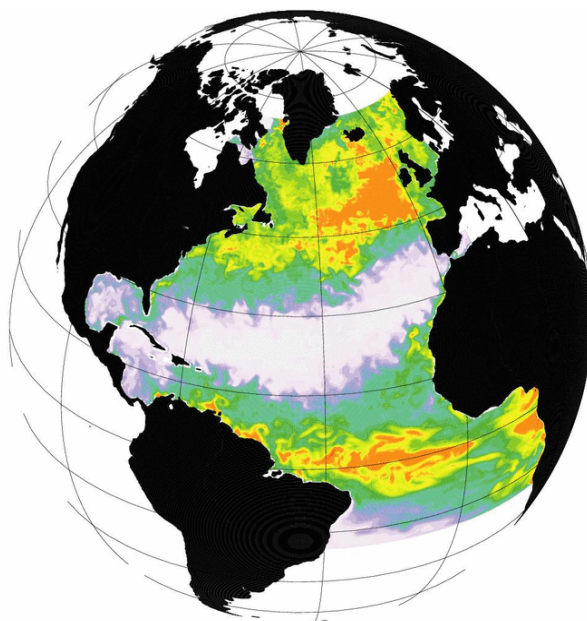




ASSIMILATION DE DONNÉES :
Formulation Ensembliste Du 4D-Var



4 JUILLET 2012

AUTEUR :
JEAN-THOMAS CAMINO

ENCADRANTS :
M. GÉRALD DESROZIERS / M. LOIK BERRE

Table des matières

Introduction	2
1 Tour d’horizon des notions importantes	3
1.1 Théorie de l’estimation	3
1.2 Les observations	3
1.3 Principe et formalisme	4
1.4 Forme alternative de la solution	6
1.5 Signaux périodiques 1D	7
1.6 Non-linéarité de l’opérateur H et formulation 3D-Var classique	8
1.7 Filtre de Kalman et impossibilité numérique	9
2 Le 4D-Var	11
2.1 La formulation 4D-Var incrémentale classique	11
2.2 Evolution implicite des covariances d’erreur d’ébauche	13
2.3 Conditionnement et 4D-Var incrémental pré-conditionné	15
2.4 Le 4D-En-Var sans localisation	18
2.5 Le 4D-En-Var avec localisation	20
2.6 Vers une nouvelle formulation du 4D-Var	23
2.6.1 Présentation de l’idée	23
2.6.2 Formulation plus générale	26
2.6.3 Nouvelle formulation ensembliste	27
3 Une première approche : le En-1D-Var	32
3.1 Ossature du code	32
3.2 Tests de sensibilité avec une matrice des covariances d’erreur d’ébauche homogène	34
3.2.1 Influence de la taille de l’ensemble	34
3.2.2 Influence de la portée de la matrice C du produit de Schur	35
Annexe	37
Propriétés des matrices circulantes	37
Création de perturbations ϵ_0 et ϵ_b cohérentes	42
Equivalence du problème en points de grille et du problème spectral en 4D-Var	43
Equivalence du problème en points de grille et du problème spectral en 4D-En-Var	45

Introduction

Dans le processus de prévision de l'évolution de l'état de l'atmosphère, la dépendance à l'état initial semble évidente. Or, il faut bien avoir conscience qu'il est actuellement impossible de déterminer avec exactitude à un instant donné cet état. C'est là que l'assimilation de données intervient. En effet, son but est de fournir une représentation aussi fidèle et cohérente que possible de l'état de l'atmosphère avant de lancer une nouvelle prévision, et ce en utilisant toutes les sources d'information à notre disposition. Concrètement, il va s'agir de corriger les ébauches mises en place antérieurement à l'aide de toutes les observations effectuées sur des plages temporelles bien définies.

Un des défauts de la méthode d'estimation de l'état atmosphérique actuelle est de ne pas prendre en compte l'évolution spatio-temporelle des covariances d'erreurs d'ébauche. Les méthodes d'ensemble que nous allons apprendre à maîtriser tout le long de ce stage vont nous permettre de nous affranchir de ces hypothèses d'homogénéité, d'isotropie et de statisme, et donc rendre le calcul de prévision bien plus précis et efficace.

1 Tour d'horizon des notions importantes

1.1 Théorie de l'estimation

Certains principes simples sont à l'origine de toute la théorie d'assimilation de données. On se contente ici d'en faire une brève énumération :

- Les observations sont imparfaites
- L'estimé est donc un résultat approché
- L'estimé est une combinaison optimale des observations
- Les modèles ne sont jamais parfaits

1.2 Les observations

Pour les observations, on note une hétérogénéité :

- spatiale
- temporelle
- des paramètres mesurés
- en précision

Elles sont naturellement de différents types. Les dix catégories usuelles sont les suivantes :

- SYNOP : mesures de surface
- AIREP : mesures fournies par les avions de ligne
- SATOB : mesures indirectes du vent
- DRIBU : mesures fournies par les bouées
- TEMP : mesures fournies par les radiosondages
- PILOT : mesures fournies par les sondages PILOT
- SATEM : profils de température inversés, ou radiances
- PAOB : pseudo-observations
- SCATT : vent à la surface des océans
- RARAD : radiances brutes

1.3 Principe et formalisme

Plaçons-nous d'emblée dans le cas multi-dimensionnel. Nous noterons alors le champ analysé x , vecteur de taille n . C'est l'ensemble des variables estimées du modèle, ce qui dans le cadre de l'assimilation donne lieu à 3 vecteurs dérivés : x^t l'état réel de l'atmosphère, x^b l'état prévu, x^a l'état analysé. De même, le champ d'observation sera noté y^o et sera un vecteur de taille p . Les localisations spatiales (et même spatio-temporelles, cf. le 4D) des observations sont rarement les mêmes que celles des variables estimées. De plus, elles ne sont pas toujours de même nature. Ainsi, en notant ϵ^0 l'erreur d'observation, on supposera qu'il existe un opérateur H tel que :

$$y^o = H x^t + \epsilon^0$$

Et que nous disposons d'une ébauche qui a une expression similaire :

$$x^b = x^t + \epsilon^b$$

L'ébauche et les instruments ne sont pas biaisés ce qui se traduit par :

$$E(\epsilon^b) = E(\epsilon^0) = 0$$

Les précisions (variances) des instruments sont connues. Tout comme les covariances d'erreurs d'observation (deux instruments différents sont décorrélés, mais si un instrument mesure plusieurs composantes de y^o alors les erreurs sur ces variables risquent d'être corrélées et on suppose qu'on sait le quantifier correctement). Finalement, tout ceci revient à connaître la matrice suivante :

$$E(\epsilon^0 T \epsilon^0) = \mathbf{R}$$

Dans l'opérateur H , on trouve :

- Le passage de l'espace spectral à l'espace physique
- Des combinaisons entre différentes variables estimées
- L'interpolation des 4 (ou des 12) points de grille encadrant l'observation
- L'interpolation suivante la verticale jusqu'au niveau de l'observation
- En 4D, l'intégration du modèle d'évolution

Rappelons que nous cherchons une estimation, notée x^a , la plus fidèle possible et qui prenne donc en compte les observations et l'ébauche précédente. C'est en fait leur combinaison optimale au sens des variances d'erreur que nous allons chercher. Notons alors :

$$\mathbf{B} = E(\epsilon^b T \epsilon^b)$$

Dès lors, ce qu'on cherche à minimiser, c'est la trace (somme des variances, qu'on veut la plus petite possible) de la matrice des covariances des erreurs d'analyse :

$$\mathbf{A} = E((x^a - x^t)^T (x^a - x^t))$$

Or, dans ce cas, on cherche cette analyse sous la forme :

$$x^a = K^0 \left(\begin{pmatrix} I_n \\ H \end{pmatrix} x^t + \begin{pmatrix} \epsilon^b \\ \epsilon^0 \end{pmatrix} \right)$$

Le caractère non-biaisé de l'analyse se traduit par :

$$K^0 \begin{pmatrix} I_n \\ H \end{pmatrix} = I_{n+p}$$

Il suffit alors de minimiser le Lagrangien contenant la trace de \mathbf{A} et la contrainte ci-dessus pour obtenir le x^a optimal. Après résolution et quelques calculs matriciels en blocs, on obtient finalement :

$$x^a = (B^{-1} + {}^T H R^{-1} H)^{-1} (B^{-1} x^b + {}^T H R^{-1} y^0)$$

$$A^{-1} = B^{-1} + {}^T H R^{-1} H$$

Un résultat très important est l'équivalence de cette démarche pour obtenir l'expression de x^a avec la minimisation sur x de la fonctionnelle suivante :

$$J(x) = \frac{1}{2} {}^T (x - x^b) B^{-1} (x - x^b) + \frac{1}{2} {}^T (Hx - y^0) R^{-1} (Hx - y^0)$$

D'un point de vue vocabulaire, s'intéresser à cette fonction de coût, c'est travailler avec des méthodes dites variationnelles. Cette dénomination provient directement du calcul variationnel où la minimisation de fonctionnelles est centrale. Remarquons que la forme de J n'est pas surprenante, puisque cela revient à minimiser quadratiquement l'écart de notre analyse à l'ébauche et aux observations, tout ceci moyennant des pondérations directement posées par l'inverse des matrices B et R . Autrement dit, plus l'erreur sur l'information prise en compte est grande, moins on s'y fierait.

1.4 Forme alternative de la solution

En reprenant l'expression de x^a établie au paragraphe précédent, on arrive à écrire :

$$x^a = x^b + B^T H (H B^T H + R)^{-1} (y^0 - H x^b)$$

On introduit alors la matrice K , communément appelée matrice de gain, et le vecteur dit d'innovation d comme suit :

$$d = y^0 - H x^b$$

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

Tout ceci permet d'écrire plus simplement :

$$x^a = x^b + K d$$

Finalement, la meilleure estimation de la réalité est une correction de la prévision qui avait été faite qui prend en compte les observations, les incertitudes des instruments et celles de la prévision.

