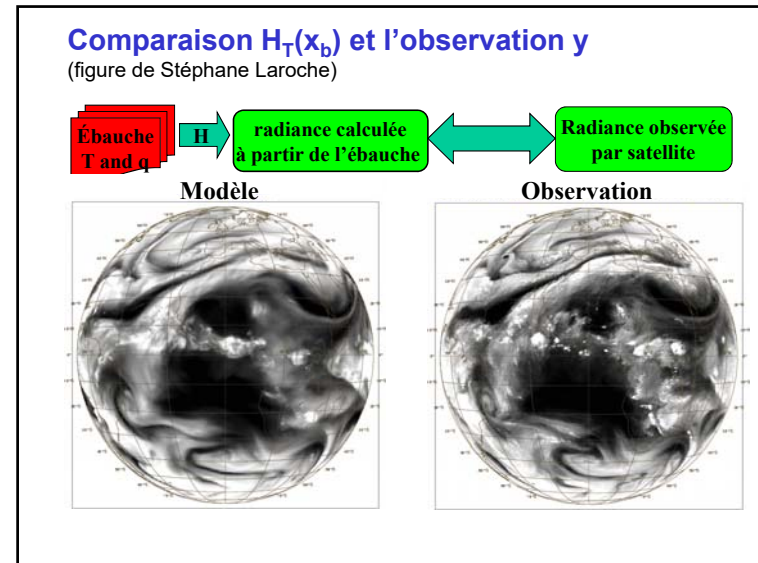
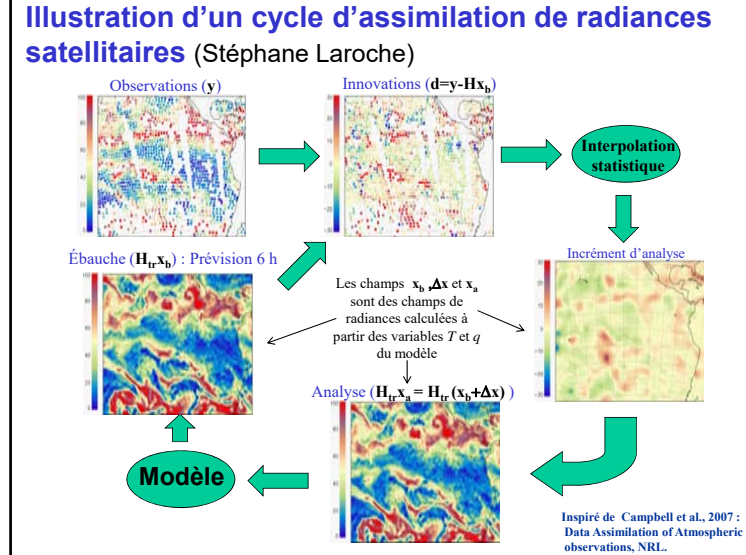
- Matrice hessienne:**  $J'' = \left( \frac{\partial^2 J}{\partial x_i \partial x_j} \right) = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{H} \mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{x} - \mathbf{y}) \\ &= \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{H} (\mathbf{x} - \mathbf{x}_b) - (\mathbf{y} - \mathbf{H} \mathbf{x}_b))^T \mathbf{R}^{-1} (\mathbf{H} (\mathbf{x} - \mathbf{x}_b) - (\mathbf{y} - \mathbf{H} \mathbf{x}_b)) \\ &= \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) (\mathbf{x} - \mathbf{x}_b) - \mathbf{d}^T \mathbf{R}^{-1} \mathbf{H} (\mathbf{x} - \mathbf{x}_b) + \mathbf{d}^T \mathbf{R}^{-1} \mathbf{d} \end{aligned}$$

**Forme incrémentale de l'analyse: on résout pour l'incrément et non pas l'analyse directement**

$$J(\Delta \mathbf{x}) = \frac{1}{2} \Delta \mathbf{x}^T (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \Delta \mathbf{x} - \mathbf{d}^T \mathbf{R}^{-1} \mathbf{H} \Delta \mathbf{x} + \frac{1}{2} \mathbf{d}^T \mathbf{R}^{-1} \mathbf{d}$$

où  $\mathbf{d} = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$  est le vecteur des innovations



### Matrice hessienne et covariance d'erreur d'analyse

- Matrice hessienne:  $J'' = \left( \frac{\partial^2 J}{\partial x_i \partial x_j} \right)_{ij} = \mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$
- Covariance d'erreur d'analyse  $\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} = \mathbf{B} - \mathbf{K}\mathbf{H}\mathbf{B}$
- Substitue  $\mathbf{K}$  par l'expression donnée précédemment

$$\begin{aligned} \mathbf{P}_a &= \mathbf{B} - \left[ (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \right] \mathbf{H}\mathbf{B} \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \left[ (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})\mathbf{B} - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{B} \right] \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \left[ \mathbf{I} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{B} - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}\mathbf{B} \right] \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \end{aligned}$$

### Exemple pour l'analyse en un point

- Analyse de température en un point sur le dernier niveau du modèle utilisant une mesure de la température à 10m

$T_b$	$\text{-----}$	Niveau du modèle
	$T_{\text{obs.}}$	Niveau à 10 m

Opérateur d'observation:  $\mathbf{H}\mathbf{T} = \alpha T$

Valeur analysée:  $T_a = T_b + \lambda(T_{\text{obs.}} - \alpha T_b)$

Estimé minimisant la variance d'erreur d'analyse:

$$\lambda = \frac{\alpha \sigma_b^2}{(\sigma_b^2 + \alpha^2 \sigma_o^2)}$$

**Formulation variationnelle: exemple pour l'analyse en un point**

- **Minimiser la fonction coût:**

$$J(T) = \frac{1}{2} \frac{(T - T_b)^2}{\sigma_b^2} + \frac{1}{2} \frac{(\alpha T - T_{obs})^2}{\sigma_o^2}$$

$$\nabla J(T) = \frac{\partial J}{\partial T} = \frac{(T - T_b)}{\sigma_b^2} + \alpha \frac{(\alpha T - T_{obs})}{\sigma_o^2}$$

