

Introduction aux principes et méthodes de l'assimilation de données en géophysique

Notes de cours du
Master M2 OACOS & WAPE
et de
l'École Nationale Supérieure des Techniques Avancées ParisTech
et l'École des Ponts ParisTech
Révision 1.31

Marc Bocquet
CEREA, École des Ponts ParisTech



2004 - 2014
14 janvier 2014

Table des matières

I	Les méthodes de l'assimilation de données	7
1	Interpolation statistique	11
1.1	Livres et notes de références	11
1.2	Introduction	11
1.2.1	La représentation du système physique	12
1.2.2	Le système observationnel	13
1.2.3	La modélisation des erreurs	15
1.2.4	Le problème de l'estimation	18
1.3	Interpolation statistique	19
1.3.1	Ansatz pour l'estimation	19
1.3.2	Estimation optimale : l'analyse BLUE	20
1.3.3	Propriétés	21
1.4	Équivalence variationnelle	23
1.4.1	Équivalence avec BLUE	23
1.4.2	Propriétés de l'approche variationnelle	24
1.4.3	Cas où l'opérateur d'observation H est non-linéaire	24
1.4.4	Formalisme dual	25
1.5	Un exemple minimaliste	25
1.5.1	Équation d'observation et matrices de covariance d'erreur	26
1.5.2	Analyse optimale	26
1.5.3	Erreur commise	27
2	Interpolation séquentielle :	
	Le filtre de Kalman	29
2.1	Modélisation stochastique du système	30
2.1.1	Étape de l'analyse	31
2.1.2	Étape de la prévision	31
2.2	Résumé, étude de cas limites et exemple	32
2.2.1	Pas d'observation	32
2.2.2	Observations parfaites	32
2.2.3	Un exemple très simple	33
2.2.4	Deuxième exemple : assimilation d'un oscillateur	35
2.3	Le filtre de Kalman étendu	36
2.3.1	Linéarisations de la prévision et de l'analyse	36
2.3.2	Résumé	37
2.3.3	Assimilation des observations sur un oscillateur non-linéaire	37

2.4	Compléments :	39
2.4.1	Théorème d'indépendance des innovations	39
2.4.2	Orthogonalité entre l'estimé de l'analyse et l'erreur d'analyse	40
2.4.3	Propagation de l'erreur d'analyse	41
3	Formalisme variationnel généralisé : le 4D-Var	43
3.1	Fonction de coût et calcul de son gradient	44
3.1.1	Gradient avec un terme d'ébauche	46
3.1.2	Cas où le modèle direct et/ou l'observation sont non-linéaires	46
3.2	Solution du problème variationnel	47
3.2.1	Solution par rapport à \mathbf{x}_0	47
3.2.2	Solution par rapport à \mathbf{x}_k	47
3.3	Propriétés	48
3.3.1	Propagation de l'erreur d'analyse	48
3.3.2	Transférabilité de l'optimalité	49
3.3.3	Équivalence entre le 4D-Var et le filtre de Kalman	50
3.4	Calculs de sensibilité	50
3.4.1	Sensibilité du premier ordre	51
3.4.2	Sensibilité du second ordre	51
3.5	Algorithmes de minimisation de la fonction de coût	52
3.5.1	Algorithmes de descente	52
3.5.2	Algorithmes quasi-Newton	53
3.5.3	Une méthode quasi-Newton : l'algorithme BFGS	55
3.6	Contrainte de modèle faible	56
3.7	Méthode incrémentale	56
II	Assimilation de données avancée	57
4	Assimilation de données non-linéaire	61
4.1	Les limites du filtre de Kalman étendu	61
4.2	Interprétation probabiliste et interpolation optimale	62
4.2.1	Assimilation d'une observation	62
4.2.2	Théorie de l'estimation et analyse BLUE	63
4.2.3	Choix d'un estimateur	64
4.3	Interprétation probabiliste et assimilation séquentielle	65
4.3.1	Étape de la prévision	65
4.3.2	Étape de l'analyse	66
4.3.3	Théorie de l'estimation et filtre de Kalman linéaire	67
4.4	Interprétation probabiliste et assimilation variationnelle	67
4.5	Filtres particuliers (Monte Carlo)	69
5	Filtres : mieux prendre en compte la non-linéarité	71
5.1	Le filtre de Kalman d'ensemble	71
5.2	Le filtre de Kalman "unscented"	74
5.2.1	La transformation unscented	75
5.2.2	Le filtre de Kalman "unscented"	80

6	Appréhender la grande taille des systèmes : filtres réduits	83
6.1	Le filtre RRSQRT	83
6.2	Le filtre SEEK	85
6.2.1	Filtre sans erreur modèle	85
6.2.2	Filtre avec erreur modèle	87
6.3	Le filtre SEIK	88
6.3.1	Filtre sans erreur modèle	88
6.3.2	Filtre avec erreur modèle	92
6.4	Le filtre UKF racine (SRUKF)	92
III	Quelques problèmes et solutions	95
7	Quelques problèmes et solutions	97
7.1	Linéarisation, boucle interne et méthode incrémentale	97
7.2	Sous-estimation des erreurs et inflation	97
7.3	Sous-échantillonnage et localisation	97
IV	Modélisation inverse	99
8	La modélisation inverse en géophysique : exemple de l'assimilation de radiances et concepts associés	103
8.1	Le sondage de l'atmosphère : la physique	103
8.1.1	Profil vertical de température dans l'atmosphère	104
8.1.2	Processus radiatifs dans l'atmosphère et définition des radiances	104
8.1.3	Quantification du rayonnement et rayonnement du corps noir	105
8.1.4	Loi de Beer-Lambert et loi de Kirchoff	106
8.1.5	La diffusion	107
8.1.6	Équation du transfert radiatif pour le rayonnement infrarouge	107
8.1.7	Définition de la transmittance	108
8.1.8	Fonctions de poids	108
8.2	Un problème inverse typique	109
8.2.1	L'équation du problème inverse	109
8.2.2	Conditionnement du problème - Régularisation	110
8.3	Solution standard et conséquences	111
8.4	Quantifier l'information	114
8.4.1	Nombre de degrés de libertés	114
8.4.2	Résolution	116
8.4.3	Information au sens de Shannon - Entropie	116
8.5	Autres approches	118
8.5.1	Méthode de Backus-Gilbert	118
8.5.2	Méthode du maximum d'entropie	118

Première partie

Les méthodes de l'assimilation de
données

On introduit dans cette partie les outils statistiques qui servent de fondement mathématique aux méthodes de l'assimilation de données. Ils se rattachent à la *théorie de l'estimation* et à la théorie du *contrôle optimal*. On commencera par introduire les concepts de base et les objectifs poursuivis qui définissent la problématique de l'*interpolation statistique*. Sous certaines conditions, l'interpolation statistique devient interpolation *optimale* qui porte le nom de *BLUE* (Best Linear Unbiased Estimator). L'analyse BLUE se révèle équivalente à un problème variationnel, qui consiste en la minimisation d'une fonctionnelle de type moindres carrés.

Cette approche trouve dans un cadre séquentiel (par exemple un problème d'assimilation qui incorpore le temps) une extension baptisée *filtre de Kalman*. Sous certaines conditions on peut aussi faire correspondre à cette analyse un problème variationnel, où il s'agit de minimiser une fonctionnelle définie sur un espace à quatre dimensions (3 d'espace et 1 de temps). Cette technique d'assimilation, appelée *4D-Var* est plus généralement devenue un outil d'optimisation de choix.

Chapitre 1

Interpolation statistique

1.1 Livres et notes de références

Un des pères de l'assimilation de données moderne était Roger Daley, qui a notamment écrit [4]. Quelques notes de cours et articles très clairs, écrits par des chercheurs ayant fortement contribué dans le domaine permettent de bien commencer en assimilation de données [19, 1]. Des notes de cours/conférences très détaillées disponibles en ligne sur la toile peuvent aider (notamment [23]). Un des ouvrages les plus récents sur l'assimilation de données et la prédictibilité en météorologie a été écrit par une chercheuse majeure du domaine (Eugenia Kalnay) [11]. Pour une clarification conceptuelle nécessaire sur la notion d'erreur en assimilation de données, on pourra consulter [3]. Pour une mise à plat des conventions et écritures dans le domaine, mais aussi une synthèse méthodologique, on pourra consulter [10].

1.2 Introduction

Comment fait-on de la prévision sur des systèmes complexes comme l'atmosphère ?

Prenons l'exemple concret de la température sur Paris. Supposons que l'on ait mesuré chaque jour une température moyenne sur la ville. On décide alors de prédire la température moyenne journalière pour les jours à venir. La démarche la plus simple consiste à extrapoler les températures déjà observées pour les jours à venir en se fondant soit sur la régularité de la courbe de l'historique des températures ou bien sur une base statistique événementielle. C'est ce qui se pratiquait avant l'avènement des ordinateurs, car la démarche est peu coûteuse en calcul.

Une autre démarche consiste à utiliser la connaissance fondamentale que l'on a de la dynamique du système en simulant l'évolution de la température les jours suivant. Ce qu'autorise aujourd'hui la puissance de calcul des ordinateurs. Cela nécessite d'initialiser le *modèle* avec les dernières températures relevées.

Cependant, les deux démarches trouvent rapidement leurs limites et divergent de la courbe réelle (Figure 1.1).