1.5 Signaux périodiques 1D

Souvent, nous avons affaire à des signaux réels que nous savons périodiques (la période sera ici notée L). Dans ce cas, il est possible de travailler en décomposition de Fourier et les inconnues du problème ne sont plus directement les valeurs en certains points mais les coefficients de cette décomposition spectrale. L'analyse et l'ébauche sont alors représentées par des vecteurs de coefficients, et le lien entre les observations et la solution réelle se fait par application d'un opérateur qui reconstitue le signal aux points observés à partir des dits coefficients. En supposant que les observations sont en s_1, \dots, s_p , ça donne :

$$X^b = X^t + \epsilon^b$$

$$y^0 = \underbrace{\begin{pmatrix} 1 & \cos(\frac{2\pi*1}{L}s_1) & \sin(\frac{2\pi*1}{L}s_1) & \cos(\frac{2\pi*2}{L}s_1) & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots \\ 1 & \cos(\frac{2\pi}{L}s_p) & \sin(\frac{2\pi}{L}s_p) & \cos(\frac{2\pi*2}{L}s_p) & \dots \end{pmatrix}}_{F_y^i} X^t + \epsilon^0$$

Le même calcul qu'au paragraphe précédent nous fournit alors le vecteur analyse des coefficients de la décomposition :

$$X^a = (B^{-1} + {}^T F_y^i R^{-1} F_y^i)^{-1} (B^{-1} X^b + {}^T F_y^i R^{-1} y^0)$$

1.6 Non-linéarité de l'opérateur H et formulation 3D-Var classique

Il est fréquent que l'opérateur d'observation ne soit pas linéaire et les résultats établis jusqu'alors deviennent inapplicables tels quels. L'idée est alors de procéder à une linéarisation au voisinage de l'ébauche. Effectivement, en linéaire tangent, on a l'approximation suivante :

$$H(x) = H(x^b) + \mathbf{H}(x - x^b)$$

Dans cette expression, \mathbf{H} est la matrice jacobienne de l'opérateur H prise au point d'ébauche x^b . Rappelons alors que nous avons défini le vecteur d'innovation $d = y^0 - H(x^b)$ et introduisons la grandeur $\delta x = x - x^b$, appelée incrément d'analyse. La fonction quadratique J présentée plus haut devient alors :

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

L'analyse $x^a = x^b + \delta x$ optimale est donc obtenue lorsque δx minimise la fonctionnelle J . Ceci est la formulation 3D-Var classique. La quatrième dimension, le temps, n'apparaît effectivement pas dans la mesure où les observations sont concentrées en un instant précis et unique, ce qui n'est plus le cas en 4D-Var.

Cette formulation rappelle les problème de type Gauss-Newton et ouvre donc la porte à une résolution numérique à l'aide d'algorithmes de descente bien connus. La minimisation fait apparaître un problème du type $AX - B = 0$ qu'on résout à l'aide d'un algorithme de gradient conjugué. Remarquons également qu'aucune évolution, implicite ou explicite, des matrices de covariances d'erreur d'ébauche n'est constatée : c'est une des faiblesses du 3D-Var.

1.7 Filtre de Kalman et impossibilité numérique

Entre deux instants i et $i + 1$, l'ébauche est construite à partir d'un modèle de prévision :

$$x_{i+1}^t = M_{(i,i+1)}(x_i^t) + \epsilon_i^m$$

Encore une fois, on va supposer que l'erreur du modèle est non-biaisée et que la matrice des covariances d'erreur est connue :

$$E(\epsilon_i^m T \epsilon_i^m) = Q_i$$

On suppose que les erreurs d'analyse et de modèle sont décorréliées :

$$E(\epsilon_i^a T \epsilon_i^m) = 0$$

Tout l'intérêt du filtre de Kalman, du moins d'un point de vue théorique, est de proposer explicitement une évolution des covariances d'erreur de prévision. Autrement dit, une évolution de la matrice B que nous n'avions pas jusqu'alors. Effectivement, en linéarisant le modèle, on peut obtenir une expression simple de l'erreur de prévision pour l'instant suivant :

$$\epsilon_{i+1}^b = \mathbf{M}_{(i,i+1)} \epsilon_i^a - \epsilon_i^m$$

La matrice des covariances d'erreur de prévision à l'instant suivant en est immédiatement déduite :

$$P_{i+1}^b = M \underbrace{E(\epsilon_i^a T \epsilon_i^a)}_{P_i^a} T M + Q_i$$

On s'aperçoit ainsi dans cette expression que l'opérateur M va être appliqué deux fois consécutivement à un vecteur de taille n . Concrètement, il faudrait donc être capable d'effectuer $2n$ intégrations du modèle entre deux instants, ce qui est extrêmement coûteux (ordre de grandeur de $n : 10^9$). Cette étape devra donc être ignorée, ou abordée différemment (cf. EnKF).

A un instant i donné, les autres étapes reprennent point par point les résultats théoriques que nous avons établis :

- Calcul du vecteur d'innovation d_i
- Calcul de la matrice de gain K_i
- Détermination de l'analyse $x_i^a = x_i^b + K_i d_i$
- Mise à jour des covariances d'erreur d'estimation $P_i^a = P_i^b - K_i^T H_i P_i^b$
- Evolution du vecteur d'état $x_{i+1}^b = M_{(i,i+1)}(x_i^b)$

2 Le 4D-Var

2.1 La formulation 4D-Var incrémentale classique

L'étude se fait encore ici en correction, l'analyse prend donc à nouveau la forme suivante :

$$x^a = x^b + \delta x$$

Dorénavant, l'incrément d'analyse optimal est obtenu après minimisation d'une fonctionnelle qui fait intervenir des observations postérieures. Il faut donc faire appel aux liens mathématiques qu'il existe entre la valeur de l'incrément à l'instant initial et celles aux autres instants de la fenêtre d'assimilation. Les notations retenues sont les suivantes :

- k : nombre d'instant pour lesquels on a des observations
- \mathbf{H}_i : opérateur d'observation linéaire tangent à l'instant i
- \mathbf{M}_i : modèle de prévision (linéaire tangent) jusqu'à l'instant i
- B : matrice de covariances d'erreur d'ébauche
- R_i : matrice de covariances d'erreur d'observation à l'instant i
- $d_i = y_i^0 - H_i(M_i(x^b))$: vecteur d'innovation à l'instant i (ébauche intégrée)
- δx_i : incrément d'analyse à l'instant i

En 4D-Var, la fonction coût J qu'on minimise est la suivante :

$$J(\delta x) = \frac{1}{2} T(\delta x) B^{-1}(\delta x) + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \delta x_i - d_i) R^{-1}(\mathbf{H}_i \delta x_i - d_i)$$

C'est un problème sous contrainte forte : l'incrément à l'instant i doit être rigoureusement égal à ce qu'on obtient par intégration du modèle depuis l'instant initial. Le 4D-Var suppose donc que le modèle est exact. Pour intégrer cette contrainte dans l'expression de la fonctionnelle, on va considérer qu'utiliser une forme linéarisée du modèle d'évolution est valide puisque nos périodes d'assimilation ne sont relativement pas longues (6h). On obtient alors un problème de minimisation sans contrainte bien plus simple à résoudre :

$$\delta x_i = \mathbf{M}_i \delta x$$

Ce qui pour la fonctionnelle à minimiser et son gradient nous donne :

$$J(\delta x) = \frac{1}{2} {}^T(\delta x) B^{-1}(\delta x) + \frac{1}{2} \sum_{i=1}^k {}^T(\mathbf{H}_i \mathbf{M}_i \delta x - d_i) R^{-1}(\mathbf{H}_i \mathbf{M}_i \delta x - d_i)$$
$$\nabla_{\delta x} J(\delta x) = B^{-1}(\delta x) + \sum_{i=1}^k {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1}(\mathbf{H}_i \mathbf{M}_i \delta x - d_i)$$

Quelques remarques sur le 4D-Var :

- il faut attendre que toutes les observations de la fenêtre d'assimilation soient disponibles avant de lancer la minimisation
- une fois l'analyse réalisée, il suffit d'appliquer le modèle de prévision, initialisé avec le x^a fraîchement obtenu
- il nécessite les adjoints du modèle et de l'opérateur d'observation

2.2 Evolution implicite des covariances d'erreur d'ébauche

Dans beaucoup de formes de l'assimilation de données, la matrice des covariances d'erreur d'ébauche est indépendante du temps. Nous l'avons par exemple constaté avec le 3D-Var. En 4D-Var, on vient de voir qu'apparaissait dans la fonctionnelle une ébauche explicitement évoluée grâce au modèle. Il est alors communément admis et cru qu'une évolution similaire de la matrice B a lieu implicitement lors de la résolution. A priori, il est difficile d'en fournir une preuve rigoureuse dans le cas général. Pour nous convaincre qu'il s'agit cependant d'une conjecture raisonnable et fondée, nous proposons ici des calculs qui resteront généraux jusqu'à qu'il ne soit pas possible d'aller plus loin, puis nous nous placerons alors dans un cas particulier pour terminer la démonstration.

Dans le chapitre sur le Filtre de Kalman, nous avons fait un calcul permettant d'aboutir à une expression des covariances d'erreur d'ébauche évoluées. Pour cela, on avait considéré que l'erreur future se calculait à partir de l'analyse fraîchement obtenue et de la réalité, ces dernières évoluées par le modèle. C'était possible car le filtre de Kalman est un schéma séquentiel qui fournit une nouvelle analyse à chaque utilisation. En 4D-Var, les instants futurs sont regardés avant que le calcul d'analyse soit fait, donc on considère cette fois que l'erreur est directement donnée par l'écart entre l'ébauche et la réalité à l'instant initial, évoluées grâce au modèle :

$$\epsilon_i^b = x_i^b - x_i^t = \mathbf{M}_i x^b - \mathbf{M}_i x^t = \mathbf{M}_i (x^b - x^t)$$

On rappelle qu'avant évolution, on a :

$$B = E((x^b - x^t)^T (x^b - x^t))$$

La matrice des covariances d'erreur d'ébauche évoluée jusqu'au temps i est donc donnée par :

$$B^i = E((x_i^b - x_i^t)^T (x_i^b - x_i^t)) = E(\mathbf{M}_i (x^b - x^t)^T (x^b - x^t)^T \mathbf{M}_i) = \mathbf{M}_i B^T \mathbf{M}_i$$

Dans le cadre de ce calcul, nous allons manipuler le système matriciel $\nabla J = 0$ d'une manière qui ne reflète en rien ce qui est fait en réalité pour le résoudre, ce sera juste pour mettre en évidence les propriétés qui nous intéressent. Rappelons donc que nous avons :

$$\begin{aligned} B^{-1} \delta x + \sum_{i=1}^k ({}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} \mathbf{H}_i \mathbf{M}_i \delta x - {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} d_i) &= 0 \\ \Leftrightarrow (B^{-1} + \sum_{i=1}^k {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} \mathbf{H}_i \mathbf{M}_i) \delta x &= \sum_{i=1}^k \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} d_i \\ \Leftrightarrow (I_n + \sum_{i=1}^k B {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} \mathbf{H}_i \mathbf{M}_i) \delta x &= \sum_{i=1}^k B \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} d_i \end{aligned}$$

C'est à ce stade que nous devons faire une hypothèse supplémentaire pour faire apparaître dans le calcul la propriété recherchée. Effectivement, si on suppose qu'il n'existe qu'une date postérieure i_0 pour lesquelles nous avons des observations, alors notre équation devient :

$$(I_n + B^T \mathbf{M}_{i_0}^T \mathbf{H}_{i_0} R^{-1} \mathbf{H}_{i_0} \mathbf{M}_{i_0}) \delta x = B \mathbf{M}_{i_0}^T \mathbf{H}_{i_0} R^{-1} d_{i_0}$$

On applique alors à gauche la matrice M_{i_0} et on ré-arrange pour obtenir finalement :

$$\delta x = (M_{i_0} + B^{i_0} {}^T \mathbf{H}_{i_0} R^{-1} \mathbf{H}_{i_0} \mathbf{M}_{i_0})^{-1} B^{i_0} {}^T \mathbf{H}_{i_0} R^{-1} d_{i_0}$$

Ce qu'on observe donc bien, c'est que la résolution numérique du problème de minimisation tendra vers une solution dont l'expression est fonction de la matrice des covariances d'erreur d'ébauche évoluée jusqu'au temps i_0 . C'est en cela que la propagation est implicite. Resterait donc à élargir la démonstration au cas où les informations postérieures sont multiples, mais au vue de tout ceci, on comprend mieux pourquoi il a été conjecturé que cela va rester vrai : la formulation 4D-Var, même dans sa forme la plus générale, propose implicitement une évolution de B par le modèle.

