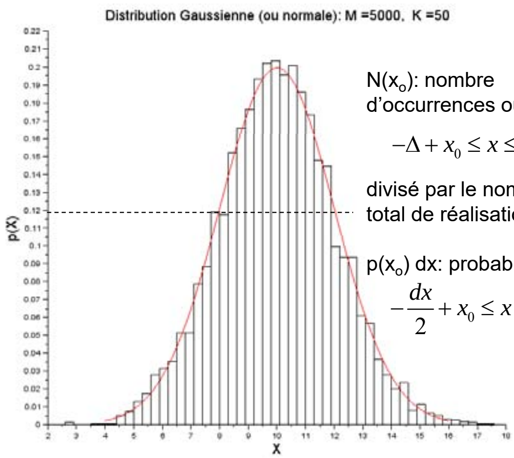


N(x_o): nombre d'occurrences où

$$-\Delta + x_0 \le x \le x_0 + \Delta$$

divisé par le nombre total de réalisations (M).

p(x_o) dx: probabilité que

$$-\frac{dx}{2} + x_0 \le x \le x_0 + \frac{dx}{2}$$

SCA-7216 Introduction à l'assimilation de données

Notions statistiques de base

- **<u>Variable aléatoire</u> ϕ est caractérisée par sa distribution (ou *densité*) de probabilité p(ϕ)**

→ Valeur moyenne: $\langle \phi \rangle = E(\phi) = \int_{-\infty}^{+\infty} \phi\, p(\phi)\, d\phi$

→ Variance:

$$\left\langle (\phi - \langle\phi\rangle)^2 \right\rangle = \int_{-\infty}^{+\infty} (\phi - \langle\phi\rangle)^2\, p(\phi) d\phi$$

$$= \langle \phi^2 \rangle - \langle \phi \rangle^2$$

→ Général: $\langle f(\phi) \rangle = \int_{-\infty}^{+\infty} f(\phi)\, p(\phi) d\phi$

- **Propriété:** si D représente l'ensemble des valeurs possibles (ou admissibles) de ϕ, alors

$$\int_D p(\phi)\, d\phi = \int_{-\infty}^{+\infty} p(\phi)\, d\phi = 1$$

- **Médiane μ:** $\int_{-\infty}^{\mu} p(x)dx = \int_{\mu}^{+\infty} p(x)dx$

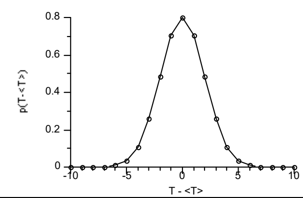
- **Mode statistique:** $\frac{dp}{dx}(m) = 0$

Distribution gaussienne

- **Définition**

$$p(T) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(T-\bar{T})^2/2\sigma^2}$$

$$\int_{-\infty}^{+\infty} p(T)dT = 1$$



- **Propriétés (exercice)**

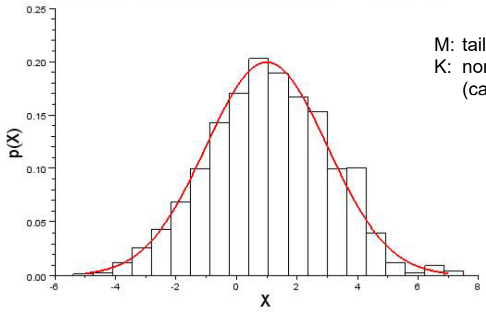
$$\left\langle (T - \langle T \rangle)^n \right\rangle = \begin{cases} 0 & \text{si n } est\ impair \\ 1\cdot 3\cdot 5 \ldots (2m-1)\sigma^{2m} & \text{si n} = 2m \end{cases}$$

- **Exercice: montrer que la probabilité que** $\langle$**T**$\rangle - \sigma <$ **T** $< \langle$**T**$\rangle + \sigma$ est de 68.26%

- **Moyenne, médiane et mode coïncident**

Estimation d'une distribution de probabilité

Distribution Gaussienne (ou normale): M = 1000, K =20



M: taille de l'échantillon
K: nombre de classes (catégories)

Distribution log-normale

- **Variable *q* définie positive: humidité, concentrations, etc.**
- **Définit:** $\xi = \ln q/\gamma$ **où** γ **représente les unités de** q
- **Distribution gaussienne exprimée en termes de** ξ**:**

$$p(\xi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\xi-\xi_0)^2}{2\sigma^2}\right)$$
$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(q/q_0))^2}{2\sigma^2}\right) = f(q)$$

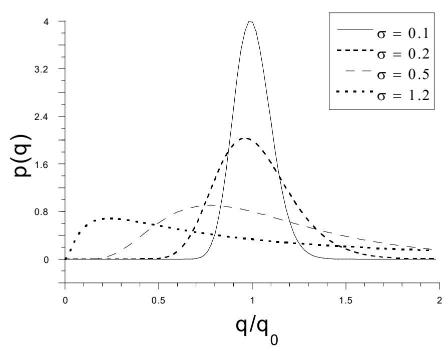
$$\int_{-\infty}^{+\infty} p(\xi)d\xi = \int_0^{+\infty} f(q)\frac{dq}{q} = \int_0^{+\infty} p(q)dq$$

- **Distribution log-normale est alors:**

$$p(q) = \frac{f(q)}{q} = \frac{1}{\sigma\sqrt{2\pi}}\frac{1}{q}\exp\left(-\frac{(\ln(q/q_0))^2}{2\sigma^2}\right)$$

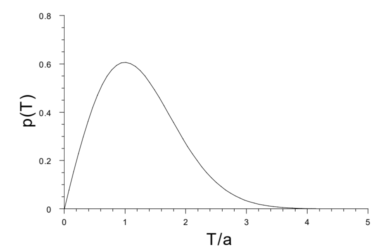
Mode $\;m: \; q_0 e^{-\sigma^2}$
Médiane $: \; q_0$
Moyenne $: \; q_0 e^{\sigma^2/2}$

Distribution log-normale



Exemple: distribution de Rayleigh

$$p(T) = \begin{cases} \dfrac{T}{a^2} e^{-T^2/2a^2} & x \geq 0 \\ 0 & x < 0 \end{cases}$$



- **Moyenne:** $\langle T \rangle = a\dfrac{\sqrt{2\pi}}{2}$
- **Médiane:** $\mu = \sqrt{2\ln 2}\; a$
- **Mode:** **m = a**