La démarche optimale consiste à combiner au mieux et le plus exhaustivement possible la connaissance théorique et la connaissance par l'observation de l'état du système. Dans

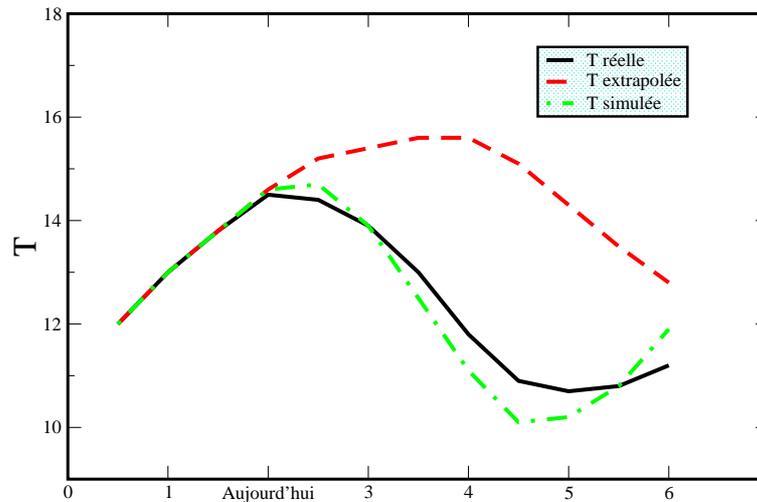


FIGURE 1.1 – Courbes des températures moyennes journalières réelles, extrapolées et simulées

notre exemple, cela consiste à initialiser la dynamique du système à l'aide de l'ensemble des observations passées et pas seulement des toutes dernières.

L'*assimilation de données* est définie comme l'ensemble des techniques statistiques qui permettent d'améliorer notre connaissance (passée, présente ou future) de l'état d'un système physique à partir de la donnée concomitante d'observations expérimentales sur le système et de sa connaissance théorique (ou a priori).

L'*interpolation statistique* est une des techniques les plus simples permettant de trouver une solution à ce problème. Si la technique est mathématiquement élémentaire (souvent équivalente aux moindres carrés), sa mise en œuvre sur des systèmes complexes de grande taille est loin d'être immédiate.

1.2.1 La représentation du système physique

On désigne par \mathbf{x} le vecteur dont les composantes sont les variables qui mesurent l'état réel d'un système. Très concrètement, on peut, par exemple, assigner à une composante du vecteur \mathbf{x} la température d'une masse d'air à une latitude, longitude et à une altitude précises, à une date et heure données.

En principe, une description exhaustive du système suppose que \mathbf{x} ne soit pas un vecteur appartenant à un espace de dimension finie, mais un champ vectoriel de composantes $x_\alpha(\overrightarrow{OM}, \tau)$. Ce champ vectoriel dépend du point spatio-temporel et ses composantes sont autant d'observables potentielles en ce point indexées par α (température, vitesse, pression, humidité, concentrations d'espèces gazeuses, etc.).

Dans une optique simulation numérique, il faut pouvoir coder informatiquement ce vecteur en le discrétisant, c'est-à-dire, formellement, en se donnant un opérateur Π qui permette de passer d'un espace de dimension infinie à un espace de dimension finie n . L'image de \mathbf{x} par Π est notée \mathbf{x}^t , et donc $\mathbf{x}^t \in \mathbb{R}^n$. L'indice t fait référence à *true*, l'état *vrai* du système. (C'est un abus de langage car en réalité il ne s'agit que d'une projection

de l'état vrai). Par exemple, \mathbf{x}^t peut être la moyenne du champ \mathbf{x} sur une cellule d'un maillage utilisé pour la représentation.

Toutefois, en pratique, au même titre que les composantes de \mathbf{x} , il est impossible de connaître la valeur exacte des composantes de \mathbf{x}^t . Et c'est précisément tout ou partie de ce vecteur qu'il s'agit d'estimer. Le résultat de l'estimation de \mathbf{x}^t est un vecteur de même dimension n noté \mathbf{x}^a , où a se réfère à *analyse*. Lorsque le résultat d'une analyse sert de point de départ à une nouvelle analyse, ce vecteur est alors noté $\mathbf{x}^b \in \mathbb{R}^n$ avec b pour *background* et est appelé en français une *ébauche*. Du point de vue d'une analyse future, elle constitue une connaissance a priori sur l'état du système.

Un système d'estimation doit se nourrir d'observations du système réel afin de réduire l'erreur commise dans l'analyse. Ces observations peuvent être entreprises en n'importe quel point du système (elles peuvent même correspondre à une moyenne locale lorsque la résolution de l'instrument est large). Sur des systèmes réalistes, il n'y a pas de raisons pour que ces points coïncident avec les points de collocation du modèle informatique représentant le système. En interpolant, extrapolant, ou filtrant, on doit être capable de fournir un jeu de données de taille p , qui puisse être mis en correspondance avec une fonction de tout où partie des composantes de \mathbf{x}^t . On notera $\mathbf{y} \in \mathbb{R}^p$ le vecteur des observations.

1.2.2 Le système observationnel

Les moyens d'observation de l'atmosphère

Un système d'assimilation complexe (donc intéressant !) doit être alimenté par de nombreuses données d'observations. On donne ici un aperçu du *système observationnel global* tel qu'il est utilisé en météorologie et en chimie atmosphérique.

Aujourd'hui, il y a par nature essentiellement deux types d'observations. Il y a d'abord les observations conventionnelles. Elles sont désormais beaucoup moins abondantes que les données satellitaires. Cependant elles sont moins redondantes, et sont mieux maîtrisées. De plus plus elles fournissent une information directe sur le système par opposition aux informations d'origine satellitaires (radiances). En 2003, 99 % des données utilisées par ECMWF (European Centre for Medium-Range Weather Forecast) sont d'origine satellitaire mais *seulement* 91.5 % des données effectivement assimilées sont d'origine satellitaires ! On compte parmi elles

- les mesures de stations (réseau SYNOP).
- les mesures faites par les avions sur les lignes commerciales.
- les mesures faites par les réseaux de bouées océaniques et les mesures faites par les bateaux marchands.
- les mesures régulières faites par ballon sonde (radiosondage).
- les mesures LIDAR (sondage actif par faisceau laser depuis le sol).
- Des données radar, permettant la localisation des précipitations.

Il y a ensuite les mesures satellitaires. Elles ne sont disponibles que depuis les années 70. Depuis la fin des années 70 et jusqu'aujourd'hui, le nombre total de mesures acquises par "round" d'observations (toutes les six heures pour la météorologie) croît sur un rythme exponentiel. En 2003, à l'ECMWF, on recueille 70×10^6 données par jour, mais on n'assimile que 3.6×10^6 données ! En 2010, à l'ECMWF, on assimile 25×10^6 observations par jours.

Du point de l'assimilation de données, le satellite a permis l'obtention d'une masse imposante de données. Il a aussi permis d'obtenir des données pour l'hémisphère sud,

juqu'alors très mal couvert par le réseau de surveillance. Des progrès très net ont ainsi été réalisés dans les prévision météorologiques lorsqu'on a pu assimiler des observations en quantité pour l'hémisphère sud.

Les satellites d'observations de la Terre

Il existe de nombreux satellites d'observation de la Terre (atmosphère, sol et océan). À leur bord, de nombreux instruments à tâche spécifique. Ces instruments sont principalement des sondes du spectre lumineux (infrarouges, visibles et UV proches). Chacun possède une batterie de canaux (de fréquences) qui fournissent autant d'informations différenciées.

Selon leur orbite, on les classe en les *satellites géostationnaires* et les *satellites de basse altitude*. Les satellites géostationnaires évoluent à 36.000 km d'altitude. Ils sont géosynchrones, c'est à dire qu'ils tournent à la vitesse angulaire de la Terre et restent donc au dessus d'un même point géographique. Leurs instruments balayent une zone circulaire autour de ce point. Il présentent donc l'avantage évident de suivre en permanence une même zone. Il sont répartis autour de l'équateur. Les satellites de basse altitude sont situés de 400 à 800 km d'altitude. Ils ont généralement une orbite polaire (passant par les pôles). Ils balayent donc une grande partie du globe. Cependant la fréquence d'observation d'une cible est moindre et parfois irrégulière.

On peut aussi classer les satellites environnementaux selon leur fonction. Il sont destinés soit à la surveillance opérationnelle, soit destinés à la recherche (les observations sont alors intermittentes).

En météorologie, il existe deux réseaux de satellites imageurs et recueillant les données météo de base (albedo, humidité de l'air, température, images des nuages) : la constellation GOES, et la constellation METEOSAT. GOES, pour Geostationary Operational Environmental Satellite, sont des satellites de la NASA (National Aeronautics and Space Agency) et de la NOAA (National Oceanic and Atmospheric Agency). METEOSAT est la contrepartie européenne, dont la constellation est gérée par l'ESA (European Space Agency).

Il existe par ailleurs des satellites dédiés à une observation environnementale plus large et plus pointue. Parmi eux, UARS (Upper Atmospheric Research Satellite de la NASA), la constellation des satellites de la NOAA : NOAA14, NOAA16, et NOAA17, les satellites EOS Terra et EOS Aqua de la NASA. Du coté européen (ESA), ERS-2 (European Remote-Sensing Satellite), et le mûr (lancé en 2002) ENVISAT, et le récent METOP-A (lancé le 19 octobre 2006).

À cela s'ajoute les satellites militaires (dont DMSP (Defense Meteorological Satellite Program). On compte également des satellites de recherche : par exemple ODIN, suédois. Plus récemment ont été lancés les satellites de l'A-Train, une constellation de 6 satellites (NASA et CNES) survolant (AQUA, AURA, CLOUDSAT, CALIPSO, PARASOL, OCO dont le lancement a échoué) l'équateur à 13h30 (orbite géosynchrone).

À cette liste, il faudrait ajouter les satellites russes et les satellites japonais (dont le satellite GOSAT mesurant le CO₂ atmosphérique).

Tous ces satellites disposent de plusieurs instruments, désignés par des acronymes. Je vais ici essayer de lister quelques uns des instruments dédiés à la composition chimique de l'atmosphère. Il est clair qu'il est plus facile de mesurer des *colonnes* (masse totale de gaz dans un cylindre de base la surface et de sommet le satellite) ou bien des concentrations dans la stratosphère, que des concentrations troposphériques.