2.3 Conditionnement et 4D-Var incrémental pré-conditionné

La rapidité de convergence de l'algorithme de gradient conjugué utilisé pour déterminer δx est directement liée au conditionnement de la matrice Hessienne de la fonction coût :

$$\mathcal{H} = B^{-1} + \sum_{i=1}^k {}^T\mathbf{M}_i {}^T\mathbf{H}_i R^{-1} \mathbf{H}_i \mathbf{M}_i$$

Le conditionnement est le rapport, en valeur absolue, des plus grande et plus petite valeurs propres de cette matrice. Cette grandeur s'avère importante en calcul numérique car elle permet d'une part d'évaluer la sensibilité du résultat aux éventuelles perturbations en entrée du système (ce qui devient donc un critère de fiabilité sur ce qu'on obtient), mais c'est surtout une mesure de l'ellipticité des iso-surfaces de la fonction coût : plus ce coefficient est grand, plus la convergence va être difficile car la forte ellipticité se traduit par des fonctions aux courbes peu marquées ce qui rend la recherche du minimum plus compliquée. Ainsi, si nous décidons de faire tourner l'algorithme avec le problème posé tel quel, des problèmes surviendront car le conditionnement est naturellement mauvais avec cette première formulation du 4D-Var. Effectivement, les hypothèses d'homogénéité et d'isotropie usuelles pour B permettent de faire un lien direct entre ses valeurs propres et les coefficients des décompositions de Fourier discrètes des fonctions de covariance que ses lignes contiennent. Or, ces dernières ont l'allure générale de large gaussiennes. Il s'agit donc finalement de signaux basses fréquences, ce qui se traduit par des longueurs d'ondes fortes en basses fréquences et casi nulles en hautes fréquences : le conditionnement est très grand et peut même tendre vers l'infini si les amplitudes tendent vers 0 en hautes fréquences. Il faut donc se tourner vers un changement de variable pour travailler sur un problème équivalent mais qui fera intervenir une Hessienne ayant un meilleur conditionnement. Courtier [1] est le premier à proposer le préconditionnement par la matrice des covariances d'erreur d'ébauche :

$$\delta x = B^{\frac{1}{2}} \chi$$

Dès lors, on a :

$$\begin{aligned} J(\chi) &= \frac{1}{2} {}^T\chi\chi + \frac{1}{2} \sum_{i=1}^k {}^T(\mathbf{H}_i \mathbf{M}_i B^{\frac{1}{2}} \chi - d_i) R^{-1} (\mathbf{H}_i \mathbf{M}_i B^{\frac{1}{2}} \chi - d_i) \\ \nabla_{\chi} J(\chi) &= \chi + \sum_{i=1}^k {}^T B^{\frac{1}{2}} {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} (\mathbf{H}_i \mathbf{M}_i B^{\frac{1}{2}} \chi - d_i) \\ \tilde{\mathcal{H}} &= I + \sum_{i=1}^k {}^T B^{\frac{1}{2}} {}^T \mathbf{M}_i {}^T \mathbf{H}_i R^{-1} \mathbf{H}_i \mathbf{M}_i B^{\frac{1}{2}} \end{aligned}$$

Ce changement de variable est avantageux pour plusieurs raisons :

- Il conserve le caractère défini-positif de la matrice Hessienne :

Soit x un vecteur de taille n non nul. Notons $\lambda_1, \dots, \lambda_p$ les valeurs propres strictement positives de R^{-1} , $D = \text{diag}(\lambda_i)$, $R^{-1} = {}^T P D P$, $A_i = \mathbf{H}_i \mathbf{M}_i B^{\frac{1}{2}}$ et enfin $\tilde{x}^i = P A_i x$. On a alors :

$${}^T x \tilde{\mathcal{H}} x = {}^T x x + {}^T (A_1 x) R^{-1} (A_1 x) + {}^T (A_2 x) R^{-1} (A_2 x) + \dots = \underbrace{\|x\|^2}_{\geq 0} + {}^T \tilde{x}^1 D \tilde{x}^1 + {}^T \tilde{x}^2 D \tilde{x}^2 + \dots$$

Or :

$${}^T \tilde{x}^i D \tilde{x}^i = \underbrace{\sum_{j=1}^p (\tilde{x}_j^i)^2}_{> 0} \lambda_j$$

Donc :

$$\boxed{{}^T x \tilde{\mathcal{H}} x > 0}$$

- Il nous assure que le conditionnement va être borné :

Soit x un vecteur propre de $\tilde{\mathcal{H}}$ associé à la valeur propre λ . On a alors :

$$\lambda x = x + {}^T A_1 R^{-1} A_1 x + {}^T A_2 R^{-1} A_2 x + \dots$$

Donc :

$$(\lambda - 1)x = {}^T A_1 R^{-1} A_1 x + {}^T A_2 R^{-1} A_2 x + \dots$$

Ce qui donne, après multiplication à gauche par ${}^T x$:

$$\lambda = 1 + \frac{\sum_{i=1}^k {}^T \tilde{x}^i D \tilde{x}^i}{\|x\|^2} > 1$$

En notant c le conditionnement de la nouvelle Hessienne, on vient de s'assurer que :

$$c < \max_{1 \leq i \leq n} |\lambda_i|$$

Ce dernier est dorénavant borné et ne tendra plus vers des valeurs aussi mauvaises

qu'avec la formulation initiale.

- La majorité des valeurs propres seront proches de 1 :

C'est directement dû au fait que le nombre d'observations est bien plus petit que la taille du vecteur de contrôle. Ainsi, la matrice identité qui intervient dans l'expression de la Hessienne prévaudra souvent.

Dans l'algorithme du 4D-Var, nous avons pu constater qu'il y a une évolution implicite de la matrice B . Mais il serait possible d'aller encore plus loin grâce aux méthodes ensemblistes qui vont être présentées et expérimentées dans les paragraphes qui suivent. Entre autres avantages, elles devraient permettre de mieux gérer les évolutions de B et fournir des résultats bien plus précis pour des coûts semblables.

2.4 Le 4D-En-Var sans localisation

La particularité des formulations ensemblistes est que le calcul prévisionnel va pouvoir s'appuyer sur un jeu de plusieurs analyses faites sur une même période d'assimilation. Effectivement, l'idée va être de perturber raisonnablement les jeux d'observations et les ébauches pour obtenir différentes analyses par le procédé de minimisation. Par intégration du modèle non-linéaire, on obtiendra les différentes trajectoires correspondantes. L'intérêt va être double : d'une part nous allons pouvoir calculer statistiquement les matrices B évoluées, mais nous verrons aussi que cette nouvelle approche présente l'atout considérable de ne plus requérir la mise en place du modèle linéarisé. En supposant donc que nous ayons l trajectoires à notre disposition, les matrices B_i (covariances d'erreur d'ébauche évoluées jusqu'au temps i) vont être calculées à partir des écarts des l prévisions à leur moyenne, nous allons voir comment. Ci-dessous, l'introduction de quelques notations et des relations qui en découlent :

- $\overline{x^{(b,i)}}$: vecteur « ébauche moyenne au temps i » des valeurs moyennes des n variables estimées, calculées au temps i sur les l ébauches disponibles
- $x_{j,k}^{(b,i)}$: valeur prévue pour la j -ème variable au temps i par la k -ème trajectoire
- $\epsilon_{j,k}^i = \frac{(x_{j,k}^{(b,i)} - \overline{x_j^{(b,i)}})}{\sqrt{l-1}}$: écart de la j -ème variable, sur la k -ème trajectoire, à sa valeur moyenne au temps i
- ϵ^i : matrice $n \times l$ des $((\epsilon_{j,k}^i))$
- $B_i = \epsilon^i T \epsilon^i$: la matrice des covariances d'erreur d'ébauche manufacturée grâce à la formulation ensembliste

La chose importante à remarquer ici est que nous disposons dorénavant d'une expression explicite des B évoluées jusqu'aux temps i . Après discussion sur ce qui va jouer le rôle de $B^{\frac{1}{2}}$, nous les injecterons directement dans notre fonctionnelle pour deux raisons : on espère que cette mise en place empirique des covariances évoluées va améliorer la précision du résultat, et surtout, on s'affranchit du calcul très lourd de $B_i^{\frac{1}{2}} = \mathbf{M}_i B_i^{\frac{1}{2}}$ de la formulation classique du 4D-Var.

A ce stade, il est légitime de se poser la question du conditionnement et d'envisager un changement de variable pour rendre une éventuelle résolution numérique possible. Somme-nous dans la même configuration que dans le paragraphe précédent ? Pas exactement. Nous n'avons plus la garantie de trouver simplement une racine de B au sens

strict (à savoir une matrice \tilde{B} telle que $B = \tilde{B}^2$) pour définir χ . Cependant, on connaît par construction une matrice rectangulaire $n \times l$ qui une fois multipliée à droite par sa transposée donne B : c'est ϵ^i . Si on conserve la notation (dorénavant légèrement abusive) $B^{\frac{1}{2}} = \epsilon^i$, on peut alors définir sur le même principe que précédemment une nouvelle variable χ , de dimension l cette fois, vérifiant : $\delta x = B^{\frac{1}{2}}\chi$. Une étude rapide montrerait alors que les bienfaits de ce changement de variable sont exactement les mêmes que ceux constatés dans le paragraphe précédent : qu'importe si la matrice « racine » est carrée ou rectangulaire tant que par produit elle redonne B et est donc à l'origine de la matrice identité qui apparaît dans la Hessienne, laquelle fait toute la différence du point de vue du conditionnement.