Distribution conjointe pour N variables aléatoires

- **Considère n variables aléatoires $\phi_1, \ldots, \phi_n$**

  Distribution de probabilité conjointe est $p(\phi_1, \ldots, \phi_n)$ est telle que

$$\int_V p(\phi_1, \cdots, \phi_n) d\phi_1 \cdots d\phi_n = 1$$

- **Distribution marginale de probabilité de ϕ_k**

  probabilité que ϕ_k prenne une valeur sans égard aux autres variables

$$p_k(\phi_k) = \int_{V \setminus \{\phi_k\}} p(\phi_1, \cdots, \phi_n) d\phi_1 \cdots \overline{d\phi_k} \cdots d\phi_n$$

  où l'intégration se fait sur toutes les variables sauf ϕ_k

- **Variables sont statistiquement indépendantes si**

$$p(\phi_1, \cdots, \phi_n) = p_1(\phi_1)\ p_2(\phi_2) \cdots\ p_n(\phi_n)$$

Propriétés de la moyenne statistique

1/ Si $a_1, \ldots, a_n$ sont des constantes alors:

$$\langle a_1\phi_1 + \ldots + a_n\phi_n \rangle = a_1\langle\phi_1\rangle + \ldots + a_n\langle\phi_n\rangle$$

2/ Si A est une matrice ($m \times n$) alors:

$$\mathbf{\Phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{n-1} \\ \phi_n \end{pmatrix} \qquad \langle \mathbf{A\Phi} \rangle = \mathbf{A}\langle \mathbf{\Phi} \rangle$$

Covariances, corrélations et moments statistiques

- **Moments statistiques (de deuxième ordre):**

$$\langle \phi_k \phi_j \rangle = \int_D \phi_k \phi_j\, p(\phi_1, \cdots, \phi_n) d\phi_1 \cdots d\phi_n$$

  → Pour des variables *statistiquement indépendantes*:

$$\langle \varphi_j \varphi_k \rangle = \begin{cases} \sigma_j^2 + \mu_j^2 & j = k \\ \mu_j \mu_k & j \neq k \end{cases}$$

- **Covariances et variances:**

$$\mathrm{cov}(\phi_j, \phi_k) = \langle (\phi_j - \mu_j)(\phi_k - \mu_k) \rangle$$

$$\langle \phi_j \phi_k \rangle = \mathrm{cov}(\phi_j, \phi_k) + \mu_j \mu_k$$

$$\sigma_k^2 = \mathrm{cov}(\phi_k, \phi_k)$$

- **Corrélations:**

$$\rho(\phi_j, \phi_k) = \frac{\mathrm{cov}(\phi_j, \phi_k)}{\sigma_j \sigma_k}$$

Estimateur statistique de la moyenne

- **Moyenne statistique est estimée par la moyenne arithmétique de *N* mesures indépendantes:**

$$\bar{\phi} = \frac{1}{N}\sum_{i=1}^{N} \phi_i$$

$$\Phi = \Phi(\phi_1, \ldots, \phi_n)$$

$$p(\Phi) = p(\phi_1, \ldots, \phi_n) = p(\phi_1) \cdots p(\phi_n)$$

$$p(\phi_k) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-(\phi_k - \mu)^2 / 2\sigma^2\right\}$$

- **Moyenne statistique de Φ:** $\langle \bar{\phi} \rangle = \frac{1}{N}\sum_{k=1}^{N} \langle \phi_k \rangle = \mu$

  → Précision de cet estimé:

$$\langle (\bar{\phi} - \mu)^2 \rangle = \left\langle \left(\frac{1}{N}\sum_{k=1}^{N} (\phi_k - \mu) \right)^2 \right\rangle = \frac{1}{N^2} \left\langle \left(\sum_{k=1}^{N} (\phi_k - \mu) \right) \left(\sum_{j=1}^{N} (\phi_j - \mu) \right) \right\rangle$$

$$= \frac{1}{N^2}\sum_{k=1}^{N} \langle (\phi_k - \mu)^2 \rangle + \frac{1}{N^2}\sum_{k=1}^{N}\sum_{\substack{j=1 \\ j \neq k}}^{N} \langle (\phi_k - \mu)(\phi_j - \mu) \rangle = \frac{\sigma^2}{N} + \frac{1}{N^2}\sum_{k \neq j}^{N} \sigma_{jk} = \frac{\sigma^2}{N}$$

$$\frac{1}{N^2}\sum_{k=1}^{N} \sum_{j \neq k}^{N} \langle (\phi_k - \mu)(\phi_j - \mu) \rangle$$

Estimateur de la variance

On pose:

$$S^2 = \frac{1}{C}\sum_{k=1}^{N}\left(\phi_k - \overline{\phi}\right)^2 \equiv \frac{1}{C}\sum_{k=1}^{N}\left(\phi_k - \frac{1}{N}\sum_{j=1}^{N}\phi_j\right)^2 = \frac{1}{C}\sum_{k=1}^{N}\left[\phi_k^2 - \frac{2}{N}\sum_{j=1}^{N}\phi_k\phi_j + \frac{1}{N^2}\sum_{j=1}^{N}\sum_{m=1}^{N}\phi_j\phi_m\right]$$

où C est une constante à déterminer

Moyenne statistique de l'estimateur doit correspondre à celle de la variance

$$\left\langle S^2\right\rangle = \frac{1}{C}\sum_{k=1}^{N}\left[\left\langle\phi_k^2\right\rangle - \frac{2}{N}\sum_{j=1}^{N}\left\langle\phi_k\phi_j\right\rangle + \frac{1}{N^2}\sum_{j=1}^{N}\sum_{m=1}^{N}\left\langle\phi_j\phi_m\right\rangle\right]$$

Comme $\left\langle\varphi_j\varphi_k\right\rangle = \begin{cases}\sigma^2+\mu^2 & j=k\\ \mu^2 & j\neq k\end{cases}$, il s'ensuit que

$$-\frac{2}{N}\sum_{j=1}^{N}\left\langle\phi_k\phi_j\right\rangle = -\frac{2}{N}\left((\sigma^2+\mu^2)+(N-1)\mu^2\right)$$

$$\frac{1}{N^2}\sum_{j=1}^{N}\sum_{m=1}^{N}\left\langle\phi_j\phi_m\right\rangle = \frac{1}{N^2}\left(N(\sigma^2+\mu^2)+N(N-1)\mu^2\right)$$

Conséquemment,

$$\begin{aligned}\left\langle S^2\right\rangle &= \frac{1}{C}\left[N(\sigma^2+\mu^2)-2\left(\sigma^2+\mu^2+(N-1)\mu^2\right)+(\sigma^2+\mu^2)+(N-1)\mu^2\right]\\ &= \frac{1}{C}\left[(N-1)\sigma^2\right]\end{aligned}$$

En choisissant C = N-1, on obtient un estimateur correct (non biaisé) de la variance.