- **GOME** (ERS-2) : (Global Ozone Monitoring Experiment). Colonnes d’ozone et NO_2 . L’instrument mesure aussi les gaz-trace et les aérosols. ERS-2 est entré en fonction en Avril 1995 et a fini sa vie en 1998.
- **HIRS-9** (sur les satellites NOAA) : (High Resolution Infrared Radiation Sounder). Le canal 9 (à 9.7 microns ; bande d’absorption du rayonnement terrestre par l’ozone dans la basse stratosphère) mesures de colonnes d’ozone. L’instrument équipe les satellites NOAA depuis 1991.
- **MOPITT** (sur EOS Terra) : (Measurements of the Pollution in the Troposphere). Mesures de la pollution dans la troposphère. En particulier, mesures de CO et CH_4 . Opérationnel depuis début 2000.
- **MIPAS** (ENVISAT) : (Michelson Interferometer for Passive Atmospheric Sounding). Mesures en chimie stratosphérique (O_3 , CH_4 , H_2O , N_2O et HNO_3)
- **GOMOS** (ENVISAT) : (Global Ozone Monitoring by Occultation of Stars), profils fins de NO_2 , O_3 . La résolution en altitude est notamment possible au dessus de la tropopause.
- **SCIAMACHY** (ENVISAT) : (Scanning Imaging Absorption Spectrometer for Atmospheric Chartography) cet instrument mesure les espèces trace dans la stratosphère et la troposphère. C’est une version notablement améliorée (spectre plus large, notamment dans l’infrarouge) de GOME. Il permet également le suivi d’aérosols.
- **IASI** (METOP) : (Infrared Atmospheric Sounding Interferometer). C’est un sondeur (CNES/EUTMETSAT) dans l’infrarouge destiné à mesurer la vapeur d’eau, le CO_2 , le dioxyde d’azote et l’ozone. Il est à bord de la plate-forme METOP, en orbite quasi-polaire et qui succède dans l’esprit à ENVISAT.
- **GOME-2** (METOP) : (Global Ozone Monitoring Experiment). C’est un spectromètre destiné à mesurer les colonnes d’ozone et les profils de NO_2 , BrO , OCIO et ClO .

Ces instruments délivrent donc des colonnes de constituants de l’atmosphère. Le calcul des profils de concentrations à partir de ces informations est réalisé par *modélisation inverse*. La procédure est appelée *assimilation de radiances*. Elle fait notamment appel à une bonne connaissance *a priori* du profil (d’ébauche donc).

1.2.3 La modélisation des erreurs

Modélisation des observations et des erreurs d’observation

En pratique, sur des systèmes complexes, la connaissance de \mathbf{y} ne permet pas d’avoir accès sans erreur à la connaissance du vecteur d’état \mathbf{x}^t . C’est par exemple le cas en météorologie, en climatologie, ou bien encore en océanographie. Il existe une chaîne de transformations, exactes ou entachées d’erreurs, entre \mathbf{y} et \mathbf{x}^t . On se propose de donner un aperçu de cette chaîne.

On a déjà mentionné le fait que, généralement, seule une partie des variables d’état du système sont accessibles à l’observation. Par exemple, en océanographie, on connaît bien les mouvements de surface des océans (par altimétrie, grâce notamment à des satellites comme Topex-Poséidon) alors qu’il est évidemment très difficile d’accéder à des données marines de profondeur, qui sont pourtant nécessaires à un diagnostic de l’état du système. Cela n’est toutefois pas une source d’erreur en soi mais seulement la signature de la sous-détermination de l’état du système (puisque $p \ll n$). En pratique pour un modèle météorologique et pour un cycle d’assimilation, le vecteur \mathbf{x}^t est de taille $n = 10^8$ ($T_L799L91$ pour le modèle du

centre européen à partir de 2006) alors que le vecteur \mathbf{y} est de taille $p = 4 \times 10^6$ (données 2009) ! En 2010, le modèle du centre européen est passé à la résolution spectrale $T_L1279L91$, soit des mailles logues d'environ 16 km, et $n = 1.5 \times 10^9$, alors que l'on désormais $p = 6 \times 10^6$

De plus, il n'y a pas toujours de relation immédiate entre les variables d'état \mathbf{x}^t et les quantités observées dans \mathbf{y} . Par exemple, les satellites mesurent les radiances émises par une colonne atmosphérique. On ne peut accéder aux concentrations des espèces présentes dans la colonne (par exemple l'ozone) qu'indirectement. Les composantes de \mathbf{y} ne se rapportent à l'état réel du système que par des combinaisons linéaires ou non-linéaires d'une partie des composantes de \mathbf{x}^t . Il existe donc une fonction, pas toujours simple, reliant \mathbf{y} à \mathbf{x} .

Les différences entre les observations et l'état du sous-système ne se limite pas à la remarque ci-dessus sur la sous-détermination (qu'on peut ranger dans la catégorie des erreurs de représentativité). On peut les classer en *erreurs de mesures* et en *erreurs de représentativité*.

L'erreur la plus évidente (et en général bien identifiée dans le processus d'observation) est l'*erreur de mesure*, qui affecte directement (et en général additivement ou multiplicativement) \mathbf{y} . On souhaite construire, connaissant l'état vrai et continu du système \mathbf{x} , le vecteur des observables qui se compare à \mathbf{y} et que l'on note $h[\mathbf{x}]$. On désigne donc par h la fonction permettant cette construction. En l'absence d'erreur de mesure, on a $\mathbf{y} = h[\mathbf{x}]$ (c'est alternativement la définition de h). L'observation s'écrit alors comme la somme de $h[\mathbf{x}]$ et de l'erreur de mesure \mathbf{e}^μ :

$$\mathbf{y} = h[\mathbf{x}] + \mathbf{e}^\mu. \quad (1.1)$$

Toutefois, on travaille numériquement sur \mathbf{x}^t et non sur \mathbf{x} . On veut donc également construire, connaissant l'état vrai discret du système \mathbf{x}^t , le vecteur des observables qui se compare à \mathbf{x} et que l'on note $H[\mathbf{x}^t]$. La chaîne décrite par H , appelée *fonction d'observation*, inclut non seulement les opérations de transformation et de projection existant dans H , mais aussi des opérations d'interpolation rendues nécessaires par l'information perdue dans la projection par Π . On écrit alors

$$\begin{aligned} \mathbf{y} &= h[\mathbf{x}] + \mathbf{e}^\mu \\ &\equiv H[\mathbf{x}^t] + \mathbf{e}^r + \mathbf{e}^\mu, \end{aligned} \quad (1.2)$$

où

$$\mathbf{e}^r \equiv h[\mathbf{x}] - H[\mathbf{x}^t] = h[\mathbf{x}] - H[\Pi\mathbf{x}], \quad (1.3)$$

est l'erreur de représentativité. Ce raisonnement est représenté sur le diagramme Fig. 1.2 et peut se résumer ainsi :

$$\mathbf{y} = H[\mathbf{x}^t] + \mathbf{e}^o, \quad (1.4)$$

avec $\mathbf{e}^o \equiv \mathbf{e}^r + \mathbf{e}^\mu$. On a supposé ici h parfaitement connu. Mais il se peut, qu'une source additionnelle d'erreur soit une connaissance imparfaite de cette fonction. Dans ce cas l'erreur commise dans la modélisation de H est de type *erreur modèle*.

Statistiques d'erreur

On s'arrange pour que l'erreur d'observation soit sans biais, c'est-à-dire que $E[\mathbf{e}^o] = 0$, quitte à soustraire a posteriori un éventuel biais constaté. $[\mathbf{R}]_{ij} = E[[\mathbf{e}^o]_i [\mathbf{e}^o]_j]$ est

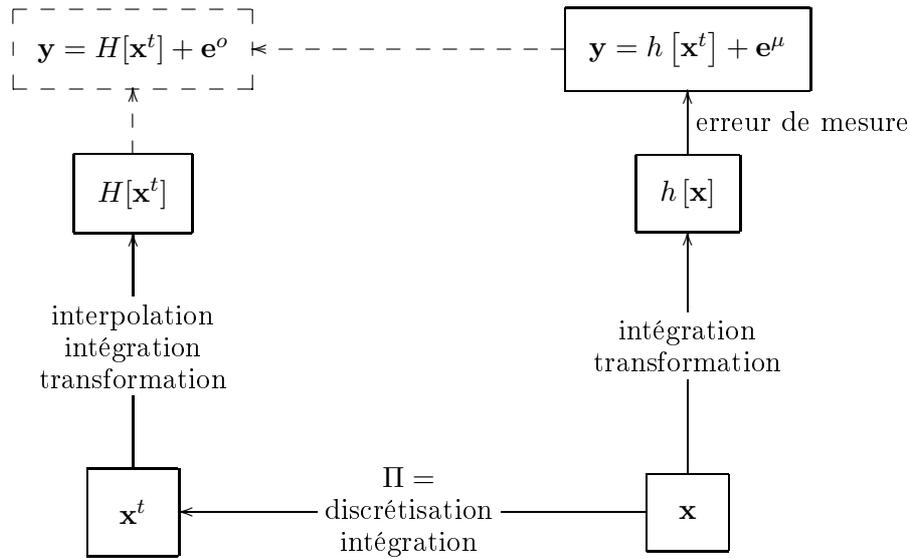


FIGURE 1.2 – Diagramme décomposant les étapes de la modélisation et les erreurs commises.

alors définie comme la matrice de covariance des erreurs d’observations. C’est une matrice symétrique de taille $\mathbb{R}^{p \times p}$. Elle est de plus supposée définie positive (donc inversible).

En général, dans la suite, on supposera connue \mathbf{R} . Mais c’est en pratique un enjeu important de s’en donner une estimation correcte, la qualité de l’analyse qui en découle en dépendant fortement. En général, on considère que \mathbf{R} est diagonale, ce qui suppose que les observations sont toutes indépendantes les unes des autres. Des sources de corrélations sont cependant possibles : corrélation temporelle, corrélation spatiale entre pixels pour les mesures satellitaires, etc.

Erreurs d’ébauche et d’analyse

L’erreur associée à l’ébauche est définie par

$$\mathbf{e}^b = \mathbf{x}^b - \mathbf{x}^t. \quad (1.5)$$

Elle mesure l’écart de l’estimation a priori (avant le début de l’analyse) à la réalité. On lui associe la matrice de covariance $[\mathbf{P}^b]_{ij} = \mathbf{E} [[\mathbf{e}^b]_i [\mathbf{e}^b]_j]$. C’est une matrice symétrique de taille $\mathbb{R}^{n \times n}$. Elle est de plus supposée définie positive (donc inversible). On suppose que l’erreur \mathbf{e}^b est sans biais : $\mathbf{E}[\mathbf{e}^b] = 0$, quitte à avoir corrigé \mathbf{x}^b pour permettre cela.