Finalement, en notant $\delta x_i = B_i^{\frac{1}{2}}\chi$, la fonctionnelle et le gradient s'écrivent alors :

$$J(\chi) = \frac{1}{2} {}^T\chi\chi + \frac{1}{2} \sum_{i=1}^k {}^T(\mathbf{H}_i\delta x_i - d_i)R^{-1}(\mathbf{H}_i\delta x_i - d_i)$$

$$\nabla_{\chi}J(\chi) = \chi + \sum_{i=1}^k {}^TB_i^{\frac{1}{2}} {}^T\mathbf{H}_iR^{-1}(\mathbf{H}_i\delta x_i - d_i)$$

2.5 Le 4D-En-Var avec localisation

Tel quel, le problème ensembliste n'est pas soluble. Effectivement, la matrice B_i a un rang aussi élevé que celui des matrices dont elle est le produit, à savoir un rang de seulement $l \ll n$. De plus, on aimerait trouver une formulation équivalente qui permettrait de lisser les erreurs dûes au faible nombre d'analyses utilisées pour obtenir statistiquement B . Ces deux préoccupations sont gérées simultanément par un procédé dit de localisation. On va multiplier point par point les covariances qui apparaissent dans B par des gaussiennes qui ont l'allure générale qu'on attend pour ces covariances d'erreur d'ébauche. Pour ce produit point par point (dit de Schur), on introduit une matrice de corrélation C aux covariances homogènes isotropes (matrice donc circulante) qui par transformation de Fourier devient diagonale (cf. Annexe). En notant \circ le produit de Schur, la matrice des covariances évoluée au temps i que nous allons utiliser est donc la suivante :

$$\tilde{B}_i = B_i \circ C$$

Comme précédemment, il va donc se poser la question du changement de variable pour gérer le conditionnement du problème. Nous l'avons vu, il suffit d'avoir à notre disposition une matrice qui une fois multipliée à droite par sa transposée donne \tilde{B}_i pour qu'un changement de variable efficace puisse directement être défini. Après introduction de différentes notations (qui viennent s'ajouter à celles du 2.4.) dont nous aurons besoin, nous définirons une matrice et prouverons qu'elle convient :

- les coefficients de \tilde{B}_i (de taille $n \times n$) sont notés \tilde{b}_{jk}^i
- les coefficients de B_i (de taille $n \times n$) sont notés b_{jk}^i
- les coefficients de C (de taille $n \times n$) sont notés c_{jk}
- les coefficients de $C^{\frac{1}{2}}$ (de taille $n \times n$) sont notés c_{jk}^r
- pour $1 \leq h \leq l$, on définit les matrices suivantes :

$$D_h^i = \begin{pmatrix} \epsilon_{1,h}^i & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \epsilon_{n,h}^i \end{pmatrix} \quad \tilde{D}_h^i = D_h^i \cdot C^{\frac{1}{2}}$$

- les coefficients de \tilde{D}_h^i (de taille $n \times n$) sont notés \tilde{d}_{jk}^h
- on note T la matrice $n \times n.l$ suivante $\left(\tilde{D}_1^i \quad \dots \quad \tilde{D}_l^i \right)$ de coefficients t_{jk}
- on note $\tilde{T} = T^T T$ (de taille $n \times n$) et ses coefficients \tilde{t}_{jk}

Montrons alors que $\tilde{T} = \tilde{B}^i$. Soit donc $(j, k) \in \llbracket 1; n \rrbracket^2$. On a par définition :

$$b_{jk}^i = \sum_{h=1}^l \epsilon_{jh}^i \epsilon_{kh}^i$$

On a donc :

$$\tilde{b}_{jk}^i = c_{jk} \sum_{h=1}^l \epsilon_{jh}^i \epsilon_{kh}^i$$

De plus, des définitions, il résulte aussi que :

$$\begin{aligned} \tilde{d}_{jk}^h &= c_{jk}^r \epsilon_{jh}^i \quad (1 \leq h \leq l) \\ c_{jk} &= \sum_{h=1}^n c_{jh}^r c_{kh}^r \end{aligned}$$

Calculons alors \tilde{t}_{jk} :

$$\begin{aligned} \tilde{t}_{jk} &= \sum_{h=1}^{n.l} t_{jh} t_{kh} \\ &= \sum_{h=1}^n t_{jh} t_{kh} + \dots + \sum_{h=1+(\alpha-1)n}^{\alpha n} t_{jh} t_{kh} + \dots + \sum_{h=1+(l-1)n}^{n.l} t_{jh} t_{kh} \\ &= \sum_{h=1}^n \tilde{d}_{jh}^1 \tilde{d}_{kh}^1 + \dots + \sum_{h=1}^n \tilde{d}_{jh}^\alpha \tilde{d}_{kh}^\alpha + \dots + \sum_{h=1}^n \tilde{d}_{jh}^l \tilde{d}_{kh}^l \\ &= \sum_{h=1}^n (c_{jh}^r c_{kh}^r \epsilon_{j1}^i \epsilon_{k1}^i + \dots + c_{jh}^r c_{kh}^r \epsilon_{jl}^i \epsilon_{kl}^i) \\ &= c_{jk} (\epsilon_{j1}^i \epsilon_{k1}^i + \dots + \epsilon_{jl}^i \epsilon_{kl}^i) \\ &= c_{jk} \sum_{h=1}^l \epsilon_{jh}^i \epsilon_{kh}^i \end{aligned}$$

On a donc bien :

$$\tilde{t}_{jk} = \tilde{b}_{jk}^i \text{ et donc } \tilde{T} = \tilde{B}^i$$

Finalement, la matrice T va convenir pour effectuer le changement de variable, on la notera donc dorénavant : $\tilde{B}_i^{\frac{1}{2}}$. La variable χ est donc cette fois de taille $n.l$ et vérifie toujours :

$$\delta x = B\chi \text{ et } \delta x_i = \tilde{B}_i^{\frac{1}{2}}\chi$$

Voici alors à ce stade les expressions de la fonction de coût et de son gradient :

$$J(\chi) = \frac{1}{2} {}^T\chi\chi + \frac{1}{2} \sum_{i=1}^k {}^T(\mathbf{H}_i \tilde{B}_i^{\frac{1}{2}}\chi - d_i) R_i^{-1} (\mathbf{H}_i \tilde{B}_i^{\frac{1}{2}}\chi - d_i)$$

$$\nabla_{\chi} J(\chi) = \chi + \sum_{i=1}^k {}^T\tilde{B}_i^{\frac{1}{2}} {}^T\mathbf{H}_i R_i^{-1} (\mathbf{H}_i \tilde{B}_i^{\frac{1}{2}}\chi - d_i)$$

Du coup, on se retrouve avec un vecteur de contrôle de très grande taille, ce qui n'est clairement pas un avantage. L'idée va alors être d'exploiter le fait que les fonctions de corrélation intervenant dans C ont des portées suffisamment longues pour qu'elles puissent être représentées convenablement dans leur espace spectral par un faible nombre de fréquences n_c . Si on dispose alors de la transformation S qui permet de passer de l'espace points de grille de dimension n à l'espace spectral de dimension n_c , on pourra y définir un vecteur de contrôle de taille $n_c.l$ bien moins coûteux. C'est donc ce qu'on fait :

Expressions...

2.6 Vers une nouvelle formulation du 4D-Var

2.6.1 Présentation de l'idée

La formulation variationnelle classique du 4D-Var telle que présentée au 2.1 permet d'obtenir une analyse qui concilie au mieux les contraintes explicites suivantes :

- L'analyse en début de fenêtre doit être proche de l'ébauche à cet instant. Une pondération par les incertitudes contenues dans B est appliquée.
- Aux instants où il y en a, l'analyse, transformée par l'opérateur H , doit également être proche des différentes observations. La matrice R est cette fois sollicitée.

Rappelons que l'un des forts intérêts de la formulation variationnelle est qu'il est aisément possible de rajouter des contraintes sur le résultat recherché : il suffit d'incorporer de nouveaux termes dans la fonctionnelle J . En l'occurrence, on pourrait exiger explicitement que l'analyse soit proche de l'ébauche à des instants postérieurs, et non plus uniquement en début de fenêtre. Intervindraient alors des ébauches et des matrices B évoluées à ces instants, idéalement par application du modèle non-linéaire. Si, dans un premier temps, nous choisissons ces instants comme étant tous ceux où des observations ont été faites (on notera k leur nombre), la fonctionnelle devient :

$$J(x) = \sum_{i=0}^{k-1} \left[\frac{1}{2} (M_i x - M_i x_b)^T B_i^{-1} (M_i x - M_i x_b) + \frac{1}{2} (\mathbf{H}_i M_i x - y_i)^T R^{-1} (\mathbf{H}_i M_i x - y_i) \right]$$

Typiquement, nous allons chercher à réécrire la fonctionnelle sous forme incrémentale, en posant $\delta x_i = M_i x - M_i x_b$ pour tout instant i où l'on dispose d'observations et en tolérant l'approximation $\delta x_i \simeq \mathbf{M}_i(x - x_b) \simeq \mathbf{M}_i \delta x$ où \mathbf{M}_i est le modèle linéaire-tangent en x_b :

$$J(\delta x_1, \dots, \delta x_k) = \sum_{i=0}^{k-1} \left[\frac{1}{2} (\delta x_i)^T B_i^{-1} (\delta x_i) + \frac{1}{2} (\mathbf{H}_i \delta x_i - d_i)^T R^{-1} (\mathbf{H}_i \delta x_i - d_i) \right]$$

Introduisons alors les grandeurs suivantes :

$$\delta X = \begin{pmatrix} \delta x_0 \\ \vdots \\ \delta x_{k-1} \end{pmatrix} \quad \mathcal{D} = \begin{pmatrix} d_0 \\ \vdots \\ d_{k-1} \end{pmatrix} \quad \mathcal{B} = \begin{pmatrix} B_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_{k-1} \end{pmatrix}$$

$$\mathcal{H} = \begin{pmatrix} \mathbf{H}_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{H}_{k-1} \end{pmatrix} \quad \mathcal{R} = \begin{pmatrix} R_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R_{k-1} \end{pmatrix}$$

On obtient ainsi :

$$J(\delta X) = \frac{1}{2}(\delta X)^T \mathcal{B}^{-1}(\delta X) + \frac{1}{2}(\mathcal{H}\delta X - \mathcal{D})^T \mathcal{R}^{-1}(\mathcal{H}\delta X - \mathcal{D})$$

Finalement, on vient de montrer que, du moment où $\delta x = x - x_b$, les deux problèmes suivants sont équivalents :

1. Trouver x de taille n qui minimise :

$$J(x) = \sum_{i=0}^{k-1} \left[\frac{1}{2}(M_i x - M_i x_b)^T B_i^{-1}(M_i x - M_i x_b) + \frac{1}{2}(\mathbf{H}_i M_i x - y_i)^T R^{-1}(\mathbf{H}_i M_i x - y_i) \right]$$

2. Trouver $\delta X = (\delta x_0, \dots, \delta x_{k-1})^T$ de taille $n.k$ qui, sous la contrainte $\delta x_i = M_i \delta x$, minimise :

$$J(\delta X) = \frac{1}{2}(\delta X)^T \mathcal{B}^{-1}(\delta X) + \frac{1}{2}(\mathcal{H}\delta X - \mathcal{D})^T \mathcal{R}^{-1}(\mathcal{H}\delta X - \mathcal{D})$$

C'est sur cette deuxième formulation que nous avons décidé de réfléchir. Avec en tête l'idée que nous pourrions certainement recourir à la théorie ensembliste pour construire les matrices \mathcal{B} évoluées, il semblait relativement attirant d'envisager la résolution de (2) en s'affranchissant de la contrainte, mais en récupérant par d'autres moyens ce qu'elle apporte d'important mathématiquement et physiquement. Justement, regardons de plus près les effets de $\delta x_i = M_i \delta x$:

◇ Inconvénients de la contrainte

- Alourdissement du calcul : le respect de cette contrainte passe par l'application du modèle et de son adjoint à plusieurs reprises dans la fonctionnelle. Puisque nous travaillons avec des algorithmes de descente pour la minimisation, cela signifie que chaque itération est notablement plus longue et coûteuse.
- Ensemble des résultats admissibles restreint : avoir une contrainte, c'est refuser les éventuels résultats qui ne la respectent pas. En l'occurrence, il est donc envisageable qu'il existe des analyses qui ne respectent pas la contrainte et qui minimisent encore plus la fonctionnelle et qui, de ce fait, répondent d'autant plus à nos attentes de précision.