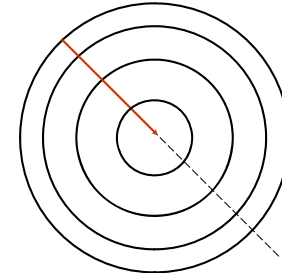
$$T_a - T_b = \frac{\alpha \sigma_b^2}{(\sigma_o^2 + \alpha^2 \sigma_b^2)} (T_{obs} - \alpha T_b)$$

**Incrément d'analyse** ( $T_a - T_b$ ): changement apporté au champ d'essai pour s'ajuster à la valeur observée

**Equivalent modèle de l'observation** ( $\alpha T_b$ ): état-modèle est converti en équivalent de l'observation.

**Convergence de la minimisation et préconditionnement**

- **Minimisation de**  
 $J(\xi) = 1/2 (\xi^T \xi)$ .
- **Peu importe le point de départ, la minimisation converge en une seule itération.**



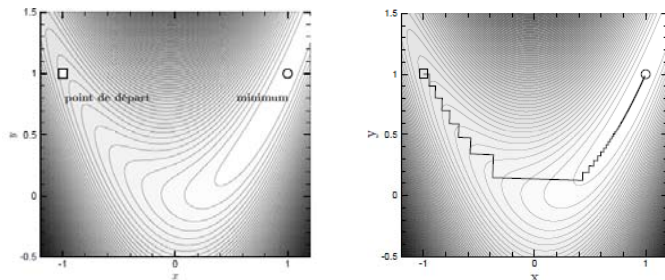
**Fonction de Rosenbrock pour mettre à l'épreuve les algorithmes d'optimisation**

La fonction de Rosenbrock banana  $\mathcal{R} : \mathbb{R}^2 \rightarrow \mathbb{R}$  à minimiser est la suivante:

$$\mathcal{R}(x, y) = (x - 1)^2 + p(x^2 - y^2)^2.$$

La profondeur de la vallée est fonction du paramètre  $p$  qui est ici choisi égal à 10. Le minimum de la fonction  $\mathcal{R}$  est  $x = y = 1$ .

**Méthode du gradient**  
**312 itérations**  
 $f_{\min} = 11 \times 10^{-5}$



**Préconditionnement du 3D-Var par changement de variable: cas où H est linéaire**

- **Formulation originale:**

$$J(\mathbf{x}) = \frac{1}{2} (\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \frac{1}{2} (\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1} (\mathbf{x} - \mathbf{x}_b) + \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

- **Changement de variables**  $\xi = \mathbf{B}^{-1/2} (\mathbf{x} - \mathbf{x}_b) = \mathbf{B}\Delta\mathbf{x}$   
 $\mathbf{x} = \mathbf{x}_b + \mathbf{B}^{1/2} \xi \equiv \mathbf{x}(\xi)$

- **Nouvelle forme:**

$$J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})$$

$$\nabla J(\xi) = \xi + (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})$$

- **Algorithmes de minimisation**

- \* Gradient conjugué (réf. Gollub et van Loan, 1996)
- \* Quasi-newton (Navon et Legler, 1987: *Mon. Wea. Rev.*, **115**,1479-1502)
- \* Lewis *et al.* (2006): ch.10 à 12

- **Minimisation comprend généralement moins d'une centaine d'itérations**

### Opérateurs élémentaires de l'analyse

• **Formulation**

$$J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})$$

$$\nabla J(\xi) = \xi + (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})$$

• **Changement de variables**

$$\xi = \mathbf{B}^{-1/2} (\mathbf{x} - \mathbf{x}_b) = \Delta \mathbf{x}$$

$$\mathbf{x} = \mathbf{x}_b + \mathbf{B}^{1/2} \xi \equiv \mathbf{x}(\xi)$$

$$\mathbf{B} = \mathbf{B}^{1/2} (\mathbf{B}^{1/2})^T$$

**Description du changement de variables**

$$\xi \xrightarrow{\mathbf{B}^{1/2}} \mathbf{x} = \mathbf{x}_b + \Delta \mathbf{x} \xrightarrow{\mathbf{H}} \mathbf{H}\mathbf{x} - \mathbf{y} \xrightarrow{\mathbf{R}^{-1}} \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

### Calcul du gradient

$$\begin{aligned} \delta J(\xi_0) &= J(\xi_0 + \delta \xi) - J(\xi_0) \\ &= \delta \xi^T \left[ \xi_0 + (\mathbf{B}^{1/2})^T \mathbf{H} \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}_0(\xi_0) - \mathbf{y}) \right] + \dots \\ &\approx \delta \xi^T \nabla J(\xi_0) \end{aligned}$$

où  $\mathbf{x}_0(\xi_0) = \mathbf{x}_b + \mathbf{B}^{1/2} \xi_0$

### Expression pour le gradient de $J_0(\xi)$

- **Evaluation de  $J_0$ :**  $J_0(\xi) = \frac{1}{2} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x}(\xi) - \mathbf{y})$

$$\xi \xrightarrow{\mathbf{B}^{1/2}} \mathbf{x} = \mathbf{x}_b + \Delta \mathbf{x} \xrightarrow{\mathbf{H}} \mathbf{H}\mathbf{x} - \mathbf{y} \xrightarrow{\mathbf{R}^{-1}} \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