L’erreur de l’analyse est définie par

$$\mathbf{e}^a = \mathbf{x}^a - \mathbf{x}^t. \quad (1.6)$$

Elle mesure l’écart du résultat de l’analyse à la réalité. On lui associe la matrice de covariance $[\mathbf{P}^a]_{ij} = \mathbf{E} [[\mathbf{e}^a]_i [\mathbf{e}^a]_j]$. C’est une matrice symétrique de taille $\mathbb{R}^{n \times n}$.

On suppose en général que \mathbf{e}^o et \mathbf{e}^b sont décorrélées. Il n’y a pas de raison de principe pour laquelle la mesure sur un système, en tant que variable aléatoire, soit corrélée à une connaissance (imparfaite) a priori de ce même système. Cela semble une hypothèse

raisonnable, même dans des situations complexes. Mais si cette ébauche a pu être construite à l'aide d'observations de même type, voire avec les mêmes instruments, alors une telle hypothèse pourrait ne plus être satisfaite.

Erreur modèle

Dans le cadre de l'assimilation de données, on est souvent amené à utiliser un modèle Φ qui permette par exemple dans un processus séquentiel, de relier $\mathbf{x}(\tau + 1)$ à $\mathbf{x}(\tau)$, états du système aux instants $\tau + 1$ et τ :

$$\mathbf{x}(\tau + 1) = \Phi[\mathbf{x}(\tau)]. \quad (1.7)$$

Il peut s'agir par exemple des équations dites *primitives* (équations du mouvement (Newton), équation de conservation de la masse d'air, équation de conservation de la masse d'eau, équation d'état de l'air (loi des gaz parfaits), conservation de l'énergie (première loi de la thermodynamique)) pour une analyse météorologique. Il peut s'agir des équations de chimie-transport des polluants gazeux et particulaires et des équations de chimie-transport et évolution des aérosols, pour une analyse en qualité de l'air.

On suppose dans un premier temps que cette fonction de transition Φ est parfaitement connue. Elle représente donc sans erreur l'évolution réelle du système physique. Mais, une nouvelle fois, on ne peut travailler en pratique que dans un espace discret, défini par $\mathbf{x}^t = \Pi\mathbf{x}$. On projette donc l'équation de propagation Eq.(1.7) au moyen de Π

$$\begin{aligned} \mathbf{x}^t(\tau + 1) &= \Pi\mathbf{x}(\tau + 1) = \Pi\Phi[\mathbf{x}(\tau)] \\ &\equiv M[\mathbf{x}^t(\tau)] + \mathbf{e}^m, \end{aligned} \quad (1.8)$$

où *l'erreur modèle* \mathbf{e}^m est donnée par

$$\mathbf{e}^m \equiv \Pi\Phi[\mathbf{x}(\tau)] - M[\mathbf{x}^t(\tau)] = \Pi\Phi[\mathbf{x}(\tau)] - M[\Pi\mathbf{x}(\tau)]. \quad (1.9)$$

On notera que l'erreur modèle identifiée ci-dessus est une *erreur de représentativité*.

Cependant, de part la complexité des modèles utilisés et des approximations mises en œuvres pour les obtenir, la modélisation elle-même est entachée d'erreurs et l'opérateur de propagation Φ est imparfait, à plus forte raison M . Il faut donc admettre qu' \mathbf{e}^m se décompose en réalité en une erreur de représentativité telle que déjà identifiée et une erreur due à la modélisation imparfaite du modèle. Cette dernière source d'erreur est distincte de l'erreur liée à la modélisation de l'observation, bien qu'elle soit de nature similaire.

1.2.4 Le problème de l'estimation

On étudie un système décrit par un vecteur d'état \mathbf{x}^t . On suppose que l'on s'en fait une idée \mathbf{x}^b . Cette ébauche résulte d'une analyse antérieure ou d'une estimation issue de principes généraux. Elle constitue la meilleure estimation de l'état du système en l'absence de toute autre information. Des observations \mathbf{y} effectuées sur le système renseignent partiellement sur celui-ci, au travers d'une fonction d'observation H . De plus on suppose que, idéalement, on connaît les statistiques (à l'ordre deux) de l'erreur d'observation \mathbf{e}^o ainsi que de l'erreur d'ébauche \mathbf{e}^b .

On cherche alors, à l'aide des observations effectuées, à améliorer notre estimation \mathbf{x}^a de l'état du système comparée à la connaissance a priori de celui-ci \mathbf{x}^b . On cherche également à connaître l'erreur commise \mathbf{e}^a au cours de l'analyse.

Il existe bien sûr de nombreuses façons de réaliser l'analyse. Mais on cherche à réaliser une (la ?) meilleure analyse possible, c'est-à-dire une analyse permettant de minimiser l'erreur commise a posteriori \mathbf{e}^a , par exemple en minimisant $\text{Tr}(\mathbf{P}^a)$.

1.3 Interpolation statistique

1.3.1 Ansatz pour l'estimation

Hypothèse linéaire d'interpolation

On suppose maintenant que l'opérateur d'observation H est linéaire. On pourra alors le noter \mathbf{H} , ce qui indique que l'opérateur est linéaire ou bien qu'il est le linéaire tangent¹ de l'opérateur non-linéaire $H : \mathbf{H} = H'$. Si l'opérateur d'observation est l'application : $\mathbf{x} \rightarrow \mathbf{y} = H[\mathbf{x}]$, alors l'opérateur linéaire tangent associé est obtenu en développant l'opérateur autour de \mathbf{x} :

$$\delta y_i = \sum_{n=1}^N \frac{\partial H_i}{\partial x_n} \delta x_n. \quad (1.10)$$

L'opérateur linéaire tangent est alors l'opérateur linéaire dont l'action est définie par la matrice \mathbf{H} , de coefficients :

$$[\mathbf{H}]_{in} = \frac{\partial H_i}{\partial x_n}. \quad (1.11)$$

Un Ansatz pour le vecteur estimation \mathbf{x}^a est de choisir pour l'analyse un vecteur de la forme

$$\mathbf{x}^a = \mathbf{L}\mathbf{x}^b + \mathbf{K}\mathbf{y}, \quad (1.12)$$

où \mathbf{L} est une matrice $n \times n$ et \mathbf{K} est une matrice $n \times p$. C'est-à-dire une combinaison linéaire des informations dont on dispose (\mathbf{L} et \mathbf{K} sont donc des opérateurs linéaires). Compte tenu de l'équation d'observation $\mathbf{y} = \mathbf{H}\mathbf{x}^t + \mathbf{e}^o$, l'erreur associée à cette combinaison peut être obtenue comme suit

$$\begin{aligned} \mathbf{x}^a - \mathbf{x}^t &= \mathbf{L}(\mathbf{x}^b - \mathbf{x}^t + \mathbf{x}^t) + \mathbf{K}(\mathbf{H}\mathbf{x}^t + \mathbf{e}^o) - \mathbf{x}^t \\ \mathbf{e}^a &= \mathbf{L}\mathbf{e}^b + \mathbf{K}\mathbf{e}^o + (\mathbf{L} + \mathbf{K}\mathbf{H} - \mathbf{I})\mathbf{x}^t. \end{aligned} \quad (1.13)$$

Il est naturel de souhaiter que l'erreur commise soit sans biais. Or, suivant l'hypothèse retenue que les erreurs d'observation et d'ébauche sont sans biais ($\mathbf{E}[\mathbf{e}^o] = 0$ et $\mathbf{E}[\mathbf{e}^b] = 0$), il découle du calcul qui précède que $\mathbf{E}[\mathbf{e}^a] = (\mathbf{L} + \mathbf{K}\mathbf{H} - \mathbf{I})\mathbf{E}[\mathbf{x}^t]$. On souhaite donc que $\mathbf{L} = \mathbf{I} - \mathbf{K}\mathbf{H}$ (condition suffisante).

Il en résulte une forme plus précise d'un Ansatz *linéaire* et *sans biais*

$$\begin{aligned} \mathbf{x}^a &= (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{x}^b + \mathbf{K}\mathbf{y}, \\ \mathbf{x}^a &= \mathbf{x}^b + \underbrace{\mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b)}_{\text{innovation}}. \end{aligned} \quad (1.14)$$

1. On dit aussi matrice jacobienne, voire *jacobienne* tout court, à ne pas confondre avec jacobien qui est le déterminant de la matrice jacobienne et qui est un scalaire.

\mathbf{K} est un opérateur linéaire de \mathbb{R}^p dans \mathbb{R}^n appelé *gain*. $\mathbf{y} - \mathbf{H}\mathbf{x}^b$ est appelé vecteur *innovation*. Il est un condensé de l'information additionnelle due à l'observation comparée à l'ébauche. La matrice de covariance associée à ses composantes constitue ce que l'on désigne parfois en théorie de l'information par *matrice d'information*.

Puisque \mathbf{K} est linéaire, l'analyse se résume ici à une interpolation linéaire. La terminologie d'usage est toutefois incorrecte puisqu'il s'agit en réalité d'une régression linéaire. L'appellation est historique puisque les premières analyses utilisées en météorologie (Cressman) s'apparentaient à de vraies interpolations aux sens mathématique du terme (i.e. $\mathbf{y} = \mathbf{H}[\mathbf{x}^a]$).

Avec un estimateur linéaire, le problème de l'estimation se ramène donc à déterminer "un gain satisfaisant".