Estimateur

- **Moyenne d'ensemble de N mesures indépendantes** $\quad \overline{\phi} = \frac{1}{N}\sum_{k=1}^{N}\phi_k$
- **Ecart par rapport à la valeur réelle de la moyenne** $\quad \left\langle\left(\overline{\phi}-\langle\varphi\rangle\right)^2\right\rangle = \frac{\sigma^2}{N}$
- **Estimateur non biaisé de la variance** $\quad S^2 = \frac{1}{(N-1)}\sum_{k=1}^{N}(\phi_k-\overline{\phi})^2$

  $\left\langle S^2\right\rangle = \sigma^2$
- **Ecart par rapport à la valeur réelle de la variance** (von Storch et Zwiers, ch.5) $\quad \left\langle\left(S^2-\sigma^2\right)^2\right\rangle = \frac{2}{N-1}\sigma^4$

Distribution gaussienne à plusieurs variables

Cas où $\phi_1, \ldots, \phi_n$ sont statistiquement indépendantes

$$\begin{aligned}p(\phi_1,\cdots,\phi_n) &= p_1(\phi_1)\cdots p_n(\phi_n)\\ &= \frac{1}{\sigma_1\sqrt{2\pi}}\exp\left\{-\frac{(\phi_1-\mu_1)^2}{2\sigma_1^2}\right\}\cdots\frac{1}{\sigma_n\sqrt{2\pi}}\exp\left\{-\frac{(\phi_n-\mu_n)^2}{2\sigma_n^2}\right\}\\ &= \frac{1}{\sigma_1\sigma_2\ldots\sigma_n(2\pi)^{n/2}}\exp\left\{-\tfrac{1}{2}\mathbf{c}^T\mathbf{B}^{-1}\mathbf{c}\right\}\end{aligned}$$

où B est la matrice de covariance

$$\mathbf{c} = \begin{pmatrix}\phi_1-\mu_1\\ \vdots\\ \phi_n-\mu_n\end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix}\sigma_1^2 & 0 & \cdots & 0\\ 0 & \sigma_2^2 & \ddots & \vdots\\ \vdots & \ddots & \ddots & 0\\ 0 & \cdots & 0 & \sigma_n^2\end{pmatrix}$$

Covariances, variances et corrélations

Vecteur d'écart à la moyenne

$$\mathbf{c} = \begin{pmatrix} (\phi_1 - \mu_1) \\ (\phi_2 - \mu_2) \\ \vdots \\ (\phi_{n-1} - \mu_{n-1}) \\ (\phi_n - \mu_n) \end{pmatrix}$$

Covariances

$$\mathbf{B} = \langle \mathbf{c}\mathbf{c}^T \rangle = \begin{pmatrix} \langle c_1^2 \rangle & \langle c_1 c_2 \rangle & \cdots & \langle c_1 c_{n-1} \rangle & \langle c_1 c_n \rangle \\ \langle c_2 c_1 \rangle & \langle c_2^2 \rangle & \ddots & \langle c_2 c_{n-1} \rangle & \langle c_2 c_n \rangle \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \langle c_{n-1} c_1 \rangle & \ddots & \ddots & \langle c_{n-1}^2 \rangle & \langle c_{n-1} c_n \rangle \\ \langle c_n c_1 \rangle & \langle c_n c_2 \rangle & \cdots & \langle c_n c_{n-1} \rangle & \langle c_n^2 \rangle \end{pmatrix}$$

Remarque: l'erreur est considérée débiaisée

$$\mathbf{c} = \phi - \langle \phi \rangle$$

- **Matrice de corrélation C**

$$C_{ij} = \frac{\langle c_i c_j \rangle}{\langle c_i^2 \rangle^{1/2} \langle c_j^2 \rangle^{1/2}} = \frac{B_{ij}}{\sigma_i \sigma_j}$$

Corrélations

- **Forme matricielle**

$$\mathbf{B} = \Sigma \mathbf{C} \Sigma = \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \end{pmatrix} \begin{pmatrix} 1 & C_{12} & \cdots & C_{1n} \\ C_{21} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & C_{n-1,n} \\ C_{n1} & \cdots & C_{n,n-1} & 1 \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \end{pmatrix}$$

- **Corrélations exprimées en termes d'erreur normalisées**

$$\mathbf{d} = \begin{pmatrix} c_1 / \sigma_1 \\ c_2 / \sigma_2 \\ \vdots \\ c_n / \sigma_n \end{pmatrix} \qquad \mathbf{C} = \langle \mathbf{d}\mathbf{d}^T \rangle$$

Représentation des matrices de covariances

__Propriété__: **B** et **C** sont des matrices symétriques définies positives car pour tout vecteur **x**,

$$\mathbf{x}^T \mathbf{B} \mathbf{x} = \langle (\mathbf{x}^T \mathbf{e})(\mathbf{e}^T \mathbf{x}) \rangle > 0$$

__Propriété des matrices symétriques définies positives__:
valeurs propres sont réelles, distinctes et positives. De plus, les vecteurs propres $\mathbf{v}_1$, …, $\mathbf{v}_n$ sont orthonormaux. Ainsi,

$$\mathbf{B} = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n) \begin{pmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^2 \end{pmatrix} (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)^T \equiv \mathbf{V}^T \mathbf{D} \mathbf{V}.$$

V est une transformation orthogonale

__Propriété__: $\mathbf{V}^{-1} = \mathbf{V}^T$

$\text{dét}(\mathbf{V}) = 1$

Orthogonalité des vecteurs propres d'une matrice réelle symétrique

- **Considère deux vecteurs propres distincts de A**

$$\mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \qquad \mathbf{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$$

$$\mathbf{v}_2^T \mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_2^T \mathbf{v}_1 \qquad \mathbf{v}_1^T \mathbf{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_1^T \mathbf{v}_2$$

$$(\lambda_1 - \lambda_2) \mathbf{v}_1^T \mathbf{v}_2 = 0$$

- **Conséquemment**

$$\mathbf{v}_1^T \mathbf{v}_2 = 0$$

Orthogonalité des vecteurs propres d'une matrice réelle symétrique (suite)

- Puisque v_1, v_2, …, v_n sont des vecteurs propres orthonormaux alors

$$\mathbf{A}(\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n) = \mathbf{AV} = (\mathbf{Av}_1 \quad \mathbf{Av}_2 \quad \cdots \quad \mathbf{Av}_1)$$
$$= (\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n)$$
$$= \mathbf{V}\begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} = \mathbf{V}\Lambda$$

- Donc:

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

Transformation orthogonale

- Transformation associée à un changement de coordonnées défini par une base orthonormale

$$(\mathbf{e}_1 \quad \mathbf{e}_2 \quad \cdots \quad \mathbf{e}_n) \to (\mathbf{e}_1' \quad \mathbf{e}_2' \quad \cdots \quad \mathbf{e}_n')$$

- Donc et inversement

$$\mathbf{e}_1' = \alpha_{11}\mathbf{e}_1 + \cdots + \alpha_{1n}\mathbf{e}_n \qquad \mathbf{e}_1 = \alpha'_{11}\mathbf{e}'_1 + \cdots + \alpha'_{1n}\mathbf{e}'_n$$
$$\vdots \qquad\qquad\qquad \vdots$$
$$\mathbf{e}_n' = \alpha_{n1}\mathbf{e}_1 + \cdots + \alpha_{nn}\mathbf{e}_n \qquad \mathbf{e}_n = \alpha'_{n1}\mathbf{e}'_1 + \cdots + \alpha_{nn}\mathbf{e}'_n$$