- **Evaluation du gradient de  $J_0$ :**

$$\begin{aligned} \nabla_{\xi} J_0(\xi_0) &= (\partial \mathbf{x} / \partial \xi)^T \nabla_{\mathbf{x}} J_0(\mathbf{x}) = (\mathbf{B}^{1/2})^T \nabla_{\mathbf{x}} J_0(\mathbf{x}) \\ &= (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y}) \end{aligned}$$

$$\nabla_{\xi} J_0(\xi) \leftarrow \mathbf{B}^{1/2^T} \nabla_{\mathbf{x}} J_0(\mathbf{x}) \leftarrow \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y})$$

- **Fonction coût est représentée par une composition de changements de variables et le gradient peut être obtenue en appliquant la règle de dérivation en chaîne**

### Modélisation des covariances d'erreur de prévision

- **Problème de grande dimension ( $n \sim 10^7$ ) empêche de représenter *explicitement* la matrice  $\mathbf{B}$  qui d'ordre  $n \times n$ .**
  - \*  $\mathbf{B}$  comprend  $n(n+1)/2$  éléments à définir
  - \* Représentation en termes de vecteurs propres ne peut être faite
- **Modèle de covariance**
  - \* Représentation des éléments de  $\mathbf{B}$  permettant de réduire le nombre de paramètres à estimer

**Exemple: représentation explicite**

- Considérant que

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma \Rightarrow \mathbf{B}^{-1} = \Sigma^{-1} \mathbf{C}^{-1} \Sigma^{-1}$$

$$\mathbf{C} = \mathbf{V} \Lambda \mathbf{V}^T \Rightarrow \mathbf{C}^{-1} = \mathbf{V} \Lambda^{-1} \mathbf{V}^T$$

- On peut donc écrire que

$$\mathbf{B} = \Sigma \mathbf{V} \Lambda \mathbf{V}^T \Sigma \quad \mathbf{B}^{-1} = \Sigma^{-1} \mathbf{V} \Lambda^{-1} \mathbf{V}^T \Sigma^{-1}$$

$$= (\Sigma \mathbf{V} \Lambda^{1/2}) (\Sigma \mathbf{V} \Lambda^{1/2})^T = (\Sigma^{-1} \mathbf{V} \Lambda^{-1/2}) (\Lambda^{-1} \mathbf{V}^T \Sigma^{-1/2})$$

$$= \mathbf{B}^{1/2} (\mathbf{B}^{1/2})^T = (\mathbf{B}^{-1/2})^T \mathbf{B}^{-1/2}$$

- On note qu'en factorisant  $\Sigma^{-1}$ , les corrélations  $\mathbf{C}$  deviennent *adimensionnelles* (sans unités)

**Succession d'opérations pour le calcul de  $\mathbf{B}^{1/2}$**

- On a que:  $\mathbf{B}^{1/2} \xi = \Sigma \mathbf{V} \Lambda^{1/2} \xi$

Evaluation de  $\mathbf{J}_0$ :  $J_0(\xi) = +\frac{1}{2} (\mathbf{H}\mathbf{X}(\xi) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{X}(\xi) - \mathbf{y})$

$$\xi \xrightarrow{\Lambda^{1/2}} \tilde{\xi} = \Lambda^{1/2} \xi \xrightarrow{\mathbf{V}} \tilde{\mathbf{X}} = \mathbf{V} \Lambda^{1/2} \xi \xrightarrow{\Sigma} \Delta \mathbf{X} = \Sigma \mathbf{V} \Lambda^{1/2} \xi = \mathbf{B}^{1/2} \xi \xrightarrow{\mathbf{H}} \mathbf{X}_b + \Delta \mathbf{X} \xrightarrow{\mathbf{H}} \mathbf{H}\mathbf{X} - \mathbf{y} \xrightarrow{\mathbf{R}^{-1}} \mathbf{R}^{-1} (\mathbf{H}\mathbf{X} - \mathbf{y})$$

Evaluation du gradient de  $\mathbf{J}_0$ :

$$\mathbf{g}_0 = \nabla_{\mathbf{X}} J_0(\mathbf{X}) \leftarrow \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{X} - \mathbf{y})$$

$$\nabla_{\xi} J_0(\xi) = \Lambda^{1/2} \mathbf{V}^T \Sigma \mathbf{g}_0 \leftarrow \Lambda^{1/2} \tilde{\xi} = \mathbf{V}^T \Sigma \mathbf{g}_0 \leftarrow \mathbf{V}^T \tilde{\mathbf{X}} = \Sigma \mathbf{g}_0 \leftarrow \Sigma \mathbf{g}_0$$

**Résumé des opérations permettant de représenter  $\mathbf{B}^{1/2} \xi = \Sigma \mathbf{V} \Lambda^{1/2} \xi$**

- Représentation des covariances d'erreur de prévision comme *composition* d'une suite d'opérateurs
- Résultat est de ramener  $\mathbf{B}$  à la matrice identité
  - \* Signifie que les composantes de la variable  $\xi$  sont statistiquement indépendantes et d'égale variance
- Multiplication par  $\Lambda^{1/2}$  rétablit l'importance relative de chaque composante associée aux directions des vecteurs propres
- Multiplication par  $\mathbf{V}$  nous ramène aux variables *physiques* (e.g., T, u, v) normalisées par les écart-types de l'erreur de prévision
- Multiplication par  $\Sigma$  rétablit les unités des variables et la modulation spatiale de l'erreur de prévision