Erreur commise

Supposons momentanément que l'on ait déterminé le gain \mathbf{K} . Quelle est alors la variance de l'erreur commise \mathbf{P}^a ? Pour la calculer, on réécrit les deux membres de l'équation (1.14) à l'aide des vecteurs d'erreurs introduits jusqu'à présent

$$\mathbf{e}^a = \mathbf{e}^b + \mathbf{K}(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b) \quad (1.15)$$

alors

$$\begin{aligned} \mathbf{P}^a &= \mathbb{E}[(\mathbf{e}^a)(\mathbf{e}^a)^T] = \mathbb{E}\left[\left(\mathbf{e}^b + \mathbf{K}(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)\right)\left(\mathbf{e}^b + \mathbf{K}(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)\right)^T\right] \\ &= \mathbb{E}\left[\left(\mathbf{L}\mathbf{e}^b + \mathbf{K}\mathbf{e}^o\right)\left(\mathbf{L}\mathbf{e}^b + \mathbf{K}\mathbf{e}^o\right)^T\right] = \mathbb{E}\left[\mathbf{L}\mathbf{e}^b(\mathbf{e}^b)^T\mathbf{L}^T\right] + \mathbb{E}\left[\mathbf{K}\mathbf{e}^o(\mathbf{e}^o)^T\mathbf{K}^T\right] \\ &= \mathbf{L}\mathbf{P}^b\mathbf{L}^T + \mathbf{K}\mathbf{R}\mathbf{K}^T, \end{aligned} \quad (1.16)$$

où l'on a utilisé la décorrélation de \mathbf{e}^o avec \mathbf{e}^b et la linéarité de \mathbf{K} . En résumé :

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T. \quad (1.17)$$

1.3.2 Estimation optimale : l'analyse BLUE

On cherche maintenant, dans ce cadre, à minimiser l'erreur scalaire commise dans l'analyse, c'est-à-dire $\text{Tr}(\mathbf{P}^a)$. Puisque l'on recherche un \mathbf{K} optimal que l'on notera \mathbf{K}^* , on étudie la variation de $\text{Tr}(\mathbf{P}^a)$ sous une variation $\delta\mathbf{K}$ de \mathbf{K} .

$$\begin{aligned} \delta(\text{Tr}(\mathbf{P}^a)) &= \text{Tr}\left(\left(-\delta\mathbf{K}\mathbf{H}\right)\mathbf{P}^b\mathbf{L}^T + \mathbf{L}\mathbf{P}^b\left(-\delta\mathbf{K}\mathbf{H}\right)^T + \delta\mathbf{K}\mathbf{R}\mathbf{K}^T + \mathbf{K}\mathbf{R}\delta\mathbf{K}^T\right) \\ &= \text{Tr}\left(\left(-\mathbf{L}\mathbf{P}^b\mathbf{H}^T - \mathbf{L}\mathbf{P}^b\mathbf{H}^T + \mathbf{K}\mathbf{R}^T + \mathbf{K}\mathbf{R}\right)(\delta\mathbf{K})^T\right) \\ &= 2\text{Tr}\left(\left(-\mathbf{L}\mathbf{P}^b\mathbf{H}^T + \mathbf{K}\mathbf{R}\right)(\delta\mathbf{K})^T\right) \end{aligned} \quad (1.18)$$

On a utilisé le fait que $\text{Tr}(\mathbf{A}) = \text{Tr}(\mathbf{A}^T)$ et que les matrices de covariance \mathbf{P}^b et \mathbf{R} sont symétriques. À l'optimalité, on en déduit donc que $-(\mathbf{I} - \mathbf{K}^*\mathbf{H})\mathbf{P}^b\mathbf{H}^T + \mathbf{K}^*\mathbf{R} = \mathbf{0}$, dont on tire la solution

$$\mathbf{K}^* = \mathbf{P}^b\mathbf{H}^T(\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T)^{-1}. \quad (1.19)$$

L'estimation \mathbf{x}^a et \mathbf{P}^a qui découle de cette analyse s'appelle une estimation **BLUE** (*Best Linear Unbiased Estimator*), puisqu'elle est : linéaire (hypothèse première : \mathbf{L} et \mathbf{K} introduits), sans biais (première étape de la construction : $\mathbf{L} = \mathbf{I} - \mathbf{KH}$) et optimale (seconde étape de la construction).

1.3.3 Propriétés

Formule de Sherman-Morrison-Woodbury

La formula Eq.(1.19) est la forme sous laquelle on présente en général le gain optimal. Pourtant, en s'y prenant autrement, on peut tout aussi justement aboutir à la formule

$$\mathbf{K}^* = (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}. \quad (1.20)$$

L'identité des deux résultats (formule de Sherman-Morrison-Woodbury) tient à ce que, formellement,

$$\begin{aligned} & \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{H}^T + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{P}^b \mathbf{H}^T) (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{R} + \mathbf{H} \mathbf{P}^b \mathbf{H}^T) (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \end{aligned} \quad (1.21)$$

La possibilité d'une telle alternative n'est pas sans conséquence. Dans la mesure où généralement l'espace des observations est significativement plus petit que celui des états, l'inversion de la matrice $\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$ est numériquement nettement plus coûteuse que celle de la matrice $\mathbf{R} + \mathbf{H} \mathbf{P}^b \mathbf{H}^T$. Il est donc en pratique avantageux de pouvoir se ramener à (1.19).

Erreur d'analyse optimale

Si l'on choisit l'estimation associée au gain optimal, alors l'erreur commise se réduit simplement à

$$\mathbf{P}^a = (\mathbf{I} - \mathbf{K}^* \mathbf{H}) \mathbf{P}^b. \quad (1.22)$$

En effet, on a

$$\begin{aligned} \mathbf{P}^a &= (\mathbf{I} - \mathbf{KH}) \mathbf{P}^b (\mathbf{I} - \mathbf{KH})^T + \mathbf{KRK}^T \\ &= (\mathbf{I} - \mathbf{KH}) \mathbf{P}^b + \left[\mathbf{KR} - (\mathbf{I} - \mathbf{KH}) \mathbf{P}^b \mathbf{H}^T \right] \mathbf{K}^T. \end{aligned} \quad (1.23)$$

Or l'expression entre crochet est nulle lorsque le gain est optimal ($\mathbf{K} = \mathbf{K}^*$) et on obtient le résultat voulu. On comprend donc maintenant pourquoi \mathbf{K} s'appelle matrice de gain. L'opérateur $\mathbf{I} - \mathbf{K}^* \mathbf{H}$ mesure également la réduction du vecteur innovation au cours du processus d'analyse $\mathbf{y} - \mathbf{H}\mathbf{x}^a = (\mathbf{I} - \mathbf{H}\mathbf{K}^*)(\mathbf{y} - \mathbf{H}\mathbf{x}^b)$ en le vecteur dit de *résidu d'analyse*.

On peut obtenir deux expressions classiques de \mathbf{P}^a en utilisant la formule de \mathbf{P}^a Eq.(1.22), avec les deux expressions du gain Eq.(1.19) et Eq.(1.20). En effet on a d'une

part

$$\begin{aligned}
\mathbf{P}^a &= (\mathbf{I} - \mathbf{K}^* \mathbf{H}) \mathbf{P}^b \\
&= (\mathbf{I} - \mathbf{P}^b \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{P}^b \mathbf{H}^T)^{-1} \mathbf{H}) \mathbf{P}^b \\
&= \mathbf{P}^b - \mathbf{P}^b \mathbf{H}^T (\mathbf{R} + \mathbf{H} \mathbf{P}^b \mathbf{H}^T)^{-1} \mathbf{H} \mathbf{P}^b,
\end{aligned} \tag{1.24}$$

et d'autre part

$$\begin{aligned}
\mathbf{P}^a &= (\mathbf{I} - (\mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{P}^b \\
&= (\mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} - \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}) \mathbf{P}^b \\
&= (\mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1}.
\end{aligned} \tag{1.25}$$

Cette dernière formule établit essentiellement que les inverses des matrices de covariance (matrices de *confiance*, ou de *précision*) sont additives :

$$\mathbf{P}^{a^{-1}} = \mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}. \tag{1.26}$$

Il est possible de relier le gain \mathbf{K}^* directement à \mathbf{P}^a . La formule très pratique suivante joue ce rôle :

$$\mathbf{K}^* = \mathbf{P}^a \mathbf{H}^T \mathbf{R}^{-1}. \tag{1.27}$$

On a en effet

$$\mathbf{P}^a \mathbf{H}^T \mathbf{R}^{-1} = (\mathbf{P}^{b^{-1}} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} = \mathbf{K}^*. \tag{1.28}$$

L'innovation et le résidu d'analyse sont sans biais

Une des conséquences de la linéarité de l'opérateur d'observation et de l'hypothèse de biais nul sur l'erreur d'analyse \mathbf{e}^a est que le vecteur résidu de l'analyse $\mathbf{y} - \mathbf{H}\mathbf{x}^a$ est aussi de biais nul. En effet

$$\mathbf{y} - \mathbf{H}\mathbf{x}^a = \mathbf{H}\mathbf{x}^t + \mathbf{e}^o - \mathbf{H}\mathbf{x}^a = -\mathbf{H}\mathbf{e}^a + \mathbf{e}^o, \tag{1.29}$$

d'où l'on tire immédiatement que $E[\mathbf{y} - \mathbf{H}\mathbf{x}^a] = \mathbf{0}$. Sachant qu'on a supposé un biais nul sur l'erreur d'ébauche, il en est de même pour le vecteur innovation

$$\mathbf{y} - \mathbf{H}\mathbf{x}^b = \mathbf{H}\mathbf{x}^t + \mathbf{e}^o - \mathbf{H}\mathbf{x}^b = -\mathbf{H}\mathbf{e}^b + \mathbf{e}^o, \tag{1.30}$$

d'où l'on tire que $E[\mathbf{y} - \mathbf{H}\mathbf{x}^b] = \mathbf{0}$.

Orthogonalité de l'estimation et de l'erreur d'analyse

On cherche à calculer la matrice de covariance dans le cas non-nécessairement optimal

$$\mathbf{C} = E[\mathbf{x}^a (\mathbf{e}^a)^T]. \tag{1.31}$$

On suppose que l'ébauche satisfaisait :

$$E[\mathbf{x}^b (\mathbf{e}^b)^T] = \mathbf{0}. \tag{1.32}$$

C'est à dire qu'il y a décorrélation entre l'ébauche et l'erreur associée.