◇ Avantages de la contrainte

- La cohérence dynamique : sans la prise en compte de la contrainte, le résultat obtenu après résolution du problème (2) n'a pas de réalité physique, dans le sens où, de façon générale, aucune trajectoire issue du modèle ne passera par l'ensemble des valeurs « analysées » aux instants étudiés. Ce n'est finalement pas ce qu'il y a de plus important dans une optique de précision du résultat par rapport à la réalité, surtout qu'un rapide post-traitement permettrait de remédier à cette lacune.
- Le « smoothing » : c'est la propriété qui fait qu'une observation qui a lieu à une date quelconque pendant la période d'assimilation se répercute sur l'incrément analysé partout ailleurs dans la fenêtre temporelle. Par exemple, l'incrément d'analyse en début de fenêtre prend en compte l'influence d'une observation en bout de fenêtre. Cette interdépendance des incréments à différentes dates, qui vient directement du lien algébrique $\delta x_i = M_i \delta x$, est une propriété forte du 4D-Var que l'on perdra en résolvant sans contrainte. Nous verrons plus loin comment nous la récupérons.
- Taille de la variable de contrôle : avoir une telle contrainte permet de ne travailler qu'avec l'incrément à l'instant initial, soit une variable de contrôle de taille n à comparer au $n.k$ du problème (2). Cela peut sembler être un argument réhibitoire, mais il n'en est rien, car comme vu, notre nouvelle formulation ouvre la porte à une résolution à moindres coûts, et a surtout un fort potentiel de parallélisation.

Finalement, la perspective de résoudre (2) sans contrainte semble acceptable à condition de trouver un moyen de conserver la propriété de « smoothing ». Il s'agit donc de modifier notre formulation de manière à inclure l'idée que si δx_i varie, δx_j doit évoluer en conséquence. Cette interdépendance n'est pas sans rappeler celle qu'essaient de matérialiser les matrices de covariances d'erreur. Dans cette logique-ci, il a alors été envisagé de recourir à des covariances d'erreur qui disposeraient d'une dimension temporelle. Finalement, cela reviendrait à remplir la matrice \mathcal{B} qui, pour le moment, n'est que diagonale par blocs. La rendre pleine créerait effectivement un lien mathématique entre les différents δx_i qui, sinon, agissent de manière indépendante dans la fonctionnelle J . Nous montrerons plus loin que les techniques ensemblistes permettent de définir proprement ces covariances spatio-temporelles tout en conservant l'idée que c'est bien l'évolution par le modèle qui lie ces grandeurs : on retrouve une certaine cohérence dynamique, et surtout la propriété de « smoothing ».

2.6.2 Formulation plus générale

Pour présenter l'idée, nous nous sommes placés dans le cas où les instants analysés qui apparaissent explicitement dans la nouvelle variable de contrôle sont uniquement ceux où des observations ont été faites. On peut élargir au cas où la variable de contrôle serait composée de ces instants, mais aussi d'autres instants où aucune mesure n'a été effectuée. Par contre, il faut bien comprendre que pour profiter des observations, les instants où elles ont lieu doivent être prévus dans la variable de contrôle. Ainsi, en notant de façon plus générale ka le nombre d'instant analysés, on a :

$$\delta X = \begin{pmatrix} \delta x_0 \\ \vdots \\ \delta x_{ka-1} \end{pmatrix} \quad \mathcal{D} = \begin{pmatrix} d_0 \\ \vdots \\ d_{ka-1} \end{pmatrix} \quad \mathcal{B} = \begin{pmatrix} B_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & B_{ka-1} \end{pmatrix}$$

$$\mathcal{H} = \begin{pmatrix} \mathbf{H}_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{H}_{ka-1} \end{pmatrix} \quad \mathcal{R} = \begin{pmatrix} R_0 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & R_{ka-1} \end{pmatrix}$$

Avec $H_i = d_i = R_i = 0$ s'il n'y a pas d'observations à l'instant correspondant à l'index i .

2.6.3 Nouvelle formulation ensembliste

Le principe de construction des matrices B ensemblistes est strictement le même que celui présenté en 2.4. La différence majeure va résider dans le fait qu'ici, nous n'allons plus avoir à utiliser une matrice $B^{\frac{1}{2}}$ localisée. Cette dernière rendait nécessaire la définition d'une nouvelle variable de contrôle autant de fois plus grande que l'originale qu'il y a de membres dans l'ensemble. Il s'agit là de la faiblesse majeure de la formulation ensembliste préconditionnée manuellement par $B^{\frac{1}{2}}$, et c'est aussi la raison pour laquelle nous l'avons temporairement abandonnée.

Le préconditionnement aura cette fois lieu directement dans l'algorithme de gradient conjugué : *preconditioned conjugate gradient method*. Nous n'aurons alors à utiliser ni les matrices B^{-1} , ni les matrices $B^{\frac{1}{2}}$, et ceci pour un résultat parfaitement équivalent (cf. Annexe).

Reste à signaler que, même dans notre nouvelle formulation, c'est bien les techniques ensemblistes qui permettent de s'affranchir des calculs du modèle et de son adjoint. En effet, l'ensemble des trajectoires perturbées est ce qui permet de définir les covariances spatio-temporelles évoquées plus haut qui rendent tolérable et quasi-indécelable l'abandon de la contrainte $\delta x_i = M_i \delta x$.

La multiplication par B_i

Il s'agit simplement ici de mettre en évidence le résultat sur le procédé de localisation qui va nous être utile pour multiplier efficacement à gauche par nos matrices ensemblistes (par efficace, on entend à la fois rapide, et aussi, qui ne nécessite pas le stockage de matrices $n \times n$ entières). On se place dans le cas où la matrice de localisation est circulante, avec donc les propriétés qui en découlent.

On note l la taille de l'ensemble, U la matrice qui opère la TFD, C la matrice des covariances utilisée pour le produit de Schur, c sa ligne circulante, \hat{c} sa TFD, \hat{C} la matrice diagonale dont la diagonale est \hat{c} , et x un vecteur quelconque de taille n . On rappelle que nous avons noté, pour $1 \leq h \leq l$, à l'instant i :

$$D_h^i = \begin{pmatrix} \epsilon_{1,h}^i & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \epsilon_{n,h}^i \end{pmatrix}$$

On a alors les égalités suivantes :

$$\begin{aligned}
B_i x &= (\epsilon_i \epsilon_i^T \circ C) x \\
&= \sum_{h=1}^l (D_h^i C D_h^i) x \\
&= \sum_{h=1}^l \left(D_h^i U \widehat{C} U^{-1} D_h^i \right) x
\end{aligned}$$

De cette manière, trois matrices diagonales interviennent, ainsi que des transformées de Fourier discrètes que nous savons calculer efficacement. Il existe certainement d'autres façon de calculer numériquement ce produit. Dans un premier temps, nous avons jugé que cette formulation avait des chances de faire partie des plus efficaces, mais il s'agit clairement d'un point important qui mérite réflexion. Ce calcul intervient effectivement plusieurs fois au coeur même de notre algorithme de minimisation. L'optimiser, ce serait directement améliorer les performances de notre nouvelle formulation. Pour le moment, d'un point de vue algorithmique, en notant y le résultat de $B_i x$, ça donne :

```

y = 0;
for h = 1 : l
    x = D_h^i x;
    x = U^{-1} x;
    x = \widehat{C} x;
    x = U x;
    x = D_h^i x;
    y = y + x;
end
return y

```

On remarque ici surtout le fait qu'à aucun moment il n'a été nécessaire de stocker une matrice $n \times n$. Seuls des vecteurs de tailles n interviennent.

NB : Par la suite, il arrivera que l'on note $B_i = B_{i,i}$.

Définition des $B_{i,j}$ et multiplication à gauche

Nous l'avons vu, la définition des covariances spatio-temporelles est le point crucial de cette nouvelle formulation. On veut conserver ce phénomène de smoothing et donc l'idée que les erreurs et les variations en un point x à la date t sont influencées par les erreurs et les variations en y à la date t' , et inversement, et ce pour tout quadruplet de type (x, t, y, t') . On décide alors de définir pour deux instants i et j , la matrice suivante :

$$B_{i,j} = \epsilon_i \epsilon_j^T \circ C = \sum_{h=1}^l D_h^i U \widehat{C} U^{-1} D_h^j$$

L’empreinte du modèle est bien là car les matrices ϵ sont construites à partir des trajectoires perturbées qui ont été intégrées par le modèle non-linéaire. En ce qui concerne la multiplication à gauche, c’est très similaire à la multiplication par B_i . On fournit tout de même l’algorithme, en notant x le vecteur à multiplier et y le résultat :

```

y = 0;
for h = 1 : l
    x = D_h^j x;
    x = U^{-1} x;
    x = \widehat{C} x;
    x = U x;
    x = D_h^i x;
    y = y + x;
end
return y

```

Nouvelle forme de \mathcal{B} et analyse de ses propriétés

Notons à nouveau ka le nombre d’instantants analysés. En rajoutant ces covariances spatio-temporelles, la matrice globale ensembliste \mathcal{B} devient :

$$\mathcal{B} = \begin{pmatrix} B_0 & B_{0,1} & \cdots & B_{0,ka-1} \\ B_{1,0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & B_{ka-2,ka-1} \\ B_{ka-1,0} & \cdots & B_{ka-1,ka-2} & B_{ka-1} \end{pmatrix}$$

Analysons ses propriétés :

◇ \mathcal{B} est symétrique :

Cela est dû à deux choses :

– Pour tout i , B_i est symétrique : $B_i' = (\epsilon_i \epsilon_i^T \circ C)' = \epsilon_i \epsilon_i^T \circ C = B_i$.
C’est nécessaire car les B_i se situent sur la diagonale.

– Pour tous i et j , $B_{i,j} = \epsilon_i \epsilon_j^T \circ C = (\epsilon_j \epsilon_i^T \circ C)' = B_{j,i}'$.

◇ \mathcal{B} est définie-positive dès que $l \geq ka$:

On montre d’abord que pour tout i , B_i l’est :

Soit x vecteur de taille n , non nul :

$$\begin{aligned}
x^T B_i x &= \sum_{h=1}^l x^T D_h^i C D_h^i x \\
&= \sum_{h=1}^l x^T D_h^i C^{\frac{1}{2}} C^{\frac{1}{2}} D_h^i x \\
&= \sum_{h=1}^l (C^{\frac{1}{2}} D_h^i x)^T C^{\frac{1}{2}} D_h^i x \\
&= \sum_{h=1}^l \|C^{\frac{1}{2}} D_h^i x\|^2
\end{aligned}$$

Aucun 0 ne figure sur la diagonale de D_h^i . En effet, il s'agit d'écart à une moyenne qui est calculée sur un ensemble de l membres tous différents, ce qui rend impossible une stricte nullité. Cela signifie que $C^{\frac{1}{2}} D_h^i$ est de rang n et donc que si x est non nul, alors $C^{\frac{1}{2}} D_h^i x$ est également non nul. Finalement :

$$x^T B_i x > 0 \text{ CQFD}$$

Vérifions maintenant la propriété sur tout \mathcal{B} . ADMIS TEMPORAIREMENT.