**Modèle de covariances d'erreur de prévision basé sur des corrélations homogènes et isotropes**

Représentation des covariances

$$\mathbf{B}(x_1, x_2) = \langle \varepsilon(x_1) \varepsilon(x_2) \rangle = \langle \varepsilon \varepsilon^T \rangle$$

où  $\varepsilon(x)$  est l'erreur de prévision au point  $x$

Propriétés

- \* Symétrique  $\mathbf{B}(x_2, x_1) = \mathbf{B}(x_1, x_2)$
- \* Définie positive  $\mathbf{x}^T \mathbf{B} \mathbf{x} = \langle (\mathbf{x}^T \varepsilon) (\varepsilon^T \mathbf{x}) \rangle = \langle (\mathbf{x}^T \varepsilon)^2 \rangle > 0$
- \* Variances  $\mathbf{B}(x, x) = \sigma^2(x)$

Corrélations

$$\mathbf{C}(x_1, x_2) = \frac{\mathbf{B}(x_1, x_2)}{\sigma(x_1) \sigma(x_2)}$$

$$\mathbf{C}(x, x) = 1$$

Forme matricielle des covariances

$$\mathbf{B} = \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n \end{pmatrix} \begin{pmatrix} 1 & C_{12} & \dots & C_{1n} \\ C_{21} & 1 & \dots & \vdots \\ \vdots & \vdots & \ddots & C_{(n-1)n} \\ C_{n1} & \dots & C_{n(n-1)} & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_n \end{pmatrix}$$

### Corrélations homogènes et isotropes

- **Forme matricielle des covariances:**  $\mathbf{B} = \Sigma \mathbf{C} \Sigma$   
 où  $\Sigma$  est la matrice diagonale contenant les écarts-type.
- **Modèles de corrélations homogène et isotropes**  
 $C(x_1, x_2) = f(|x_1 - x_2|)$   
Isotrope: fonction ne dépend que de la distance entre deux points  
Homogène: cette dépendance est la même en tout point (la fonction de corrélation ne change pas)
- **Corrélations homogènes et isotropes n'impliquent pas que les covariances le soient puisque les écarts-type peuvent varier d'un point à un autre**
- **Forme spectrale des corrélations homogènes et isotropes est une matrice diagonale**

### Représentation spectrale des corrélations

- **Séries de Fourier unidimensionnelle**  

$$\tilde{\varepsilon}_m = \frac{1}{L} \int_0^L \varepsilon(x) e^{-imKx} dx = S \varepsilon \quad \varepsilon(x) = \sum_{m=-\infty}^{m=+\infty} \tilde{\varepsilon}_m e^{imKx} = S^{-1} \tilde{\varepsilon}$$
  - \* où  $K = 2\pi/L$
  - \* Propriété:  $\tilde{\varepsilon}_{-m} = \tilde{\varepsilon}_m^*$
- **Forme spectrale d'une matrice**  

$$\begin{aligned} \tilde{\mathbf{C}}(m, n) &= \langle \tilde{\varepsilon}(m) \tilde{\varepsilon}(-n) \rangle = \langle \tilde{\varepsilon} \tilde{\varepsilon}^T \rangle \\ &= \langle (S \varepsilon) (S \varepsilon)^T \rangle = \langle S (\varepsilon \varepsilon^T) S^T \rangle = S \langle \varepsilon \varepsilon^T \rangle S^T \\ &= S \mathbf{C} S^T = S \mathbf{C} S^{-1} \end{aligned}$$
- **Relation inverse:**  

$$\mathbf{C} = S^{-1} \tilde{\mathbf{C}} (S^{-1})^T \simeq S^{-1} \tilde{\mathbf{C}} S$$

- **Forme spectrale:**

$$\begin{aligned} \tilde{\mathbf{C}}(m, n) &= \frac{1}{L^2} \left\langle \left( \int_0^L \varepsilon(x_1) e^{-imKx_1} dx_1 \right) \left( \int_0^L \varepsilon(x_2) e^{+inKx_2} dx_2 \right) \right\rangle \\ &= \frac{1}{L^2} \int_0^L dx_1 e^{-imKx_1} \int_0^L \langle \varepsilon(x_1) \varepsilon(x_2) \rangle e^{+inKx_2} dx_2 \\ &= \frac{1}{L^2} \int_0^L dx_1 e^{-imKx_1} \int_0^L f(|x_1 - x_2|) e^{+inKx_2} dx_2 \end{aligned}$$

- **Changement de variable:**  $\xi = x_1 - x_2$   

$$\begin{aligned} \tilde{\mathbf{C}}(m, n) &= \frac{1}{L^2} \int_0^L dx_1 e^{-imKx_1} \int_0^L f(\xi) e^{+inK(x_1 - \xi)} d\xi \\ &= \frac{1}{L^2} \int_0^L dx_1 e^{-i(m-n)Kx_1} \int_0^L f(\xi) e^{-inK\xi} d\xi \\ &= \tilde{f}_n \frac{1}{L} \int_0^L dx_1 e^{-i(m-n)Kx_1} = \tilde{f}_n \delta(m - n) \end{aligned}$$

- **La forme spectrale de la matrice des corrélations conduit à une matrice diagonale dont les éléments sont obtenus des coefficients spectraux de la fonction de corrélation  $f(\xi)$**

### Remarques

- **Cas 2D sur le plan**
  - \* Double transformée de Fourier conduit également à une matrice de corrélation diagonale.
  - \* Passage en coordonnées polaires:  $(x, y) \rightarrow (r, \theta)$   
 → Transformée de Hankel unidimensionnelle (c.f. Daley (1991), ch.4, sect. 4.3 et Annexe G) développement en termes de fonctions de Bessel
- **Cas 2D sur la sphère**
  - \* Représentation spectrale est basée sur les harmoniques sphériques (zonales)
  - \* Fonctions de corrélations sont représentées en termes des composantes zonales seulement (Boer (1983); Gauthier *et al.* (1993))
  - \* Gauthier *et al.* (1998): modèle de corrélation homogène et isotrope sur la sphère (utilisé dans le 3D-Var du SMC)
  - \* Modèle de covariance dont la représentation est compacte (e.g., 2N coefficients au lieu de  $N(N+1)/2$ )