On a vu que $\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b) = \mathbf{x}^b + \mathbf{K}(-\mathbf{H}\mathbf{e}^b + \mathbf{e}^o)$. Alors :

$$\begin{aligned} \mathbf{C} &= \mathbb{E} \left[\left(\mathbf{x}^b + \mathbf{K}(-\mathbf{H}\mathbf{e}^b + \mathbf{e}^o) \right) \left((\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{e}^b + \mathbf{K}\mathbf{e}^o \right)^T \right] \\ &= -\mathbf{K}\mathbf{H}\mathbb{E} \left[\mathbf{e}^b(\mathbf{e}^b)^T \right] (\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbb{E} \left[\mathbf{e}^o(\mathbf{e}^o)^T \right] \mathbf{K}^T \\ &= \mathbf{K} \left[-\mathbf{H}\mathbf{P}^b (\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{R}\mathbf{K}^T \right]. \end{aligned} \quad (1.33)$$

Puis, si l'analyse est optimale, on a $\mathbf{K}\mathbf{R} - (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P}^b\mathbf{H}^T = \mathbf{0}$, et donc $\mathbf{C} = \mathbf{0}$. On a ainsi dans ce cas orthogonalité de l'estimé \mathbf{x}^a avec l'erreur d'analyse \mathbf{e}^a . Il y a même équivalence entre optimalité et orthogonalité ! Ainsi l'erreur de l'analyse optimale est décorrélée de l'estimation issue de la même analyse.

1.4 Équivalence variationnelle

On se place dans les conditions de BLUE, c'est-à-dire que H est un opérateur linéaire et que les conditions du problème à résoudre sont identiques. On souhaite montrer que le résultat BLUE peut-être obtenu par une approche variationnelle (minimisation d'une fonctionnelle).

1.4.1 Équivalence avec BLUE

Alors on définit la fonctionnelle

$$J(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^b, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) \right) + \frac{1}{2} \left(\mathbf{y} - \mathbf{H}\mathbf{x}, \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \right), \quad (1.34)$$

qui est appelée *fonction coût*. (\cdot, \cdot) désigne le produit scalaire canonique de \mathbb{R}^n ou de \mathbb{R}^p : $(a, b) = a^T b$. Comme H est linéaire, J est une fonctionnelle quadratique en \mathbf{x} . Comme \mathbf{P}^b et \mathbf{R} sont définies positives, cette fonctionnelle est convexe et possède donc un unique minimum. Quel est-il ?

$$\begin{aligned} \delta J(\mathbf{x}) &= \frac{1}{2} \left(\delta\mathbf{x}, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) \right) + \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^b, \mathbf{P}^{b-1} \delta\mathbf{x} \right) \\ &\quad + \frac{1}{2} \left(-\mathbf{H}\delta\mathbf{x}, \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \right) + \frac{1}{2} \left(\mathbf{y} - \mathbf{H}\mathbf{x}, -\mathbf{H}\delta\mathbf{x} \right) \\ &= \left(\delta\mathbf{x}, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \right). \end{aligned} \quad (1.35)$$

On a donc comme condition d'extremum $\mathbf{P}^{b-1}(\mathbf{x}^* - \mathbf{x}^b) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}^*) = \mathbf{0}$, ce qui se réécrit

$$\mathbf{x}^* = \mathbf{x}^b + \underbrace{(\mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1}}_{\mathbf{K}^*} (\mathbf{y} - \mathbf{H}\mathbf{x}^b). \quad (1.36)$$

Donc \mathbf{x}^* est identifié à l'analyse optimale \mathbf{x}^a de BLUE. On notera qu'ici la formule de Sherman-Morrison-Woodbury permet de faire le lien avec le résultat de BLUE.

1.4.2 Propriétés de l'approche variationnelle

Qualité de l'analyse et Hessien

Du calcul (1.35), on tire le gradient de la fonctionnelle

$$\nabla J(\mathbf{x}) = \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}). \quad (1.37)$$

Dans le cas présent, comme la fonction coût est quadratique, il est très simple d'en déduire le Hessien : $\text{Hess}(\mathbf{x}) = \mathbf{P}^{b-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H}$. Or c'est une expression de l'inverse de la matrice de covariance d'analyse que l'on a déjà rencontrée ! Donc

$$\mathbf{P}^a = \text{Hess}(\mathbf{x})^{-1}. \quad (1.38)$$

C'est plus qu'une autre expression formelle pour \mathbf{P}^a . Elle permet de comprendre que la qualité de l'analyse est proportionnelle à la convexité de la fonction de coût J . Plus le minimum est piqué, meilleure est l'analyse.

Extension non-linéaire

Outre l'éclairage qu'elle apporte, la formulation variationnelle du problème présente deux avantages importants. D'abord, elle permet une extension naturelle et immédiate d'un problème linéaire à un problème non-linéaire, essentiellement lorsque H n'est plus linéaire. Ceci n'est pas possible avec BLUE sans linéarisation.

Avantage algorithmique

D'autre part elle présente un avantage algorithmique significatif. En effet, sous BLUE, on est contraint de calculer l'inverse de la matrice $\mathbf{R} + \mathbf{H}\mathbf{P}^b\mathbf{H}^T$ qui apparaît dans le gain \mathbf{K}^* . Avec la méthode variationnelle, on minimise la fonction de coût J , qui n'exige que de calculer plusieurs fois le produit d'un vecteur par les inverses de \mathbf{P}^b et \mathbf{R} , ce qui peut avoir un coût algorithmique plus faible qu'une inversion de matrice lorsque le nombre d'itérations est limité.

La formulation variationnelle du problème d'interpolation statistique se nomme **3D-Var**. Elle a été utilisée de façon opérationnelle dans les services de prévisions météorologique dans les années 90 en supplantant l'interpolation optimale de type BLUE. Il ne lui a été que récemment préféré le **4D-Var**, une généralisation du 3D-Var.

1.4.3 Cas où l'opérateur d'observation H est non-linéaire

Comme on vient de le mentionner, un avantage majeur de la méthode variationnelle est de permettre de traiter le cas où H est un opérateur non-linéaire. Voyons ce que cela modifie dans l'analyse qui précède.

$$J(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^b, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) \right) + \frac{1}{2} \left(\mathbf{y} - H[\mathbf{x}], \mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]) \right), \quad (1.39)$$

On introduit alors le *linéaire tangent* de H en \mathbf{x} , noté logiquement \mathbf{H} . Contrairement au cas linéaire précédent, \mathbf{H} dépend maintenant du point d'observation. Le gradient de J est très similaire

$$\nabla J(\mathbf{x}) = \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) - \mathbf{H}^T \mathbf{R}^{-1}(\mathbf{y} - H[\mathbf{x}]). \quad (1.40)$$

La qualité de l'analyse se déduit toujours avantageusement de $\mathbf{P}^a = \text{Hess}(\mathbf{x})^{-1}$. Cette fois-ci le Hessien dépend véritablement de \mathbf{x} .

1.4.4 Formalisme dual

On se place dans les cas où l'opérateur d'observation est linéaire. La fonction coût de départ est :

$$J(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^b, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) \right) + \frac{1}{2} \left(\mathbf{y} - \mathbf{H}\mathbf{x}, \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) \right). \quad (1.41)$$

On peut introduire un vecteur de multiplicateurs de Lagrange \mathbf{w} (dont la dimension est celle de l'espace des observations) de façon à contraindre l'équation d'observation :

$$L(\mathbf{x}, \epsilon, \mathbf{w}) = \frac{1}{2} \left(\mathbf{x} - \mathbf{x}^b, \mathbf{P}^{b-1}(\mathbf{x} - \mathbf{x}^b) \right) + \frac{1}{2} \left(\epsilon, \mathbf{R}^{-1}\epsilon \right) + \mathbf{w}^T (\mathbf{y} - \mathbf{H}\mathbf{x} - \epsilon) \quad (1.42)$$

L'optimum (sur \mathbf{w}) de cette dernière fonctionnelle est équivalent à la fonction de coût précédente. Mais on a de plus, d'après le théorème du *minmax* que le minimum de la fonction de coût $J(\mathbf{x})$ est égal au maximum de la fonctionnelle, fonction de \mathbf{w} , engendrée par le minimum de L sur \mathbf{x} et ϵ , noté $G(\mathbf{w})$. Que vaut-il? La condition de minimum (gradient nul) de L implique les équations

$$\mathbf{x} = \mathbf{x}^b + \mathbf{P}^b \mathbf{H}^T \mathbf{w} \quad \epsilon = \mathbf{R}\mathbf{w}. \quad (1.43)$$

Ce qui conduit à la fonction de coût *duale*

$$G(\mathbf{w}) = \frac{1}{2} \left(\mathbf{w}, \left(\mathbf{R} + \mathbf{H}\mathbf{P}^b \mathbf{H}^T \right) \mathbf{w} \right) - \left(\mathbf{w}, \mathbf{y} - \mathbf{H}\mathbf{x}^b \right), \quad (1.44)$$

qui est l'opposé de L . L'avantage majeur de cette approche est que l'optimisation de la nouvelle fonctionnelle s'effectue dans l'espace des observations \mathbb{R}^p . Or celui-ci est très souvent réputé de taille nettement plus petite que celui des états \mathbb{R}^n . Ce formalisme porte le nom de *PSAS* pour *Physical Statistical space Assimilation System*.

1.5 Un exemple minimaliste

Imaginons que l'on vienne de faire naufrage en mer à quelques kilomètres de la côte, et qu'avant d'embarquer dans un canot de sauvetage on ait eu le temps de relever les coordonnées de sa position de façon très précise $(u, v) = (0, v_b)$, en prenant l'axe Ox parallèle à la côte et l'axe Oy qui lui est perpendiculaire. Une heure plus tard, on souhaite évaluer sa position. Pour cela, on estime la distance au rivage v_0 *au jugé* avec une variance σ_0^2 . De plus, on veut utiliser le fait que l'on connaissait la position du canot il y a une heure de cela. Entre temps, et en l'absence de courant, le canot a dérivé et la probabilité qu'il se trouve à la position (u_b, v_b) suit une loi normale de variance σ_b^2 qui dépend linéairement du temps écoulé depuis le naufrage. On suppose naturellement qu'il n'y a aucune corrélation entre le processus d'observation et celui de la dérive.