- Comme $\alpha_{ij} = \mathbf{e'}_i^T\mathbf{e}_j$ alors $\alpha'_{ij} = \mathbf{e}_i^T\mathbf{e'}_j = \alpha_{ji}$
- Définissant

$$\mathbf{M} = (\mathbf{m}_1 \quad \mathbf{m}_2 \quad \cdots \quad \mathbf{m}_n) \quad \text{où} \quad \mathbf{m}_j^T = (\alpha_{1j} \quad \alpha_{2j} \quad \cdots \quad \alpha_{nj})^T$$

Transformation orthogonale (suite)

- Les vecteur m_j représentent les vecteurs unitaires e_j dans le système primé
  - → Donc: $\mathbf{m}_i^T\mathbf{m}_j = \delta_{ij}$
  - → Il s'ensuit que

$$\mathbf{M}^T\mathbf{M} = \begin{pmatrix} \mathbf{m}_1^T \\ \mathbf{m}_2^T \\ \vdots \\ \mathbf{m}_n^T \end{pmatrix}(\mathbf{m}_1 \quad \mathbf{m}_2 \quad \cdots \quad \mathbf{m}_n) = \mathbf{I}$$

  - → Et donc $\mathbf{M}^{-1} = \mathbf{M}^T$
- Considérons l'expression $\mathbf{Av} = \mathbf{w}$ où les vecteurs v et w sont exprimés en termes de composantes selon e_1 , …, e_n

Transformations orthogonales (suite)

- Donc

$$\mathbf{v}' = \mathbf{Mv} \quad \Leftrightarrow \quad \mathbf{v} = \mathbf{M}^T\mathbf{v}'$$
$$\mathbf{Av} = \mathbf{w} \quad \Rightarrow \quad \mathbf{AM}^T\mathbf{v}' = \mathbf{M}^T\mathbf{w}'$$
$$\mathbf{MAM}^T\mathbf{v}' = \mathbf{w}' \quad \Rightarrow \quad \mathbf{A}'\mathbf{v}' = \mathbf{w}'$$

Variables statistiquement indépendantes associées à B

- **Changement de variable orthogonal:** $\xi = \mathbf{V}\,\phi$

- **Distribution gaussienne: cas général**

$$p_\xi(\xi_1,\cdots,\xi_n) = \frac{1}{\lambda_1\sqrt{2\pi}}\exp\left\{-\frac{\xi_1^2}{2\lambda_1^2}\right\}\cdots\frac{1}{\lambda_n\sqrt{2\pi}}\exp\left\{-\frac{\xi_n^2}{2\lambda_n^2}\right\}$$

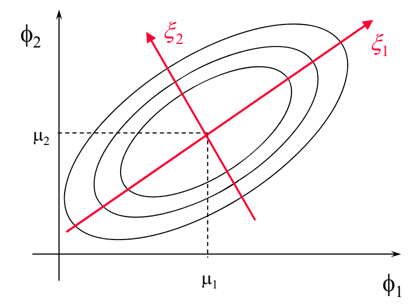
$$= \frac{1}{(\det\mathbf{D})^{1/2}(2\pi)^{n/2}}\exp\left\{-\tfrac{1}{2}(\mathbf{V}\varphi)^T\mathbf{D}^{-1}(\mathbf{V}\varphi)\right\}$$

$$= \frac{1}{(d\acute{e}t\,\mathbf{B})^{1/2}(2\pi)^{n/2}}\exp\left\{-\tfrac{1}{2}\varphi^T\mathbf{B}^{-1}\varphi\right\} = p(\phi_1,\cdots,\phi_n)$$

- **Normalisation peut être obtenue en utilisant**

$$\chi = \mathbf{D}^{-1/2}\mathbf{V}\varphi \quad\Rightarrow\quad p(\chi_1,\cdots,\chi_n) = \frac{1}{(2\pi)^{n/2}}\exp\left\{-\tfrac{1}{2}\chi^T\chi\right\}$$

Exemple illustrant $\phi^T\,\mathbf{B}^{-1}\phi$ pour $\phi = (\phi_1,\phi_2)^T$



Covariance d'erreur d'un champ de température

- Variable aléatoire: $T = (T_1, \ldots, T_n)$ où $T_i = T(x_i,y_i)$
- Modèle de covariance d'erreur:

$$B_{ij} = \sigma_i\sigma_j f(|p_i - p_j|)$$

où $p_i \equiv (x_i,y_i)$.

- Corrélations sont dites homogènes et isotropes car elles ne dépendent que de la distance entre les points p_i et p_j.

- <u>Remarque</u>: la fonction de corrélation peut avoir un profil gaussien ou non. Il n'en demeure pas moins que la distribution de probabilité demeure gaussienne.

Référence: von Storch et Zwiers, 1999: *Statistical analysis in Climate Research*

Section V: Eigen techniques. Ch-13 à 16

FONCTIONS EMPIRIQUES ORTOGONALES

SCA-7216 Introduction à l'assimilation de données

Fonctions empiriques orthogonales

- **Evaluation des modes associés à la variabilité atmosphérique moyennée sur une période de temps**
- **Exemple: variabilité de la température de la surface de la mer (Sea Surface Temperature, SST)**
  - → Considérons que nous avons une suite de champs de température de surface la mer $T_s(\lambda,\varphi,t_k)=T_k$.
  - → Sur une période de temps t_1 à t_k, on calcule la moyenne temporelle $\overline{T}_s(\lambda,\varphi)$ de T_s comme étant

$$\overline{T}_s \equiv \overline{T}_s(\lambda,\varphi) = \frac{1}{K}\sum_{k=1}^{K} T_s(\lambda,\varphi,t_k)$$

  - → La variabilité par rapport à cette moyenne est ensuite définie, à chaque instant, comme étant

$$T_k^{'}(\lambda,\varphi) = \left(T_s(\lambda,\varphi,t_k) - \overline{T}_s\right)$$

Définition de la covariance de la variabilité

- La covariance spatiale est estimée par

$$B(\lambda,\varphi;\lambda',\varphi') = \frac{1}{(K-1)}\sum_{k=1}^{K} T_k^{'}(\lambda,\varphi) T_k^{'}(\lambda',\varphi')$$

- Ceci représente la covariance d'erreur entre deux points de longitude et latitude (λ,φ) et (λ',φ') respectivement moyennée à partir d'un échantillon de taille **K**

- Sur une grille comportant N_x points en λ et N_y points en φ, chaque champ comprend donc **$N = N_x$x N_y** degrés de liberté
  - → Les points de grille sont de coordonnées

$$(\lambda_1,\varphi_1) \quad (\lambda_1,\varphi_2) \quad \ldots \quad (\lambda_N,\varphi_N)$$