### Extension au cas où H(X) est une fonction non linéaire

• **Formulation:**  $J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$

$$\nabla J(\xi) = \xi + (\mathbf{B}^{1/2})^T \left( \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \right)^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$$

• **Changement de variables**  $\xi = \mathbf{B}^{-1/2} (\mathbf{x} - \mathbf{x}_b)$

$$\mathbf{x} = \mathbf{x}_b + \mathbf{B}^{1/2} \xi \equiv \mathbf{x}(\xi)$$

• **Décomposition de B<sup>1/2</sup>:**

$$\begin{aligned} \mathbf{B} &= \Sigma \mathbf{S}^{-1} \tilde{\mathbf{C}} (\mathbf{S}^{-1})^T \Sigma = \Sigma \mathbf{S}^{-1} \tilde{\mathbf{C}}^{1/2} (\Sigma \mathbf{S}^{-1} \tilde{\mathbf{C}}^{1/2})^T \\ &= \mathbf{B}^{1/2} (\mathbf{B}^{1/2})^T \\ \mathbf{B}^{1/2} &= \Sigma \mathbf{S}^{-1} \tilde{\mathbf{C}}^{1/2} \end{aligned}$$

#### Description du changement de variables

$$\xi \xrightarrow{\tilde{\mathbf{C}}^{1/2}} \xi_1 \xrightarrow{\mathbf{S}^{-1}} \mathbf{x}_1 \xrightarrow{\Sigma} \mathbf{x} = \mathbf{x}_b + \Delta \mathbf{x} \xrightarrow{\mathbf{H}} \mathbf{H}(\mathbf{x}) - \mathbf{y} \xrightarrow{\mathbf{R}^{-1}} \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}) - \mathbf{y})$$

### Calcul du gradient de J<sub>b</sub> et J<sub>o</sub>

$$\begin{aligned} \delta J(\xi_0) &= J(\xi_0 + \delta \xi) - J(\xi_0) \\ &= \delta \xi^T \left[ \xi_0 + (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}_0) - \mathbf{y}) \right] \\ &= \delta \xi^T \nabla J(\xi_0) \end{aligned}$$

Variation de H(x) causée par une variation  $\delta \mathbf{x} = \mathbf{B}^{1/2} \delta \xi$  de l'état-modèle

$$\mathbf{H}(\mathbf{x}_0 + \delta \mathbf{x}) - \mathbf{H}(\mathbf{x}_0) \approx \left[ \frac{\partial \mathbf{H}}{\partial \mathbf{x}} (\mathbf{x}_0) \right] \cdot \delta \mathbf{x} \equiv \mathbf{H}'(\mathbf{x}_0) \cdot \delta \mathbf{x}$$

**H'(X<sub>0</sub>) est la matrice jacobienne de H évaluée en x<sub>0</sub> = x<sub>b</sub> + B<sup>1/2</sup>ξ<sub>0</sub>.**

Règle de dérivation en chaîne:

$$\begin{aligned} \frac{\partial \mathbf{H}(\mathbf{x}(\xi_0))}{\partial \xi} &= \left( \frac{\partial \mathbf{x}}{\partial \xi} \right)^T \left( \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \right)_{\mathbf{x}=\mathbf{x}_0(\xi)} \\ &= \mathbf{B}^{1/2 T} \left( \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \right)_{\mathbf{x}=\mathbf{x}_0(\xi)} \end{aligned}$$

### Exemple: assimilation d'une mesure d'intensité de vent

• **Intensité de vent V et composantes horizontales du vent u et v:**

$$V = \sqrt{u^2 + v^2}$$

• **Perturbation de V causée par un changement δu et δv:**

$$\begin{aligned} \delta V &= \frac{\partial V}{\partial u} (u_0, v_0) \delta u + \frac{\partial V}{\partial v} (u_0, v_0) \delta v \\ &= \left( \frac{\partial V}{\partial u} (u_0, v_0) \quad \frac{\partial V}{\partial v} (u_0, v_0) \right) \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} \\ &\equiv \mathbf{H}'(u_0, v_0) \begin{pmatrix} \delta u \\ \delta v \end{pmatrix} = \mathbf{H}'(\mathbf{X}_0) \delta \mathbf{X}_0 \end{aligned}$$

• **Gradient de J<sub>o</sub> par rapport à X:**

$$\begin{aligned} \nabla_{\mathbf{x}} J_0(\mathbf{X}_0) &= \mathbf{H}'(\mathbf{X}_0)^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{X}_0) - \mathbf{y}) \\ &= \begin{pmatrix} u_0 / \sqrt{u_0^2 + v_0^2} \\ v_0 / \sqrt{u_0^2 + v_0^2} \end{pmatrix} \left( \sqrt{u_0^2 + v_0^2} - V_{obs.} \right) \frac{1}{\sigma_o^2} \end{aligned}$$

∴ **Aucune correction sur la direction**

### Expression pour le gradient de J<sub>o</sub>(ξ)

• **Evaluation de J<sub>o</sub>:**  $J_o(\xi) = \frac{1}{2} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$

$$\xi \xrightarrow{\tilde{\mathbf{C}}^{1/2}} \xi_1 \xrightarrow{\mathbf{S}^{-1}} \mathbf{x}_1 \xrightarrow{\Sigma} \mathbf{x} = \mathbf{x}_b + \Delta \mathbf{x} \xrightarrow{\mathbf{H}} \mathbf{H}(\mathbf{x}) - \mathbf{y} \xrightarrow{\mathbf{R}^{-1}} \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}) - \mathbf{y})$$