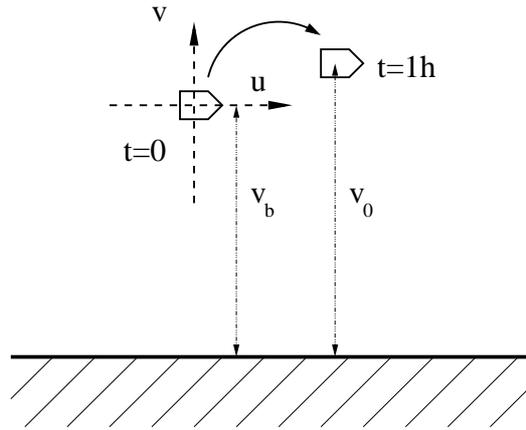


FIGURE 1.3 – La géométrie du problème du naufragé.

1.5.1 Équation d'observation et matrices de covariance d'erreur

Le vecteur d'état est celui des coordonnées de la barque $\mathbf{x} = (u, v)^T$. Si l'équation de la mesure est $y = \mathbf{H}\mathbf{x} + \varepsilon$, alors la fonction de mesure est $\mathbf{H} = (0, 1)$, le vecteur observation est (v_o) , et l'erreur ε suit une loi normale de variance σ_o^2 . La matrice de covariance d'erreur sur l'observation est $R = (\sigma_o^2)$, alors que la matrice de covariance de l'ébauche est

$$\mathbf{P}^b = \begin{pmatrix} \sigma_b^2 & 0 \\ 0 & \sigma_b^2 \end{pmatrix}. \quad (1.45)$$

1.5.2 Analyse optimale

L'analyse linéaire est de la forme $\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathbf{H}\mathbf{x}^b)$ qui se traduit ici par

$$\begin{pmatrix} u_a \\ v_a \end{pmatrix} = \begin{pmatrix} 0 \\ v_b \end{pmatrix} + \mathbf{K} \left(v_o - (0, 1) \begin{pmatrix} 0 \\ v_b \end{pmatrix} \right) = \begin{pmatrix} 0 \\ v_b + \mathbf{K}(v_o - v_b) \end{pmatrix}. \quad (1.46)$$

Calculons alors le gain optimal \mathbf{K}^* . Pour cela, on utilise la forme adéquate de l'expression de \mathbf{K} lorsque le nombre d'observation est inférieur à celui des variables d'états.

$$\begin{aligned} \mathbf{K}^* &= \sigma_b^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \left(\sigma_o^2 + (0, 1) \sigma_b^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^{-1} \\ &= \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \end{aligned} \quad (1.47)$$

On en déduit que

$$\begin{pmatrix} u_a \\ v_a \end{pmatrix} = \begin{pmatrix} 0 \\ v_b + \frac{\sigma_b^2}{\sigma_o^2 + \sigma_b^2} (v_o - v_b) \end{pmatrix}. \quad (1.48)$$

L'observation n'ayant apporté aucune information sur la coordonnée de Ox , on en reste à la coordonnée du point de naufrage $u_a = 0$. Plus le temps passe, plus σ_b grandit. v_a tend alors vers v_o , qui reste dans cette limite la seule information à laquelle on puisse accorder du crédit.

1.5.3 Erreur commise

On déduit de $\mathbf{P}^a = (\mathbf{I} - \mathbf{K}^*\mathbf{H})\mathbf{P}^b$ que

$$\mathbf{P}^a = \begin{pmatrix} \sigma_b^2 & 0 \\ 0 & \frac{\sigma_o^2 \sigma_b^2}{\sigma_o^2 + \sigma_b^2} \end{pmatrix}. \quad (1.49)$$

On constate que, comme attendu, l'analyse ne fait rien gagner sur la connaissance de la coordonnée parallèle au rivage u et son incertitude est celle liée à la dérive, donc proportionnelle au temps. Alors que l'incertitude sur la coordonnée v est réduite par un facteur $\sqrt{\frac{\sigma_o^2}{\sigma_o^2 + \sigma_b^2}}$. Plus l'observation est fine, plus est réduite l'erreur. Si l'on note σ_a^2 l'erreur sur la coordonnée v après analyse, alors on a

$$\frac{1}{\sigma_a^2} = \frac{1}{\sigma_o^2} + \frac{1}{\sigma_b^2}. \quad (1.50)$$

L'interprétation de la formule (1.25) est donc : **la confiance (ou précision) sur l'analyse est la somme de la confiance sur l'observation et de la confiance sur l'ébauche.**

La fonction de coût du problème variationnel qui est équivalente à cette analyse BLUE est

$$J(u, v) = \frac{1}{2\sigma_b^2} (u^2 + v^2) + \frac{1}{2\sigma_o^2} (v_o - v)^2. \quad (1.51)$$

La fonctionnelle construite dans le formalisme PSAS est

$$G(y) = \frac{1}{2}(\sigma_o^2 + \sigma_b^2) w^2 - w(v_o - v_b), \quad (1.52)$$

dont il est très facile d'obtenir la solution

$$w^* = \frac{v_o - v_b}{\sigma_o^2 + \sigma_b^2}. \quad (1.53)$$

Chapitre 2

Interpolation séquentielle : Le filtre de Kalman

Jusqu'à présent il ne s'est agi que d'effectuer une unique analyse d'un problème d'estimation compte tenu d'un ensemble d'informations comprenant une ébauche et un ensemble d'observations. Il n'y avait donc pas, a priori, de dimension temporelle à ce problème d'estimation.

Or, en météorologie par exemple, on souhaite non seulement réaliser une analyse en temps réel, mais on souhaite aussi en effectuer une succession, au fur et à mesure de l'évolution des conditions météorologiques. La réception des observations météo, ainsi que les analyses successives destinées à faire des prévisions s'inscrivent dans le temps. Pour modéliser cela, on définit une série de dates $t_0, t_1, ..t_k, ..t_n$ marquant ces événements (dans les centres de prévision météo européens, on a actuellement $t_{k+1} - t_k = 6$ heures).

Outre le temps, l'interpolation séquentielle incorpore en général un ingrédient supplémentaire comparée à l'interpolation statistique. Il s'agit du modèle d'évolution de l'état du système entre les instants t_k et t_{k+1} . En météorologie, il s'agit du modèle numérique qui simule les équations physiques d'évolution des variables d'états de l'atmosphère (implémentation numérique des équations primitives). En transport réactif d'espèces chimiques, il s'agit d'un code de chimie-transport, appelé génériquement CTM (Chemistry Transport Model).

Le schéma d'assimilation typique est le suivant : à l'instant t_k , on dispose du résultat d'une prévision notée \mathbf{x}_k^f , l'indice f faisant référence à "forecast". \mathbf{x}_k^f est donc l'analogue de l'ébauche \mathbf{x}^b de l'interpolation statistique. À la date t_k , on obtient le résultat d'un ensemble d'observations regroupées dans \mathbf{y}_k . Compte tenu de \mathbf{x}_k^f et des observations \mathbf{y}_k , on procède à une analyse qui produit \mathbf{x}_k^a . On a alors recours au modèle d'évolution pour établir une prédiction sur l'état du système x_{k+1} à l'instant t_{k+1} (on abandonne l'indice t rappelant *truth* utilisé dans le chapitre 1). Le résultat de la prévision est noté \mathbf{x}_{k+1}^f et constitue l'ébauche qui sera utilisée à l'étape suivante (voir Fig. 2.1). Et ainsi de suite...

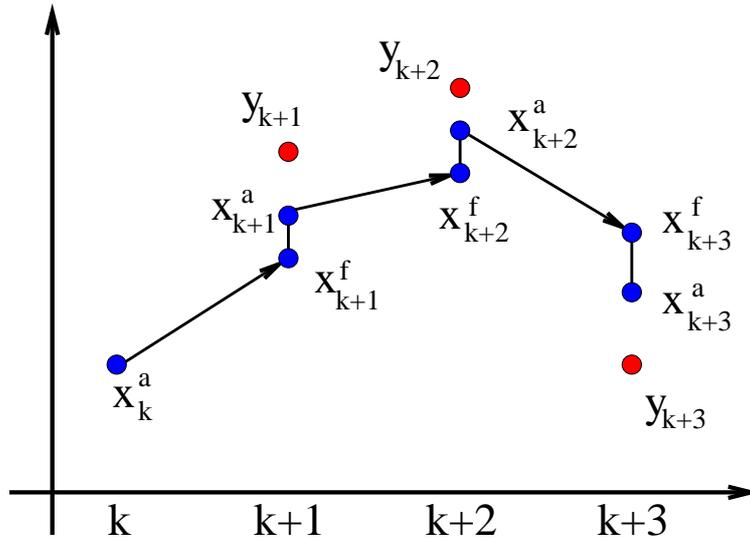


FIGURE 2.1 – Le processus séquentiel d’assimilation du filtre de Kalman.