  - → Le vecteur de variabilité est alors défini par $\mathbf{E}_k = \frac{1}{\sqrt{K-1}}\begin{pmatrix} T_k^{'}(\lambda_1,\varphi_1) \\ T_k^{'}(\lambda_1,\varphi_2) \\ \vdots \\ T_k^{'}(\lambda_N,\varphi_N) \end{pmatrix}$

Équivalence entre les formes matricielle et scalaire

On définit $p_\nu = \left(\lambda_{i(\nu)},\varphi_{j(\nu)}\right)$ avec $\nu = 1,\cdots,N_x \times N_y = N$

$$\mathbf{B} = \sum_{k=1}^{K}\begin{pmatrix} E_k'(p_1) \\ \vdots \\ E_k'(p_n) \end{pmatrix}\begin{pmatrix} E_k'(p_1) & \cdots & E_k'(p_n) \end{pmatrix}$$

$$= \sum_{k=1}^{K}\begin{pmatrix} E_k'(p_1)E_k'(p_1) & \cdots & E_k'(p_1)E_k'(p_n) \\ \vdots & \ddots & \vdots \\ E_k'(p_1)E_k'(p_1) & \cdots & E_k'(p_n)E_k'(p_n) \end{pmatrix} = \sum_{k=1}^{K}\mathbf{E}_k\mathbf{E}_k^T$$

$$= \begin{pmatrix} E_1'(p_1) & E_2'(p_1) & \cdots & E_k'(p_1) \\ E_1'(p_2) & E_2'(p_2) & \cdots & \\ \vdots & \vdots & \vdots & \vdots \\ E_1'(p_n) & E_2'(p_n) & \cdots & \end{pmatrix}\begin{pmatrix} E_1'(p_1) & E_1'(p_1) & \cdots & E_1'(p_n) \\ E_2'(p_1) & E_2'(p_2) & \cdots & E_2'(p_n) \\ \vdots & \vdots & \vdots & \vdots \\ E_n'(p_1) & E_n'(p_2) & \cdots & E_n'(p_n) \end{pmatrix}$$

$$= \mathbf{E}\mathbf{E}^T$$

SCA-7216 Introduction à l'assimilation de données

Définition de la covariance de la variabilité

- Avec cette notation, la covariance d'erreur devient

$$\mathbf{B} = \sum_{k=1}^{K}\mathbf{E}_k\mathbf{E}_k^T$$

- On forme une matrice **E** dont les *colonnes* correspondent aux vecteurs E_1, E_2, …, E_K, c'est-à-dire que

$$\mathbf{E} = \begin{pmatrix}\mathbf{E}_1 & \mathbf{E}_2 & \cdots & \mathbf{E}_k\end{pmatrix}$$

- Conséquemment $\mathbf{B} = \mathbf{E}\mathbf{E}^T$

- Cette matrice est d'ordre NxN

- Remarque Comme E est une matrice *rectangulaire* d'ordre NxK, il s'ensuit que E^T est d'ordre (KxN)

Fonction empirique orthogonale

- **Définition.** Les vecteurs propres de la matrice **B** sont appelés *fonctions empiriques orthogonales* (ou EOF de l'anglais *Empirical Orthogonal functions*)
- Calcul nécessite de *définir* les N(N+1)/2 coefficients de la matrice et ensuite d'en trouver les vecteurs propres
  - → Les vecteurs propres de B définis comme

$$\mathbf{Bv} = \left(\mathbf{EE}^T\right)\mathbf{v} = \lambda\mathbf{v}$$

- Structure de la matrice symétrique permet de réduire considérablement le calcul.

Fonctions empiriques orthogonales et composantes principales

- La forme de **B** = **EE**$^{\mathbf{T}}$ permet d'établir une relation entre les vecteurs propres de **B** et ceux de la matrice **E**$^{\mathbf{T}}$**E** qui est d'ordre (KxK) ce qui est beaucoup moins élevé que l'ordre de **B** (NxN).
  - → Typiquement, K est d'ordre 10^2 alors que N peut facilement être d'ordre égal ou plus grand à 10^6

Considérant que

$$\mathbf{Bv} = \mathbf{EE}^T\mathbf{v} = \lambda\mathbf{v}$$

On multiplie à gauche par $\mathbf{E}^T$ pour obtenir

$$\left(\mathbf{E}^T\mathbf{E}\right)\mathbf{E}^T\mathbf{v} = \lambda\mathbf{E}^T\mathbf{v}$$

ce qui se réécrit

$$\left(\mathbf{E}^T\mathbf{E}\right)\mathbf{w} = \lambda\mathbf{w} \qquad (1)$$

où

$$\boxed{\mathbf{w} = \mathbf{E}^T\mathbf{v} \quad \Rightarrow \quad \mathbf{v} = \mathbf{Ew}}$$

Composantes principales et EOF

- **Echantillon basé sur une série temporelle de K instants**
  - → champ fourni à toutes les 6 heures durant 60 jours donne un échantillon de taille K=360
- **Les vecteurs propres w de E$^{\mathbf{T}}$E sont de dimension K de telle sorte que les K composantes sont**

$$\mathbf{w} = \begin{pmatrix} w(t_1) \\ w(t_2) \\ \vdots \\ w(t_K) \end{pmatrix}$$

- **Les EOFs sont déduites de (1) et sont telles que**

$$\mathbf{v} = \mathbf{Ew} = \sum_{k=1}^{K} w(t_k)\mathbf{E}_k$$

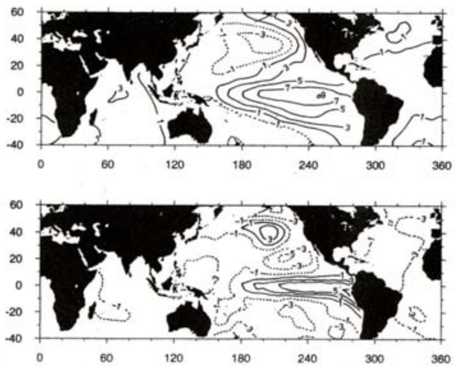
Interprétation

- Composantes principales définissent les poids $w(t_k)$ d'une combinaison linéaire qui définit les vecteurs propres de la matrice B
  - → w(t) est un série temporelle
- Vecteurs propres de B sont des structures purement spatiales
- Les définitions précédentes permettent d'écrire que

$$\mathrm{v}(\lambda,\varphi) = \frac{1}{(K-1)}\sum_{k=1}^{K} w(t_k)T_k^{'}(\lambda,\varphi)$$

- Exemple: analyse des moyennes mensuelles de la température de la surface de la mer (SST) sur la période de 1970 à 1990 (K=252)

Deux premières EOFs (tiré de von Storch et Zwiers p.305)