◇ \mathcal{B} réalise le smoothing :

Remplir \mathcal{B} met fin à l'indépendance des incréments analysés aux différentes dates en les liant mathématiquement. Cela a le résultat attendu, nous fournirons plus loin des graphiques qui le montrent.

La multiplication par \mathcal{B}

Ce n'est qu'une combinaison des multiplications décrites plus haut. En notant $x = (x_0^T, \dots, x_{ka-1}^T)^T$ le vecteur multiplié et $y = (y_0^T, \dots, y_{ka-1}^T)^T$ le résultat de la multiplication, l'algorithme est le suivant :

```

y = 0;
for i = 0 : ka - 1
    for j = 0 : ka - 1
        y_i = y_i + B_{i,j} x_j;
    end
end
return y

```

L'algorithme de gradient conjugué préconditionné

Nous avons utilisé l'algorithme proposé par Derber et Rosati (1989). Avec les notations du problème (2), il s'écrit :

Initialisation

$$\begin{aligned}\delta X_0 &= 0 \\ g_0 &= -\mathcal{H}^T \mathcal{R}^{-1} \mathcal{D} \\ h_0 &= \mathcal{B} g_0 \\ d_{-1} &= e_{-1} = 0\end{aligned}$$

Coeur de l'algorithme

$$\begin{aligned}d_n &= -h_n + \beta_{n-1} d_{n-1} \\ e_n &= -g_n + \beta_{n-1} e_{n-1} \\ f_n &= e_n + \mathcal{H}^T \mathcal{R}^{-1} \mathcal{H} d_n \\ \alpha_n &= \frac{g_n^T h_n}{d_n^T f_n} \\ g_{n+1} &= g_n + \alpha_n f_n \\ \delta X_{n+1} &= \delta X_n + \alpha_n d_n \\ h_{n+1} &= \mathcal{B} g_{n+1} \\ \beta_n &= \frac{g_{n+1}^T h_{n+1}}{g_n^T h_n},\end{aligned}$$

L'intérêt de cet algorithme vient du fait que l'utiliser sur notre problème tel quel est équivalent à utiliser l'algorithme de gradient conjugué traditionnel sur le problème manuellement préconditionné par $\mathcal{B}^{\frac{1}{2}}$ (preuve envisagée, cf. Annexe). On sait que pour des questions de vitesse de convergence, ce préconditionnement est nécessaire dans les formulations variationnelles. Or, dans le cas de la formulation ensembliste, le recours à $\mathcal{B}^{\frac{1}{2}}$ passe par un agrandissement conséquent de la variable de contrôle qui pose des problèmes de résolution numérique. Cet algorithme nous permet de les éviter tout en profitant de ce qu'apporte d'intéressant la formulation ensembliste.

3 Une première approche : le En-1D-Var

Pour bien appréhender cette nouvelle méthode, et surtout déjà se faire une idée de l'importance de certains paramètres, nous proposons une première étude dans un cas simple. Elle se fera sur un méridien, sur lequel est défini un signal scalaire périodique. Nous reproduirons la cadre du calcul d'analyses des formes variationnelles en manufacturant des ébauches et des observations. Tout sera synchrone (la dimension du temps n'intervient pas : c'est du 1D).

3.1 Ossature du code

Sans rentrer dans le détail, on peut essayer de décrire le comportement de l'algorithme en quelques étapes :

- Définition et initialisation de différentes variables : taille du vecteur de contrôle, nombre d'observations, taille de l'ensemble, définition du domaine, caractéristiques des différentes Gaussiennes...
- Définition de la position des observations et des points apparaissant dans le vecteur de contrôle.
- Définition de la matrice des covariances d'erreur d'observation R : elle est logiquement prise diagonale.
- Définition de l'opérateur d'observation H qui par multiplication à droite par les coefficient X de la décomposition de Fourier donne les valeurs aux points d'observation.
- Définition de la matrice des covariances d'erreur d'ébauche initiale B : on utilise dans un premier temps des Gaussiennes homogènes.
- Définition de \hat{C} et de $\hat{C}^{\frac{1}{2}}$.
- Définition de l'opérateur H_c qui permet de passer de l'espace spectral de dimension n_c à l'espace point de grille de dimension n .
- Définition du signal vrai que nous prenons connu dans le cadre de ces tests : signal nul.
- Création de l'ébauche que nous allons utiliser pour l'analyse : on perturbe le signal vrai de façon cohérente avec B (cf. Annexe).
- Création du jeu d'observations que nous allons utiliser pour l'analyse : on perturbe le signal vrai de façon cohérente avec R .
- Création de l'ensemble de l ébauches par perturbation cohérente avec B de l'ébauche dont on dispose.

- Construction de la matrice ϵ (cf. 2.4).
- Calcul de l'analyse exacte selon l'expression du 1.4.
- Calcul de l'analyse en utilisant la nouvelle expression de B issue de ϵ (En-1D-Var).

3.2 Tests de sensibilité avec une matrice des covariances d'erreur d'ébauche homogène

3.2.1 Influence de la taille de l'ensemble

* Principe et résultats :

Pour mesurer la sensibilité de la taille de l'ensemble, nous définissons en discret l'équivalent de la norme 1 dans l'espace des fonctions continues. Effectivement, l'avantage de notre cas pédagogique est que nous disposons de l'analyse exacte vers laquelle devrait tendre ce que nous renvoie notre méthode. On calcule alors une moyenne sur tout l'échantillon des écarts de l'analyse obtenue par la méthode ensembliste à la valeur exacte, tout ceci en valeur absolue. Pour faciliter l'interprétation, on a fait le choix de diviser par la valeur absolue moyenne de la solution exacte, pour travailler donc en écart relatif. Pour des tailles d'ensemble ayant variées de 2 à 100, nous avons obtenu les résultats suivants :

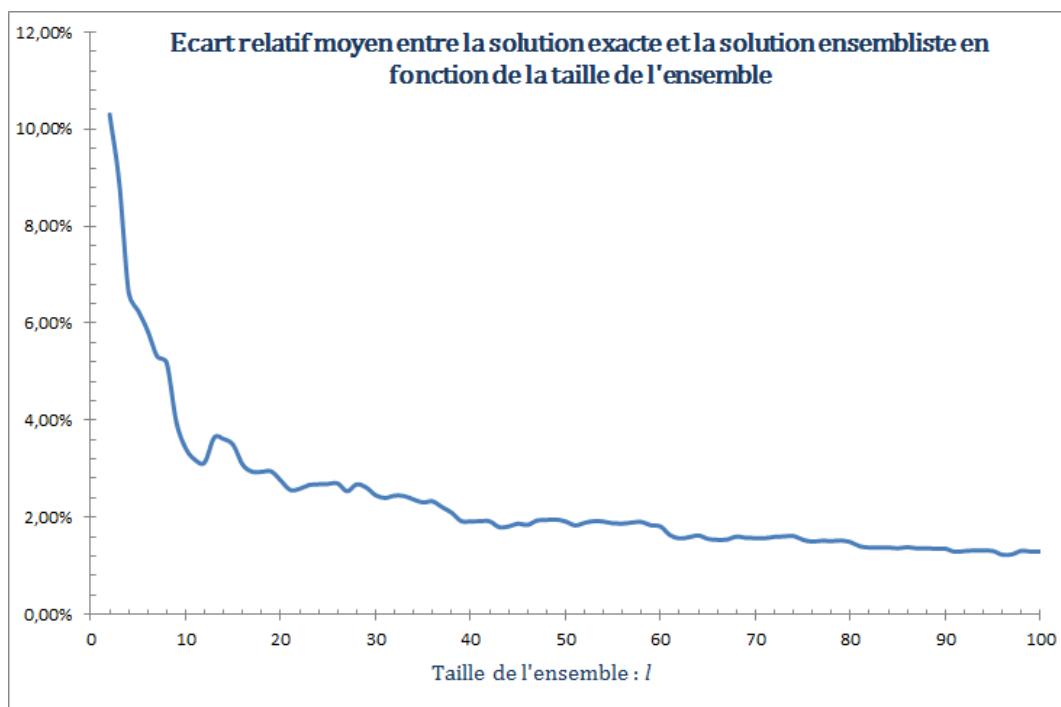


FIG. 1 – Influence de la taille l de l'ensemble

* Interprétation :

De cette courbe, on retiendra qu'avec une taille d'ensemble grandissante, l'analyse renvoyée par la méthode ensembliste tend vers l'analyse exacte, laquelle avait été obtenue en utilisant explicitement la matrice B . La matrice des covariances construite par la méthode ensembliste tend donc vers la vraie matrice B . L'intérêt de cette propriété se fera surtout sentir quand une dimension temporelle sera ajoutée à notre problème. Effec-

tivement, quand seront disponibles les trajectoires correspondant à chacune des analyses perturbées, les matrices de covariances calculées par la méthode ensembliste à des dates postérieures tendront vers les valeurs évoluées à ces dates de B , sans que l'évolution explicite de cette dernière ait à être calculée (calcul très lourd rappelons-le). Là est d'ailleurs la force de la formulation ensembliste.

3.2.2 Influence de la portée de la matrice C du produit de Schur

* Principe et résultats :

Comme dans le paragraphe précédent, l'idée est ici de faire des tests qui vont permettre de mettre en évidence l'influence de la longueur de corrélation du produit de Schur. Aussi, nous avons pris le parti de faire varier la taille de l'ensemble en même temps que cette longueur de corrélation de C : nous devrions donc aussi retrouver les conclusions du paragraphe précédent. Encore une fois, on travaille en écart relatif entre la solution renvoyée par la méthode ensemble et l'analyse exacte :

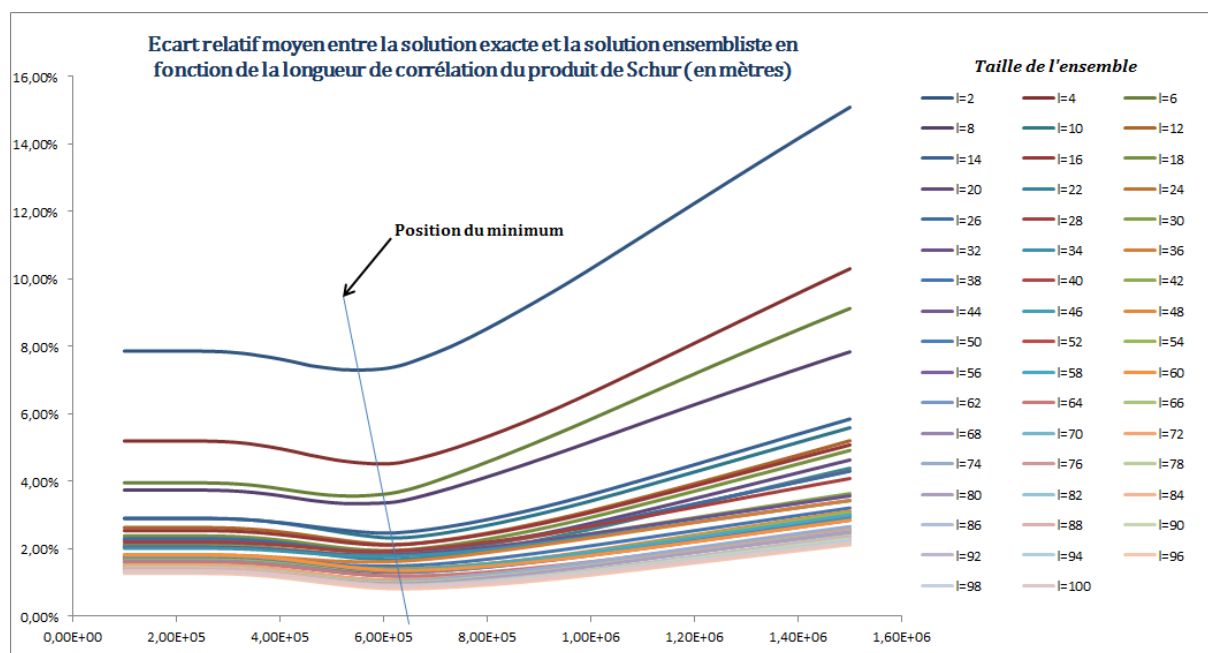


FIG. 2 – Influence de la longueur LC de corrélation de C et de la taille l de l'ensemble