• **Evaluation du gradient de J<sub>o</sub>:**

$$\begin{aligned} \nabla_{\xi} J_0(\xi_0) &= (\partial \mathbf{x} / \partial \xi)^T \nabla_{\mathbf{x}} J_0(\mathbf{x}) = (\mathbf{B}^{1/2})^T \nabla_{\mathbf{x}} J_0(\mathbf{x}) \\ &= (\mathbf{B}^{1/2})^T \left( \frac{\partial \mathbf{H}}{\partial \mathbf{x}} \right)^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}) - \mathbf{y}) \end{aligned}$$

$$\nabla_{\xi} J_0(\xi) \xleftarrow{\tilde{\mathbf{C}}^{1/2}} \nabla_{\xi_1} J_0(\xi_1) \xleftarrow{\mathbf{S}} \nabla_{\mathbf{x}_1} J_0(\mathbf{x}_1) \xleftarrow{\Sigma} \nabla_{\mathbf{x}} J_0(\mathbf{x}) \xleftarrow{(\partial \mathbf{H} / \partial \mathbf{x})^T} \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}) - \mathbf{y})$$

• **Fonction coût est représentée par une composition de changements de variables et le gradient peut être obtenue en appliquant la règle de dérivation en chaîne**



### Classification qualitative des non linéarités

- **Problème inverse linéaire**
  - \* Opérateur d'observation est de la forme  $\mathbf{y} = \mathbf{H}\mathbf{x}$
  - \* L'estimé *a priori* a une distribution Gaussienne et les équations de l'analyse sont alors purement linéaires
  - \* Peu de problèmes réels d'inversion sont dans cette catégorie
- **Problèmes quasi-linéaires: approche incrémentale**
  - \* Problèmes non linéaires mais qui peuvent être résolus à l'aide d'une linéarisation au voisinage d'un état raisonnable pour une plage de variations typiques pour une variable (e.g., température, humidité)
- **Problèmes modérément non linéaires**
  - \* Linéarisation peut être adaptée pour une estimation de l'erreur d'analyse mais pas pour résoudre le problème directement
  - \* Approche par linéarisations successives doit être adoptée
    - Approche incrémentale avec boucle externe
  - \* Plusieurs problèmes d'assimilation sont dans cette catégorie
- **Problèmes hautement non linéaires**
  - \* Situations pour lesquelles des variations de l'ordre des erreurs de prévision et d'observation nécessitent la prise en compte des non linéarités
  - \* Situations avec processus de seuil comme le déclenchement de la précipitation, la présence d'un nuage ou non.

### Approche incrémentale

- **Formulation complète originale:**

$$J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$$

$$\nabla J(\xi) = \xi + (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$$

- **Changement de variables**  $\xi = \mathbf{B}^{-1/2} (\mathbf{x} - \mathbf{x}_b) = \mathbf{B}^{-1/2} \delta \mathbf{x}$

$$\mathbf{x} = \mathbf{x}_b + \mathbf{B}^{1/2} \xi \equiv \mathbf{x}(\xi)$$

- **Comme la variable  $\xi$  est reliée à l'incrément d'analyse  $\delta \mathbf{x}$**

$$\begin{aligned} \mathbf{H}(\mathbf{x}(\xi)) &= \mathbf{H}(\mathbf{x}_b + \delta \mathbf{x}) \cong \mathbf{H}(\mathbf{x}_b) + \frac{\partial \mathbf{H}}{\partial \mathbf{x}}(\mathbf{x}_b) \cdot \delta \mathbf{x} + \dots \\ &\equiv \mathbf{H}(\mathbf{x}_b) + \mathbf{H}'(\mathbf{x}_b) \cdot \delta \mathbf{x} \end{aligned}$$

### Approche incrémentale (suite)

- **Nouvelle forme de la fonction objective devient**

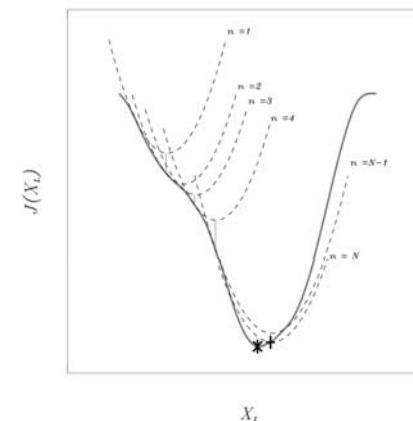
$$J_L(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}' \mathbf{B}^{1/2} \xi - \mathbf{y}')^T \mathbf{R}^{-1} (\mathbf{H}' \mathbf{B}^{1/2} \xi - \mathbf{y}')$$

$$\nabla J_L(\xi) = \xi + \mathbf{B}^{1/2T} \mathbf{H}'^T \mathbf{R}^{-1} (\mathbf{H}' \mathbf{B}^{1/2} \xi - \mathbf{y}')$$

où  $\mathbf{y}' = \mathbf{y} - \mathbf{H}(\mathbf{x}_b)$

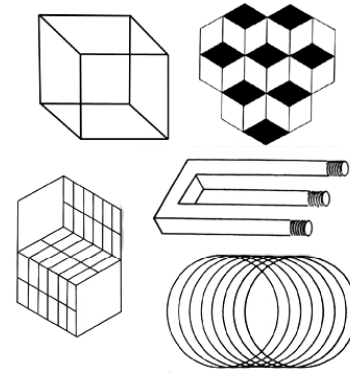
- **Forme est similaire au problème original sauf que**
  - \* Les observations sont remplacées par les innovations
  - \* L'opérateur d'observation est remplacé par son approximation linéarisée autour de l'état courant défini par  $\mathbf{x}_b$
  - \* Approximations peuvent être justifiées pour les incréments mais pas nécessairement pour un état modèle complet