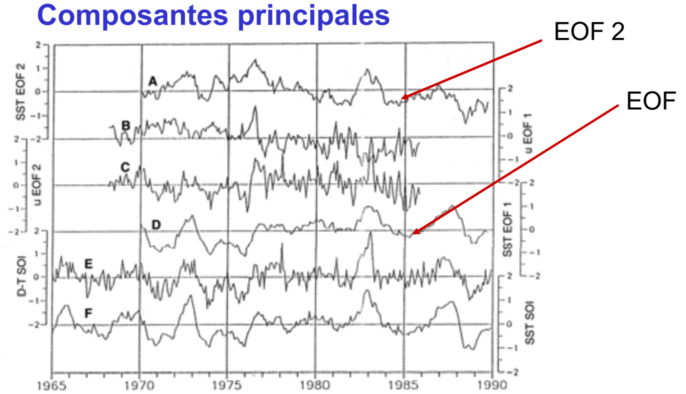


Variabilité sur l'est du Pacifique est relié au mécanisme du '*El Niño Southern Oscillation*' (ENSO)

EOF 1: 27.1%
EOF 2: 7.9 %
De la variance totale

Figure 13.3: *EOFs 1 (top) and 2 (bottom) of monthly mean sea-surface temperature (SST). Units: 10^{-2}. Courtesy Xu.*

Composantes principales



Série D représente la composante principale associée avec la première EOF. Sa forme concorde assez bien avec la courbe F qui est un indice climatologique mesurant l'intensité de la *Southern Oscillation*

Remarques

- **Chapitres 13 à 16 de von Storch et Zwiers**
  - → Canonical Correlation Analysis (CCA) (Analyse canonique des corrélations)
  - → Même approche que pour les EOFS mais pour traiter les corrélations
  - → Permet de considérer des corrélations entre différentes variables (e.g., vents et température)
- **Covariances permettent d'établir des liens pour des mécanismes existant au même moment.**

Estimation statistique: cas univarié (1 point)

X_b = prévision (ou champ d'essai)
X_o = observation
X_a = analyse
et X_t = état réel

$\varepsilon_b = (\mathbf{X}_b - \mathbf{X}_t)$: erreur de prévision
$\varepsilon_o = (\mathbf{X}_o - \mathbf{X}_t)$: erreur d'observation
$\varepsilon_a = (\mathbf{X}_a - \mathbf{X}_t)$: erreur d'analyse

---

Hypothèses sous-jacentes à l'interpolation statistique

- Erreur d'observation et de prévision sont non biaisées $\langle\varepsilon_b\rangle = \langle\varepsilon_o\rangle = 0$
- variances d'erreur σ_b^2 et σ_o^2 sont connues

Meilleur estimé linéaire non-biaisé
(*Best Linear Unbiased Estimate*) (BLUE)

- Analyse $\mathbf{X}_a = \mathbf{X}_b + \lambda(\mathbf{X}_o - \mathbf{X}_b)$

- Variance d'erreur d'analyse

$$
\begin{aligned}
J(\lambda) = \sigma_a^2 &= \left\langle (\mathbf{X}_a - \mathbf{X}_t)^2 \right\rangle \\
&= \left\langle (\mathbf{X}_b - \mathbf{X}_t + \lambda((\mathbf{X}_o - \mathbf{X}_t) - (\mathbf{X}_b - \mathbf{X}_t))^2 \right\rangle \\
&= \left\langle ((1-\lambda)\varepsilon_b + \lambda\varepsilon_o)^2 \right\rangle = (1-\lambda)^2 \left\langle \varepsilon_b^2 \right\rangle + 2\lambda(1-\lambda)\left\langle \varepsilon_b \varepsilon_o \right\rangle + \lambda^2 \varepsilon_o^2 \\
&= \sigma_b^2 (1-\lambda)^2 + \lambda^2 \sigma_o^2
\end{aligned}
$$

- Estimé de λ minimisant la variance d'erreur d'analyse:

$$
\frac{dJ}{d\lambda} = 0 \quad \Rightarrow \quad \lambda^* = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_o^2} \Rightarrow \frac{1}{\sigma_a^2} = \frac{1}{\sigma_b^2} + \frac{1}{\sigma_o^2}
$$

$$
\Rightarrow \sigma_a^2 < \min\left\{\sigma_b^2, \sigma_o^2\right\}
$$

Opérateur d'observation

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

T_b — Niveau du modèle

$T_{obs.}$ — Niveau à 10 m

Opérateur d'observation: HT = α T

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

Estimé minimisant la variance d'erreur d'analyse:

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

Interpolation statistique
(ou méthode de Gauss-Markov)

(Gandin, 1963; Rutherford, 1973; Schlatter,1977; Daley, 1991)

Définitions

- $\mathbf{y} \in R^m$: vecteur d'observation (m ~ 10^6)
- $\mathbf{X} \in R^n$: vecteur d'état du modèle (n ~ 10^8)
- **H:** $R^n \to R^m$: Opérateur d'observation
- $\mathbf{B} = \langle \varepsilon_b \varepsilon_b^T \rangle$: covariances d'erreur de prévision
- $\mathbf{R} = \langle \varepsilon_o \varepsilon_o^T \rangle$: covariances d'erreur d'observation

Hypothèses

Erreurs d'observation et de prévision sont supposées être non-corrélées

$$\langle \varepsilon_b \varepsilon_o^T \rangle = 0.$$

Erreurs de prévision et d'observation ne sont pas biaisées:

$$\langle \varepsilon_b \rangle = \langle \varepsilon_o \rangle = 0$$

Estimateur de variance minimale: cas univarié

- **Analyse univariée: analyse d'un seul champ météorologique en chaque point de la grille du modèle**
- **Dimension de la variable aléatoire:**

  $\to$ (# niveaux) ´ (# points de grille) = 80 $\times$(1024 $\times$ 800) ~ 10^8

- **Autocovariances d'erreur:** $B_{ij} = \langle T_i T_j \rangle$

  $\to$ **B** est de dimension ($10^8 \times 10^8$) = 10^{16}.