* Interprétation :

On observe donc assez clairement deux choses :

- Il existe une valeur optimale de la longueur de corrélation : ni trop large ni trop restreinte.
- Cette valeur croît avec la taille de l'ensemble l :

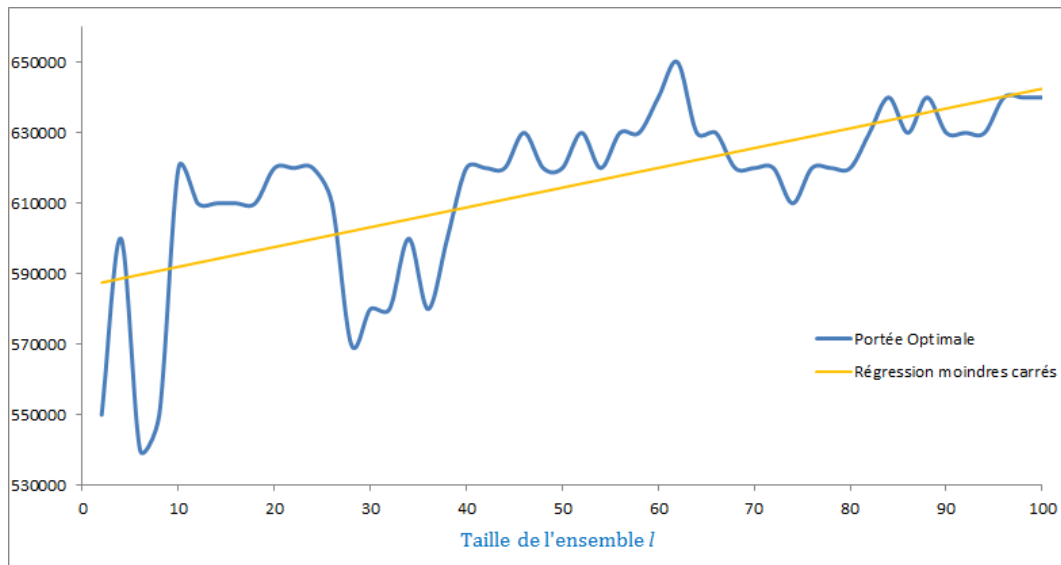


FIG. 3 – Evolution de la portée optimale en fonction de la taille de l'ensemble l

* Cohérence avec l'analyse physique :

Ces résultats semblent correct d'un point de vue physique. Effectivement :

- L'un des gros intérêts du produit de Schur par la matrice C est qu'il permet de lisser les erreurs de B dues au fait qu'elle est un résultat statistique obtenu avec peu de données. Notamment, cela se traduit par des erreurs notables loin du centre de la Gaussienne, une non-nullité qui n'a aucun sens physique (des points très éloignés devraient être décorrélés). Dès lors, avec une portée trop grande pour C , ces erreurs resteront en grande partie inchangées et ne tendront donc pas vers 0. D'où l'écart relatif grandissant pour de fortes valeurs de la longueur de corrélation.
- Une portée trop petite va à l'inverse filtrer l'information utile qui se trouve dans le centre de la Gaussienne. Cela se traduit inévitablement par une erreur relative plus élevée qu'à l'optimum, lequel se situe bel et bien, nous l'avons compris, entre ces deux extrêmes.

Annexe

Propriétés des matrices circulantes

Définition : Matrices circulantes

Une matrice circulante est une matrice carrée dans laquelle le passage d'une ligne à la suivante se fait par un décalage des coefficients vers la droite (on parle de permutation circulaire). Finalement, si on parle d'une matrice de taille n , elle est entièrement définie par les coefficients (complexes de façon générale) de sa première ligne qu'on note : c_0, \dots, c_{n-1} . Elle a donc la forme suivante :

$$C = \begin{pmatrix} c_0 & c_1 & c_2 & \dots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \dots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \dots & c_{n-3} \\ \vdots & & & \ddots & \vdots \\ c_1 & c_2 & c_3 & \dots & c_0 \end{pmatrix}$$

Proposition : Toute matrice circulante peut s'écrire comme un polynôme en une même matrice J

On s'intéresse à la matrice J suivante :

$$J = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

Dès lors, on remarque que pour toute matrice C définie comme ci-dessus, on a :

$$C = \sum_{i=0}^{n-1} c_i J^i$$

Preuve : Pour s'en convaincre, il suffit de remarquer que :

$$J^0 = I_n$$

$$J^1 = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & & & & 1 \\ 1 & 0 & 0 & \dots & 0 \end{pmatrix}$$

$$J^2 = \begin{pmatrix} 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & 0 & 0 & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & & & & & 1 \\ 1 & \ddots & & & & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \end{pmatrix}$$

$$J^{n-1} = \begin{pmatrix} 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & & & 0 \\ 0 & 1 & & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix}$$

Définition : Racine n^{eme} primaire de l'unité

La racine n^{eme} primaire de l'unité est la grandeur complexe suivante :

$$\omega = e^{\frac{i \cdot 2\pi}{n}}$$

Les puissances de ω sont les autres racines n^{eme} de l'unité (il y a cependant périodicité : $\omega^{k+n} = \omega^k$).

Proposition : J est diagonalisable dans une base construite à partir de $\bar{\omega}$

Pour k allant de 0 à $n - 1$, les vecteurs suivants constituent une base de diagonalisation de J :

$$X_k = \begin{pmatrix} 1 \\ \bar{\omega}^k \\ \bar{\omega}^{2k} \\ \vdots \\ \bar{\omega}^{(n-1)k} \end{pmatrix}$$

Preuve : On prouve donc que les X_k sont tous vecteurs propres de J associés à des valeurs propres distinctes. Soit donc $k \in \llbracket 0; n - 1 \rrbracket$:

$$JX_k = \begin{pmatrix} \bar{\omega}^k \\ \bar{\omega}^{2k} \\ \vdots \\ \bar{\omega}^{(n-1)k} \\ 1 \end{pmatrix} = \bar{\omega}^k \begin{pmatrix} 1 \\ \bar{\omega}^k \\ \bar{\omega}^{2k} \\ \vdots \\ \bar{\omega}^{(n-1)k} \end{pmatrix} = \bar{\omega}^k X_k$$

Ainsi, les X_k forment bien une base de diagonalisation de J . Il est en donc de même pour les $\tilde{X}_k = \frac{1}{\sqrt{n}} X_k$ obtenus par homothéties. Ces derniers ont l'avantage d'être à l'origine d'une matrice de passage U unitaire, i.e. dont l'adjoint est aussi son inverse. U est donc définie comme suit :

$$U = ((u_{ij})) \in \mathcal{M}_n \text{ où pour } (i, j) \in \llbracket 1; n \rrbracket^2, u_{ij} = \frac{\bar{\omega}^{(i-1)(j-1)}}{\sqrt{n}}$$

En effet, on a bien :

– Si $r \neq s$:

$$\begin{aligned}
(U U^\dagger)_{rs} &= \sum_{t=1}^n \frac{\bar{w}^{(r-1)(t-1)}}{\sqrt{n}} \frac{w^{(t-1)(s-1)}}{\sqrt{n}} \\
&= \frac{1}{n} \sum_{t=1}^n e^{-\frac{i2\pi}{n}(r-1)(t-1)} e^{\frac{i2\pi}{n}(t-1)(s-1)} \\
&= \frac{1}{n} \sum_{t=1}^n \left(e^{-\frac{i2\pi}{n}(r-s)} \right)^{t-1} \\
&= \frac{1}{n} \frac{1 - \left(e^{-\frac{i2\pi}{n}(r-s)} \right)^n}{1 - e^{-\frac{i2\pi}{n}(r-s)}} \\
&= 0
\end{aligned}$$

– Si $r = s$:

$$\begin{aligned}
(U U^\dagger)_{rs} &= \sum_{t=1}^n \frac{\bar{w}^{(r-1)(t-1)}}{\sqrt{n}} \frac{w^{(t-1)(s-1)}}{\sqrt{n}} \\
&= \frac{1}{n} \sum_{t=1}^n e^{-\frac{i2\pi}{n}(r-1)(t-1)} e^{\frac{i2\pi}{n}(t-1)(s-1)} \\
&= \frac{1}{n} \sum_{t=1}^n (1)^{t-1} \\
&= \frac{1}{n} \\
&= 1
\end{aligned}$$

Donc U est bien unitaire et diagonalise J .

Proposition : U diagonalise toute matrice circulante

Pour toute matrice C circulante, il existe une matrice diagonale D telle que :

$$C = UDU^\dagger$$

Preuve : Il suffit de rappeler que toute matrice circulante C est un polynôme en J donc tout vecteur propre de J est aussi vecteur propre de C (les valeurs propres changent par conjugué).

Proposition : Les valeurs propres d'une matrice circulante sont liées à la TFD des coefficients qui la définissent

Notons \tilde{c}_k les nombres complexes issus de la transformation de Fourier discrète des c_k . On a alors par définition :

$$\tilde{c}_k = \sum_{l=0}^{n-1} c_l e^{-i2\pi \frac{k}{n} l}$$

En notant $\lambda_0, \dots, \lambda_{n-1}$ les valeurs propres de C , on a alors :

$$\forall k \in \llbracket 0; n-1 \rrbracket, \tilde{c}_k = \lambda_k$$

Preuve : Avec les notations introduites plus haut, nous avons :

$$C = UDU^\dagger$$

On a donc :

$$c_k = C_{1,k+1} = \frac{1}{n} \sum_{l=1}^n 1 * \lambda_{l-1} \omega^{k(l-1)} = \frac{1}{n} \sum_{l=0}^{n-1} 1 * \lambda_l \omega^{kl}$$

On peut alors remarquer que les coefficients c_k sont obtenus par transformée de Fourier inverse des valeurs propres. Par injectivité de la transformée de Fourier. Finalement, on a prouvé que les valeurs propres de C étaient directement la TFD des coefficients c_k :

$$\lambda_k = \tilde{c}_k$$

Création de perturbations ϵ_0 et ϵ_b cohérentes

Dans les formulations ensemblistes, nous l'avons vu, il est nécessaire de créer artificiellement de nouveaux jeux d'observations et d'ébauches à partir de ceux dont on dispose réellement. On doit donc créer aléatoirement des ϵ_0 et ϵ_b qui vérifient :

$$E(\epsilon^0 T \epsilon^0) = R$$

$$E(\epsilon^b T \epsilon^b) = B$$

De cette façon là, on conserve le caractère non biaisé de nos grandeurs, bien qu'une discussion pourrait se faire sur le fait que deux erreurs conditionnées par B se cumulent. Quoiqu'il en soit, une idée efficace est de rechercher ces erreurs sous la forme :

$$\epsilon^0 = R^{\frac{1}{2}} \eta_p$$

$$\epsilon^b = B^{\frac{1}{2}} \eta_n$$

Dès lors :

$$E(\epsilon^0 T \epsilon^0) = R^{\frac{1}{2}} E(\eta_p T \eta_p) R^{\frac{1}{2}}$$

$$E(\epsilon^b T \epsilon^b) = B^{\frac{1}{2}} E(\eta_n T \eta_n) B^{\frac{1}{2}}$$

Ainsi, on voit bien que si η est un vecteur dont chaque composante a une variance de 1 et est décorrélée de toutes les autres alors $E(\eta T \eta) = I$ et nous avons le résultat recherché. Une bonne solution informatique est de prendre η qui suit une loi normale d'espérance nulle et de variance 1 (sous Scilab, c'est contenu dans la commande *rand('normal')*).