### Relation entre fonction de coût et les approximations quadratiques successives (Laroche et Gauthier, 1997)



## Approche bayésienne à l'estimation statistique

## Autre perspective au problème de l'estimation statistique ...



## Approche bayésienne aux problèmes inverses

**Fonction de densité de probabilité conjointe (pdf, *probability density function*):  $p(\mathbf{x}, \mathbf{y})$**

\* Densités de probabilité marginales associées

$$P(\mathbf{x}) = \int p(\mathbf{x}, \mathbf{y}) d\mathbf{y}, \quad P(\mathbf{y}) = \int p(\mathbf{x}, \mathbf{y}) d\mathbf{x}$$

**A priori pdf  $P(\mathbf{x})$ : probabilité que  $\mathbf{x} = \mathbf{x}_b$**

\* Exemple: cas Gaussien pour lequel on connaîtrait les covariances d'erreur et que nous ayons  $\mathbf{x}_b$  comme seule réalisation de  $\mathbf{x}$ .

\* La variable  $\mathbf{x}$  est distribuée normalement de moyenne  $\mathbf{x}_b$  et covariance  $\mathbf{B}$

$$P(\mathbf{x}) = \frac{1}{C} \exp\left\{-\frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b)\right\}$$

**En absence d'autre information,  $\mathbf{x} = \mathbf{x}_b$  est l'état le plus probable (maximum likelihood)**

**De même, pour  $P(\mathbf{y})$ , si  $\mathbf{y}_o$  correspond aux observations disponibles,**

$P(\mathbf{y})$ : probabilité que  $\mathbf{y} = \mathbf{y}_i$

Estimé de la moyenne:  $\mathbf{y} = \mathbf{y}_o$

Cas gaussien:

$$P(\mathbf{y}) = \frac{1}{C_2} \exp\left\{-\frac{1}{2}(\mathbf{y} - \mathbf{y}_o)^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{y}_o)\right\}$$

\* Variable aléatoire distribuée normalement, de moyenne  $\mathbf{y}_o$  et de covariance  $\mathbf{R}$

### Théorème de Bayes

- Distribution de la probabilité conditionnelle**

- \* Probabilité de  $\mathbf{y}$  étant donné une certaine valeur de  $\mathbf{x}$

$$p(\mathbf{y} | \mathbf{x} = \mathbf{x}_o) = \frac{p(\mathbf{x}_o, \mathbf{y})}{\int p(\mathbf{x}_o, \mathbf{y}) d\mathbf{y}} \equiv \frac{p(\mathbf{x}_o, \mathbf{y})}{P(\mathbf{x}_o)}$$

- \* Probabilité de  $\mathbf{x}$  étant donné une certaine valeur de  $\mathbf{y}$

$$p(\mathbf{x} | \mathbf{y} = \mathbf{y}_o) = \frac{p(\mathbf{x}, \mathbf{y}_o)}{\int p(\mathbf{x}, \mathbf{y}_o) d\mathbf{x}} \equiv \frac{p(\mathbf{x}, \mathbf{y}_o)}{P(\mathbf{y})}$$

- \* Conséquemment,

$$p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x} | \mathbf{y})P(\mathbf{y}) = p(\mathbf{y} | \mathbf{x})P(\mathbf{x})$$

- Théorème de Bayes:**

$$p(\mathbf{x} | \mathbf{y}) = \frac{p(\mathbf{y} | \mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}$$

### L'état le plus probable (ou mode)

- $p(\mathbf{x} | \mathbf{y})$ :** probabilité que  $\mathbf{x} = \mathbf{x}_t$  étant donné que  $\mathbf{y}$  a été observé

- \* Densité de probabilité *a posteriori* associée à celle de l'erreur d'analyse

- $p(\mathbf{y} | \mathbf{x})$ :** probabilité de  $\mathbf{y}$  étant donné que  $\mathbf{x}$  est la valeur exacte

- \* **Remarque:**  $\mathbf{H}(\mathbf{x})$ : estimé de la valeur moyenne  $\mathbf{y}$ .

→ Si  $\mathbf{x} = \mathbf{x}_t$ , alors  $\mathbf{H}\mathbf{x}_t = \mathbf{y}_t + \varepsilon_R$ .

- \*  $(\mathbf{y} - \mathbf{H}(\mathbf{x})) = \mathbf{y}_t + \varepsilon_m - \varepsilon_R - \mathbf{y}_t = \varepsilon_m - \varepsilon_R$

- \*  $\mathbf{y}$  est distribuée normalement, de moyenne  $\mathbf{H}(\mathbf{x})$  et de covariance  $\mathbf{R}$  qui inclut l'erreur de mesure et l'erreur de représentativité

$$\begin{aligned} \mathbf{R} &= \langle (\mathbf{y} - \langle \mathbf{y} \rangle)(\mathbf{y} - \langle \mathbf{y} \rangle)^T \rangle = \langle (\varepsilon_m - \varepsilon_R)(\varepsilon_m - \varepsilon_R)^T \rangle \\ &= \langle \varepsilon_m \varepsilon_m^T \rangle + \langle \varepsilon_R \varepsilon_R^T \rangle \end{aligned}$$