- **Estimateur linéaire:** $\mathbf{X}_a = \mathbf{X}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{X}_b)$

  $\to$ **K** est maintenant une matrice (N $\times$ M)

Dérivation de la matrice de gain K minimisant la variance d'erreur d'analyse

- **Variance totale d'erreur d'analyse:**

$$
\begin{aligned}
J(\mathbf{K}) &= \varepsilon_a^T\varepsilon_a = (\mathbf{x}_a-\mathbf{x}_t)^T(\mathbf{x}_a-\mathbf{x}_t)\\
&= \left(\varepsilon_b+\mathbf{K}(\varepsilon_0-\mathbf{H}\varepsilon_b)\right)^T\left(\varepsilon_b+\mathbf{K}(\varepsilon_0-\mathbf{H}\varepsilon_b)\right)\\
&= \varepsilon_b^T\varepsilon_b+(\varepsilon_o-\mathbf{H}\varepsilon_b)^T\mathbf{K}^T\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b)+\varepsilon_b^T\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b)+(\varepsilon_o-\mathbf{H}\varepsilon_b)^T\mathbf{K}^T\varepsilon_b
\end{aligned}
$$

- **Calcul de $\partial J/\partial \mathbf{K}$ = M:** $M_{ij} = \partial J/\partial K_{ij}$

- **Quelques identités:**

$$\frac{\partial}{\partial \mathbf{K}}\mathbf{X}^T\mathbf{K}^T\mathbf{K}\mathbf{X} = 2\mathbf{K}\mathbf{X}\mathbf{X}^T$$

$$\frac{\partial}{\partial \mathbf{K}}\mathbf{Y}^T\mathbf{K}\mathbf{X} = \mathbf{Y}\mathbf{X}^T$$

$$\frac{\partial}{\partial \mathbf{K}}\mathbf{Y}^T\mathbf{K}^T\mathbf{X} = \mathbf{X}\mathbf{Y}^T$$

Démonstration

1. $\dfrac{\partial}{\partial \mathbf{K}}\mathbf{Y}^T\mathbf{K}\mathbf{X} = \mathbf{Y}\mathbf{X}^T$

$$\left(\mathbf{Y}^T\mathbf{K}\mathbf{X}\right) = \sum_l\sum_m Y_l K_{lm} X_m$$

$$\left[\frac{\partial}{\partial K_{ij}}\mathbf{Y}^T\mathbf{K}\mathbf{X}\right]_{ij} = Y_i X_j$$

2. $\dfrac{\partial}{\partial \mathbf{K}}\mathbf{Y}^T\mathbf{K}^T\mathbf{X} = \mathbf{X}\mathbf{Y}^T$

$$\left(\mathbf{Y}^T\mathbf{K}^T\mathbf{X}\right) = \sum_l\sum_m Y_l K_{lm}^T X_m$$

$$\left[\frac{\partial}{\partial K_{ij}}\mathbf{Y}^T\mathbf{K}^T\mathbf{X}\right]_{ij} = Y_j X_i$$

3. $\dfrac{\partial}{\partial \mathbf{K}}\mathbf{X}^T\mathbf{K}^T\mathbf{K}\mathbf{X} = 2\mathbf{K}\mathbf{X}\mathbf{X}^T$

$$\left(\mathbf{X}^T\mathbf{K}^T\mathbf{K}\mathbf{X}\right) = \sum_l\sum_m\sum_n X_n K_{ln} K_{lm} X_m$$

$$
\begin{aligned}
\left[\frac{\partial}{\partial K_{ij}}\mathbf{X}^T\mathbf{K}^T\mathbf{K}\mathbf{X}\right]_{ij} &= X_j K_{im} X_m + X_n K_{in} X_j\\
&= 2(\mathbf{K}\mathbf{X})_i X_j
\end{aligned}
$$

Matrice de gain

$$
\begin{aligned}
\frac{\partial}{\partial \mathbf{K}}J(\mathbf{K}) &= \frac{\partial}{\partial \mathbf{K}}\left(\varepsilon_b^T\varepsilon_b+(\varepsilon_o-\mathbf{H}\varepsilon_b)^T\mathbf{K}^T\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b)+\varepsilon_b^T\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b)+(\varepsilon_o-\mathbf{H}\varepsilon_b)^T\mathbf{K}^T\varepsilon_b\right)\\
&= 2\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b)(\varepsilon_o-\mathbf{H}\varepsilon_b)^T+\varepsilon_b(\varepsilon_o-\mathbf{H}\varepsilon_b)^T+\varepsilon_b(\varepsilon_o-\mathbf{H}\varepsilon_b)^T\\
&= \mathbf{0}
\end{aligned}
$$

En prenant l'espérance mathématique de ce résultat et en utilisant les propriétés et définitions introduites plus haut:

$\langle\varepsilon_b\varepsilon_o^T\rangle = 0$; $\mathbf{B} = \langle\varepsilon_b\varepsilon_b^T\rangle$; $\mathbf{R} = \langle\varepsilon_o\varepsilon_o^T\rangle$

on obtient

$$
\begin{aligned}
\frac{\partial}{\partial \mathbf{K}}J(\mathbf{K}) &= 2\mathbf{K}\left[\langle\varepsilon_o\varepsilon_o^T\rangle-\mathbf{H}\langle\varepsilon_b\varepsilon_o^T\rangle-\langle\varepsilon_o\varepsilon_b^T\rangle\mathbf{H}^T+\mathbf{H}\langle\varepsilon_b\varepsilon_b^T\rangle\mathbf{H}^T\right]+2\langle\varepsilon_b\varepsilon_o^T\rangle-2\langle\varepsilon_b\varepsilon_b^T\rangle\mathbf{H}^T\\
&= 2\mathbf{K}\left[\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T\right]-2\mathbf{B}\mathbf{H}^T = 0
\end{aligned}
$$

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

Covariances d'erreur d'analyse

- **Erreur d'analyse** $\varepsilon_a = \varepsilon_b+\mathbf{K}(\varepsilon_o-\mathbf{H}\varepsilon_b) = (\mathbf{I}-\mathbf{K}\mathbf{H})\varepsilon_b+\mathbf{K}\varepsilon_o$
- **Covariances d'erreur d'analyse**

$$
\begin{aligned}
\mathbf{P}_a &= \langle\varepsilon_a\varepsilon_a^T\rangle\\
&= \left\langle\left((\mathbf{I}-\mathbf{K}\mathbf{H})\varepsilon_b+\mathbf{K}\varepsilon_o\right)\left((\mathbf{I}-\mathbf{K}\mathbf{H})\varepsilon_b+\mathbf{K}\varepsilon_o\right)^T\right\rangle\\
&= \left\langle\left((\mathbf{I}-\mathbf{K}\mathbf{H})\varepsilon_b+\mathbf{K}\varepsilon_o\right)\left(\varepsilon_b^T(\mathbf{I}-\mathbf{K}\mathbf{H})^T+\varepsilon_o^T\mathbf{K}^T\right)\right\rangle\\
&= (\mathbf{I}-\mathbf{K}\mathbf{H})\langle\varepsilon_b\varepsilon_b^T\rangle(\mathbf{I}-\mathbf{K}\mathbf{H})^T+\mathbf{K}\langle\varepsilon_o\varepsilon_o^T\rangle\mathbf{K}^T+\underbrace{(\mathbf{I}-\mathbf{K}\mathbf{H})\langle\varepsilon_b\varepsilon_o^T\rangle\mathbf{K}^T}_{=0}+\underbrace{\mathbf{K}\langle\varepsilon_o\varepsilon_b^T\rangle(\mathbf{I}-\mathbf{K}\mathbf{H})^T}_{=0}\\
&= (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I}-\mathbf{K}\mathbf{H})^T+\mathbf{K}\mathbf{R}\mathbf{K}^T
\end{aligned}
$$