Equivalence du problème en points de grille et du problème spectral en 4D-Var

!! A EDITER!!

Nous l'avons vu plus haut, il est fréquent de prendre pour variable de contrôle un ensemble de coefficients spectraux. Nous avons sous-entendu que le problème de minimisation de fonctionnelle équivalent était le même moyennant le changement des matrices B (qui devient diagonale en spectral) et H (qui devient une simple transformée de Fourier inverse aux points indiqués), nous en donnons ici la preuve, avec quelques précisions supplémentaires. On travaille en incrémental, notre fonctionnelle avant passage dans l'espace spectral est donc la suivante :

$$J(\delta x) = \frac{1}{2} T(\delta x) B^{-1}(\delta x) + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \mathbf{M}_i \delta x - d_i) R^{-1}(\mathbf{H}_i \mathbf{M}_i \delta x - d_i)$$

On note U la matrice définie comme suit :

$$U = ((u_{kj})) \in \mathcal{M}_n \text{ où pour } (k, j) \in \llbracket 1; n \rrbracket^2, u_{kj} = e^{\frac{-i \cdot (k-1) \cdot (j-1) \cdot 2\pi}{n}}$$

Si x est un vecteur, alors le produit Ux est sa transformée de Fourier discrète. On vérifie donc bien que l'inverse de U est donnée par :

$$U^{-1} = ((u_{kj}^{-1})) \in \mathcal{M}_n \text{ où pour } (k, j) \in \llbracket 1; n \rrbracket^2, u_{kj}^{-1} = \frac{1}{n} e^{\frac{i \cdot (k-1) \cdot (j-1) \cdot 2\pi}{n}}$$

Ainsi, si x est un vecteur alors $U^{-1}x$ est sa transformée de Fourier discrète inverse. Tout ceci nous permet alors de définir facilement la nouvelle variable spectrale de notre problème :

$$\begin{aligned} \widehat{\delta x} &= U \delta x \\ \delta x &= U^{-1} \widehat{\delta x} \end{aligned}$$

Ainsi, en prenant bien garde d'utiliser dorénavant des adjoints et non plus de simples transposées dans notre fonction de coût puisque des grandeurs complexes interviennent, on obtient que :

$$J(\widehat{\delta x}) = \frac{1}{2} T \overline{\widehat{\delta x} U^{-1}} B^{-1} U^{-1} \widehat{\delta x} + \frac{1}{2} \sum_{i=1}^k T \overline{(\mathbf{H}_i \mathbf{M}_i \widehat{\delta x} - d_i)} R^{-1} (\mathbf{H}_i \mathbf{M}_i \widehat{\delta x} - d_i)$$

Attention, dorénavant, le modèle M_i est celui de l'évolution des coefficients spectraux de l'incrément, et la matrice H_i reconstitue le développement de Fourier aux points des observations. En observant alors les expressions de U et U^{-1} , on obtient alors que :

$$J(\widehat{\delta x}) = \frac{1}{2} \overline{T\widehat{\delta x}} \frac{U}{n} B^{-1} U^{-1} \widehat{\delta x} + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \mathbf{M}_i \widehat{\delta x} - d_i) R^{-1} (\mathbf{H}_i \mathbf{M}_i \widehat{\delta x} - d_i)$$

Or, comme nous l'avons détaillé [ici](#), U et U^{-1} diagonalisent les matrices circulantes. C'est le cas de B qui devient la matrice diagonale notée $\widehat{B} = UBU^{-1}$ dont les coefficients diagonaux sont ceux donnés par la transformée de Fourier discrète de la première ligne de B . Son inverse est $\widehat{B}^{-1} = UB^{-1}U^{-1}$. Finalement, en travaillant avec la variable spectrale, le problème 4D-Var se résout en minimisant la fonctionnelle suivante :

$$J(\widehat{\delta x}) = \frac{1}{2} \overline{T\widehat{\delta x}} \frac{\widehat{B}^{-1}}{n} \widehat{\delta x} + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \widehat{\delta x}_i - d_i) R^{-1} (\mathbf{H}_i \widehat{\delta x}_i - d_i)$$

Equivalence du problème en points de grille et du problème spectral en 4D-En-Var

!! A EDITER!!

La perspective de travailler en spectral est d'autant plus attirante dans le cadre du 4D-En-Var car cela nous permettrait de réduire la taille de la variable de contrôle, en nous limitant au spectre utile de la fonction de corrélation du produit de Schur. Le nombre et le coût des calculs s'en verront grandement réduits. Pour trouver le problème équivalent en variable spectrale, la démarche est la même que dans le paragraphe précédent : on cherche des liens explicites entre nos variables pour faire apparaître la nouvelle et disparaître l'ancienne. Dans l'espace physique, nous nous sommes ramenés au problème équivalent suivant :

$$J(\chi) = \frac{1}{2} T \chi \chi + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \tilde{B}_i^{\frac{1}{2}} \chi - d_i) R_i^{-1} (\mathbf{H}_i \tilde{B}_i^{\frac{1}{2}} \chi - d_i)$$

Nous présentons ici un rapide calcul qui va mettre en évidence les motivations de notre changement de variable futur, et le préciser. U et U^{-1} sont les mêmes que dans le paragraphe précédent. On rappelle que χ est de taille $n.l$ sur n et on décide de le noter :

$$\chi = \begin{pmatrix} \chi_1 \\ \vdots \\ \chi_l \end{pmatrix}$$

Dès lors, un calcul très rapide permet de constater que :

$$\tilde{B}_i^{\frac{1}{2}} \chi = \sum_{k=1}^l \epsilon^{i,k} \circ \left(C^{\frac{1}{2}} \chi_k \right)$$

Or :

$$\begin{aligned} C^{\frac{1}{2}} \chi_k &= I_n C^{\frac{1}{2}} I_n \chi_k \\ &= U^{-1} U C^{\frac{1}{2}} U^{-1} U \chi_k \\ &= U^{-1} \hat{C}^{\frac{1}{2}} (U \chi_k) \end{aligned}$$

Où $\hat{C}^{\frac{1}{2}}$ est la racine de C dans la base de diagonalisation (qui est l'espace spectral, nous rappelons que dans ce cas particulier valeurs propres = coefficients de la TFD). C'est bien cette matrice qui pourrait être de taille moindre car, du fait de leurs longues

portées, ces fonctions utilisées pour la localisation sont suffisamment bien décrites par les basses fréquences seules. Si cette matrice agissait directement sur notre variable de contrôle, on pourrait alors la réduire autant que C. C'est ce qui nous incite à poser :

$$\widehat{\chi} = \begin{pmatrix} U\chi_1 \\ U\chi_2 \\ \vdots \\ U\chi_{l-1} \\ U\chi_l \end{pmatrix} = \begin{pmatrix} U & 0_{n,n} & \cdots & \cdots & 0_{n,n} \\ 0_{n,n} & U & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & U & 0_{n,n} \\ 0_{n,n} & \cdots & \cdots & 0_{n,n} & U \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \\ \vdots \\ \chi_{l-1} \\ \chi_l \end{pmatrix} = \mathbf{U}\chi$$

$$\chi = \begin{pmatrix} U^{-1}\widehat{\chi}_1 \\ U^{-1}\widehat{\chi}_2 \\ \vdots \\ U^{-1}\widehat{\chi}_{l-1} \\ U^{-1}\widehat{\chi}_l \end{pmatrix} = \begin{pmatrix} U^{-1} & 0_{n,n} & \cdots & \cdots & 0_{n,n} \\ 0_{n,n} & U^{-1} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & U^{-1} & 0_{n,n} \\ 0_{n,n} & \cdots & \cdots & 0_{n,n} & U^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\chi}_1 \\ \widehat{\chi}_2 \\ \vdots \\ \widehat{\chi}_{l-1} \\ \widehat{\chi}_l \end{pmatrix} = \mathbf{U}^{-1}\widehat{\chi}$$

En s'inspirant de notre calcul plus haut, on introduit $\widehat{B}_i^{\frac{1}{2}}$ comme suit :

$$\widehat{B}_i^{\frac{1}{2}} = \left(\begin{pmatrix} \epsilon_1^{i,1} & & \\ & \ddots & \\ & & \epsilon_n^{i,1} \end{pmatrix} U^{-1}\widehat{C}^{\frac{1}{2}} \quad \cdots \quad \cdots \quad \begin{pmatrix} \epsilon_1^{i,l} & & \\ & \ddots & \\ & & \epsilon_n^{i,l} \end{pmatrix} U^{-1}\widehat{C}^{\frac{1}{2}} \right)$$

Soit, en injectant les expressions dans la fonctionnelle :

$$\begin{aligned} J(\widehat{\chi}) &= \frac{1}{2} T\widehat{\chi} \overline{\mathbf{U}^{-1}\mathbf{U}^{-1}\widehat{\chi}} + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \widehat{B}_i^{\frac{1}{2}} \widehat{\chi} - d_i) R_i^{-1} (\mathbf{H}_i \widehat{B}_i^{\frac{1}{2}} \widehat{\chi} - d_i) \\ &= \frac{1}{2} \frac{1}{n} T\widehat{\chi} \widehat{\chi} + \frac{1}{2} \sum_{i=1}^k T(\mathbf{H}_i \widehat{B}_i^{\frac{1}{2}} \widehat{\chi} - d_i) R_i^{-1} (\mathbf{H}_i \widehat{B}_i^{\frac{1}{2}} \widehat{\chi} - d_i) \end{aligned}$$

Références

- [1] Philippe Courtier. Dual formulation of four-dimensional variational assimilation. Q. J. R. Meteorol. Soc. 123, 1997.
- [2] Gérald Desroziers. Assimilation de données. ENM - CNRM, 115.
- [3] Gérald Desroziers. Introduction au 4d-en-var. ENM - CNRM, 6.
- [4] Emilie Neveu. Application des méthodes multigrilles à l'assimilation variationnelle de données en géophysique. Thèse Mathématiques appliquées, 162, 2006.