### Distribution conditionnelle de $\mathbf{y}$

- On conclut alors que**

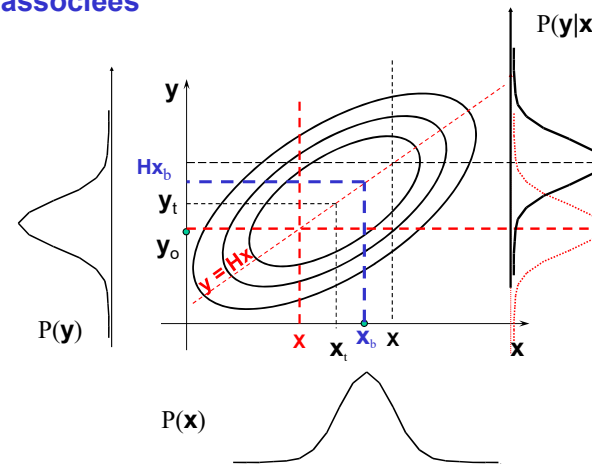
$$p(\mathbf{y} | \mathbf{x}) = \frac{1}{C_3} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}(\mathbf{x})) \right\}$$

- Remarque**

- \* Lorsque l'erreur d'observation n'est pas corrélée pour les différents types d'instruments

$$\mathbf{R}^{-1} = \begin{pmatrix} \mathbf{R}_1^{-1} & 0 & 0 \\ 0 & \mathbf{R}_2^{-1} & 0 \\ 0 & 0 & \mathbf{R}_3^{-1} \end{pmatrix}$$

### Représentation des distributions de probabilité associées



**Mode:**  $\frac{d}{dx}(-\ln p) = -\frac{1}{p} \frac{dp}{dx} = 0 \Rightarrow \frac{dp}{dx} = 0$

**De par le théorème de Bayes, la probabilité a posteriori est:**

$$p(\mathbf{x}|\mathbf{y}) = C \frac{\exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{H}(\mathbf{x}))^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}(\mathbf{x}))\right\} \exp\left\{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)\right\}}{\exp\left\{-\frac{1}{2}(\mathbf{y}-\mathbf{y}_o)^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{y}_o)\right\}}$$

$$-\ln p(\mathbf{x}|\mathbf{y}) = J(\mathbf{x})$$

$$= \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b)^T + \frac{1}{2}(\mathbf{H}(\mathbf{x})-\mathbf{y})^T \mathbf{R}^{-1}(\mathbf{H}(\mathbf{x})-\mathbf{y}) + C$$

- Pour le cas où les statistiques sont gaussiennes, le mode et l'état minimisant la variance totale coïncident.
- Formulation inclut le cas où  $\mathbf{H}(\mathbf{x})$  est nonlinéaire.

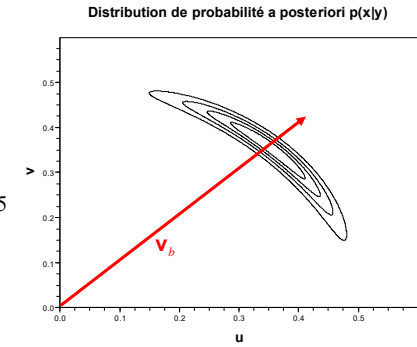
**Distribution a posteriori non-gaussienne pour  $p(\mathbf{x}|\mathbf{y})$  associée à l'assimilation d'une mesure d'intensité de vent**

Opérateur d'observation

$$\mathbf{x} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \mathbf{H}\mathbf{x} = \sqrt{u^2 + v^2}$$

Paramètres:

$$u_b = v_b = 0.4 \quad V_{obs} = 0.5 \\ \sigma_b = 0.2, \quad \sigma_o = 0.02$$



$$p(\mathbf{x}|\mathbf{y}) = \exp\left(-\frac{(u-u_b)^2 + (v-v_b)^2}{\sigma_b^2}\right) \exp\left(-\frac{\left(\left(\sqrt{u^2 + v^2}\right) - V_{obs}\right)^2}{\sigma_o^2}\right) = P(u, v)$$

**Références**

- Rodgers, C.D., 2000: sect. 2.3
  - Tarantola, A., 2005: *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadelphie, 342 pages. (ch.1)
- (disponible sur le Web à l'adresse <http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/Books/index.html>)

**Références**

- Rodgers, C.D., 2000: sect. 2.3
  - Tarantola, A., 2005: *Inverse problem theory and methods for model parameter estimation*. SIAM, Philadelphie, 342 pages. (ch.1)
- (disponible sur le Web à l'adresse <http://www.ipgp.jussieu.fr/~tarantola/Files/Professional/Books/index.html>)

## Résumé

- Estimateur de variance minimale est

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

- Méthode variationnelle ramène le problème à la minimisation d'une fonction objective

$$J(\xi) = \frac{1}{2} \xi^T \xi + \frac{1}{2} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$$

$$\nabla J(\xi) = \xi + (\mathbf{B}^{1/2})^T \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H}(\mathbf{x}(\xi)) - \mathbf{y})$$

- Modélisation des covariances d'erreur basée sur des corrélations homogènes et isotropes
- Forme variationnelle 3D décrit le problème comme une succession de changements de variables dont la transposée (adjoint) permet d'obtenir le gradient de la fonction objective

## Résumé

- L'opérateur d'observation  $H$  peut être non linéaire

- La matrice de gain  $\mathbf{K}$  a été obtenue sur la base d'une hypothèse linéaire pour  $\mathbf{H}$
- Approche bayésienne a permis d'obtenir *directement* la forme variationnelle en levant l'hypothèse de linéarité et en permettant de prendre en compte des distributions de probabilité autres que gaussiennes

- Comment peut-on obtenir les statistiques d'erreur d'observation ( $\mathbf{R}$ ) et de prévision ( $\mathbf{B}$ )?

- Estimation et modélisation des statistiques d'erreur de prévision
- Contrôle de qualité des observations