A l'optimalité, on a $\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$
Conséquemment,

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

Equations de l'interpolation statistique

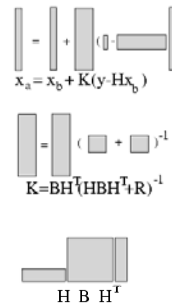
$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$
$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$$
$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}$$

Matrice de gain (ou poids statistiques): K

Innovations (ou écarts aux observation): $\mathbf{y} - \mathbf{H}\mathbf{x}_b$

Dimensionalité du problème

- **Nombre d'observations:** **M**
- **Etat-modèle:** **N**
- **Vecteur d'état $\mathbf{x}_a$ et $\mathbf{x}_b$:** $\mathbf{x} = (\mathrm{N} \times 1)$
- **Vecteur d'observation:** $\mathbf{y} = (\mathrm{M} \times 1)$
- **Equivalent modèle de l'observation:** $\mathbf{H}\mathbf{x}_b = (\mathrm{N} \times 1)$
- **Opérateur d'observation** $\mathbf{H} = (\mathrm{M} \times \mathrm{N})$
- **Matrice de gain** $\mathbf{K} = (\mathrm{N} \times \mathrm{M})$
- **Covariance d'erreur de prévision** $\mathbf{B} = ((\mathrm{N} \times \mathrm{N})$
- **Covariance d'erreur d'observation** $\mathbf{R} = (\mathrm{M} \times \mathrm{M})$
- **Erreur de prévision dans l'espace des observations** $\mathbf{H}\mathbf{B}\mathbf{H}^T = (\mathrm{M} \times \mathrm{M})$



Exemple: cas unidimensionnel T(x)

- **Covariance d'erreur de prévision**

$$B(x_1, x_2) = \sigma_b(x_1)\sigma_b(x_2)\exp\left\{-\frac{(x_1 - x_2)^2}{2L^2}\right\}$$

L est la longueur caractéristique de la fonction de corrélation

- **Assimilation d'une seule observation**

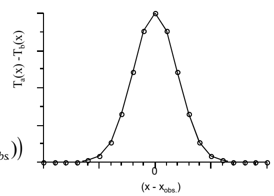
$$H(x) = \delta(x - x_{obs.})$$
$$(\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) = (\sigma_o^2 + \sigma_b^2(x_{obs.}))$$

- **Forme matricielle de H**

$$\mathbf{H} = \begin{pmatrix} 0 & \cdots & 0 & \underset{x = x_{obs}}{1} & 0 & \cdots & 0 \end{pmatrix}$$

- **Incrément d'analyse**

$$T_a(x) - T_b(x) = \frac{\sigma_b(x_{obs})\sigma_b(x)}{(\sigma_b^2(x_{obs}) + \sigma_o^2)}\exp\left\{-\frac{(x - x_{obs})^2}{2L^2}\right\}(T_{obs} - T_b(x_{obs}))$$



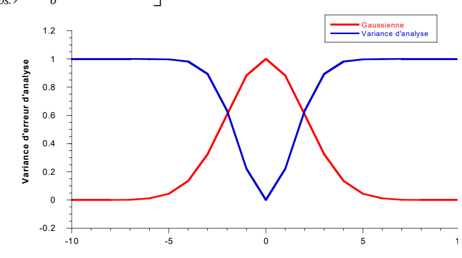
Covariance d'erreur d'analyse

$$(\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} = \left(\mathbf{I} - \frac{1}{\sigma_o^2 + \sigma_b^2}\mathbf{B}\mathbf{H}^T\mathbf{H}\right)\mathbf{B} \qquad \mathbf{B}\mathbf{H}^T(x) = \sigma_b(x_{obs})\sigma_b(x)e^{-(x - x_{obs})^2/2L^2}$$

- **Expression pour la variance d'erreur d'analyse**

$$\mathbf{B}\mathbf{H}^T(\mathbf{B}\mathbf{H}^T)^T(x, x) = \sigma_b^2(x_{obs.})\sigma_b^2(x)e^{-(x - x_{obs.})^2/L^2}$$

$$\sigma_a^2(x) = \sigma_b^2(x)\left[1 - \frac{\sigma_b^2(x_{obs.})}{\sigma_b^2(x_{obs.}) + \sigma_o^2}e^{-(x - x_{obs.})^2/L^2}\right]$$



Exemple: assimilation de radiances

- Dans ce cas, pour chaque longueur d'onde λ, la radiance $R(\lambda)$ mesurée au sommet de l'atmosphère est telle que

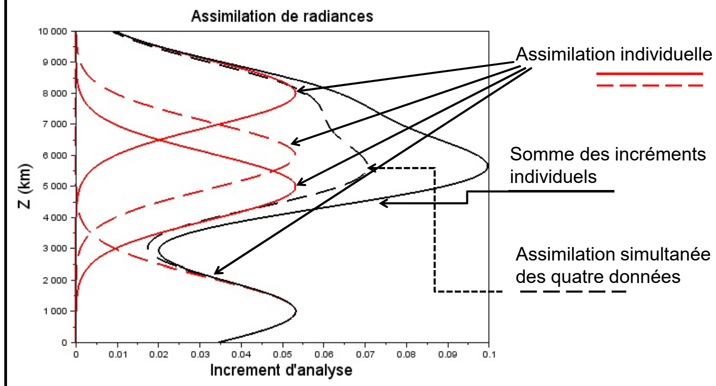
$$R(\lambda) = \int_0^\infty \tau(\lambda; z) T(z)\, dz$$

- La forme discrétisée de ceci est

$$R(\lambda) = \begin{pmatrix} \tau(z_1) w_1 & \tau(z_2) w_2 & \cdots & \tau(z_N) w_N \end{pmatrix} \begin{pmatrix} T(z_1) \\ T(z_2) \\ \vdots \\ T(z_N) \end{pmatrix} \equiv \mathbf{HT}$$

- Programme SCA7216_radiances.sce illustre l'assimilation de quatre radiances à différentes longueur d'ondes prises individuellement ou ensemble

Assimilation de quatre radiances pour quatre longueurs d'onde différentes



Hypothèses utilisées dans la dérivation

- **Linéarisation des opérateurs H intervient**
- **Hypothèses sur les statistiques d'erreur**

  $\langle \varepsilon_b \varepsilon_o^T \rangle = 0; \qquad \langle \varepsilon_b \rangle = \langle \varepsilon_o \rangle = 0$
- **Généralisation aux estimateurs donnant l'état le plus probable.**
- **Implications et exemples:**
  - $\rightarrow$ Daley (1991) ch.4, sections 4.5 et 4.6
- **Historique de l'assimilation de données**
  - $\rightarrow$ Daley (1991): chapître 1