2.1 Modélisation stochastique du système

Le schéma décrit précédemment se représente à l’aide du système d’équations stochastiques suivant :

$$\begin{cases} \mathbf{x}_{k+1} = M_{k+1}(\mathbf{x}_k) + \mathbf{w}_k, \\ \mathbf{y}_k = H_k(\mathbf{x}_k) + \mathbf{v}_k. \end{cases} \quad (2.1)$$

\mathbf{x}_k désigne l’état vrai du système à l’instant t_k . Le système se décompose en :

- une première équation qui décrit l’étape de prévision. M_{k+1} est la matrice de **transfert** (dite aussi **de transition**) qui simule l’évolution du système de l’instant t_k à l’instant t_{k+1} . Cette opérateur peut dépendre du temps (processus non-autonome), d’où l’indice k . \mathbf{w}_k est l’erreur commise dans la modélisation par M_k des processus physiques. On suppose que ce bruit est non-biaisé, décorrélé dans le temps (bruit blanc) et de matrice de covariance \mathbf{Q}_k à chaque instant t_k , soit

$$E[\mathbf{w}_k] = \mathbf{0} \quad \text{et} \quad E[\mathbf{w}_k \mathbf{w}_l^T] = \mathbf{Q}_k \delta_{kl}. \quad (2.2)$$

- une seconde équation qui correspond à une équation d’observation similaire à celles déjà rencontrées. H_k est l’opérateur d’observation à l’instant t_k . \mathbf{v}_k est l’erreur d’observation, non-biaisée, décorrélée dans le temps (bruit blanc) et de matrice de covariance \mathbf{R}_k à chaque instant t_k , soit

$$E[\mathbf{v}_k] = \mathbf{0} \quad \text{et} \quad E[\mathbf{v}_k \mathbf{v}_l^T] = \mathbf{R}_k \delta_{kl}. \quad (2.3)$$

Il n’y naturellement aucune corrélation entre l’erreur modèle et l’erreur d’observation si bien que

$$E[\mathbf{v}_k \mathbf{w}_l^T] = \mathbf{0}. \quad (2.4)$$

On se place désormais dans le cadre de la linéarité de tous les opérateurs en jeu (M_k et H_k sont donc notés \mathbf{M}_k et \mathbf{H}_k). Le bien-fondé de cette hypothèse sera discuté plus loin.

2.1.1 Étape de l'analyse

On dispose donc au temps t_k , d'une prédiction \mathbf{x}_k^f supposée sans biais, qui s'apparente à notre ébauche. On suppose connue la matrice de covariance d'erreur \mathbf{P}_k^f associée à \mathbf{x}_k^f . On dispose également des observations \mathbf{y}_k (non biaisées) de matrice de covariance \mathbf{R}_k , elle aussi connue. L'analyse qui en découle est alors semblable à celle de l'interpolation statistique. De plus, si l'estimateur associé à cette analyse correspond à une variance de l'erreur commise minimum, alors c'est une analyse en tout point semblable à une interpolation optimale. Compte tenu des hypothèses de linéarité et de biais nul, ce doit être une analyse BLUE. On utilise donc les résultats déjà obtenus dans le chapitre 1.

L'analyse \mathbf{x}_k^a est de la forme

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f), \quad (2.5)$$

ce qui lui assure automatiquement d'être non-biaisée, puisque les erreurs associées à \mathbf{y}_k et \mathbf{x}_k^f ne le sont pas. Sachant qu'on en déduit $\mathbf{e}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{e}_k^f + \mathbf{K}_k \mathbf{v}_k$, la matrice de covariance de l'erreur commise est alors

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^f (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T. \quad (2.6)$$

De plus, si l'analyse est optimale, alors la matrice de gain optimal \mathbf{K}_k^* à l'instant t_k vérifie

$$-(\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{K}_k^* \mathbf{R} = \mathbf{0} \quad (2.7)$$

et s'écrit donc

$$\mathbf{K}_k^* = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1}, \quad (2.8)$$

et dans ce cas, on a

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f. \quad (2.9)$$

Dans ce cadre séquentiel, le gain est appelé *gain de Kalman*.

2.1.2 Étape de la prévision

À ce stade, on a corrigé notre connaissance a priori de l'état du système à l'aide des observations. On souhaite maintenant établir une prédiction sur l'état ultérieur du système (au temps t_{k+1}), au moyen de notre connaissance (souvent imparfaite) des lois qui régissent l'évolution du système entre les temps t_k et t_{k+1} . Par définition, l'estimateur associé est

$$\mathbf{x}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{x}_k^a. \quad (2.10)$$

La linéarité de \mathbf{M}_k assure que l'estimateur est sans biais. L'erreur associée est

$$\begin{aligned} \mathbf{e}_{k+1}^f &= \mathbf{x}_{k+1}^f - \mathbf{x}_{k+1} \\ &= \mathbf{M}_{k+1} (\mathbf{x}_k^a - \mathbf{x}_k) - (\mathbf{x}_{k+1} - \mathbf{M}_{k+1} \mathbf{x}_k) \\ &= \mathbf{M}_{k+1} \mathbf{e}_k^a - \mathbf{w}_k. \end{aligned} \quad (2.11)$$

On en déduit la matrice de covariance traduisant l'erreur associée à la prédiction

$$\begin{aligned} \mathbf{P}_{k+1}^f &= \mathbb{E} \left[\mathbf{e}_{k+1}^f (\mathbf{e}_{k+1}^f)^T \right] \\ &= \mathbb{E} \left[(\mathbf{M}_{k+1} \mathbf{e}_k^a - \mathbf{w}_k) (\mathbf{M}_{k+1} \mathbf{e}_k^a - \mathbf{w}_k)^T \right] \\ &= \mathbf{M}_{k+1} \mathbb{E} \left[(\mathbf{e}_k^a) (\mathbf{e}_k^a)^T \right] \mathbf{M}_{k+1}^T + \mathbb{E} \left[(\mathbf{w}_k) (\mathbf{w}_k)^T \right] \\ &= \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T + \mathbf{Q}_k. \end{aligned} \quad (2.12)$$

à la différence de l'étape d'analyse où les confiances s'ajoutent, les erreurs (sous la forme des variances) s'ajoutent dans l'étape de prédiction.

2.2 Résumé, étude de cas limites et exemple

Filtre de Kalman linéaire

1. Initialisation
 - État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .
2. Pour $t_k = 1, 2, \dots$
 - (a) Analyse
 - Calcul du gain $\mathbf{K}_k^* = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$
 - Calcul de l'estimé de l'analyse

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f)$$
 - Calcul de la matrice de covariance

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f$$
 - (b) Prédiction
 - Calcul de l'estimé de prédiction $\mathbf{x}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{x}_k^a$
 - Calcul de la matrice de covariance

$$\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T + \mathbf{Q}_k$$

2.2.1 Pas d'observation

Dans le cas où il n'y a pas d'observation disponible à assimiler, le filtre de Kalman se réduit à l'étape de prédiction :

$$\begin{aligned} \mathbf{x}_{k+1}^f &= \mathbf{M}_{k+1} \mathbf{x}_k^f \\ \mathbf{P}_{k+1}^f &= \mathbf{M}_{k+1} \mathbf{P}_k^f \mathbf{M}_{k+1}^T + \mathbf{Q}_k. \end{aligned} \quad (2.13)$$

Si la dynamique est neutre ou instable (et on conçoit bien qu'il s'agisse des cas les plus importants), l'erreur ne peut que croître indéfiniment. Seule l'injection de données d'observation peut permettre de réduire l'erreur.

2.2.2 Observations parfaites

Supposons que l'on ait une très grande confiance dans les observations, au point de considérer que $\mathbf{R}_k = \mathbf{0}$. Supposons de plus qu'il y ait autant d'observations que de variables

d'états et qu'elles ne sont pas redondantes. Autrement dit, \mathbf{H}_k est inversible. Alors

$$\begin{aligned}\mathbf{K}_k^* &= \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \\ &= \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k^T)^{-1} (\mathbf{P}_k^f)^{-1} (\mathbf{H}_k)^{-1} \\ &= (\mathbf{H}_k)^{-1}.\end{aligned}\quad (2.14)$$

Il en résulte que $\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f = \mathbf{0}$. Alors $\mathbf{x}_k^f = \mathbf{M}_{k+1} \mathbf{x}_k^a = \mathbf{M}_{k+1} \mathbf{y}_k$. Quant à l'étape de prévision, on a $\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T + \mathbf{Q}_k = \mathbf{Q}_k$. L'erreur commise ne dépend donc bien que de l'erreur modèle, et l'observation suffit à assurer une parfaite connaissance du système.

2.2.3 Un exemple très simple

On revient sur l'exemple du naufragé maritime. Toutes les heures, il évalue la distance qui le sépare de la côte et répète son analyse. La distance au rivage mesurée au temps t_k , k heures après le naufrage survenue à la date t_0 , est notée y_k . Cette évaluation est supposée sans biais. La variance est comme précédemment notée σ_o^2 et elle est supposée stationnaire. Les coordonnées réelles du canot sont (u_k, v_k) , les coordonnées issues de l'analyse sont notées (u_k^a, v_k^a) , et les coordonnées issues de la prévision sont notées (u_k^f, v_k^f) . Au départ (à t_0), on a $(u_0^a, v_0^a) = (0, 0)$, par convention. Entre un temps t_k et t_{k+1} , on sait seulement que le canot a dérivé sans toutefois savoir dans quelle direction. Notre modèle consiste donc seulement à supposer la diffusion de la position du canot autour du point d'origine. Dans notre exemple, on a donc $\mathbf{M}_{k+1} = \mathbf{I}$, la matrice identité. Dans ce contexte, l'erreur modèle est par conséquent importante et s'écrit

$$\mathbf{Q}_k = \begin{pmatrix} \sigma_m^2 & 0 \\ 0 & \sigma_m^2 \end{pmatrix}, \quad (2.15)$$

où σ_m est une mesure de l'incertitude sur l'ampleur de la dérive de l'embarcation entre les instants t_k et t_{k+1} .

On suppose enfin que les matrices de covariance de l'analyse et de la prévision peuvent se mettre sous la forme

$$\mathbf{P}_k^a = \begin{pmatrix} \lambda_k & 0 \\ 0 & \mu_k \end{pmatrix} \quad \mathbf{P}_k^f = \begin{pmatrix} \nu_k & 0 \\ 0 & \rho_k \end{pmatrix}, \quad (2.16)$$

car il n'y a priori aucune corrélation induites entre les deux coordonnées de la position du canot. On pourra constater a posteriori que cette hypothèse est juste.

Analyse

Raisonnons à partir de l'instant t_k . On a pour vecteur résultant de la prévision précédente $(u_k^f, v_k^f)^T$. Le naufragé effectue alors la mesure v_k^o . Il effectue une analyse optimale à partir de ces données. Le gain de Kalman se déduit du formulaire qui précède

$$\begin{aligned}\mathbf{K}_k^* &= \begin{pmatrix} \nu_k & 0 \\ 0 & \rho_k \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \left(\sigma_o^2 + (0, 1) \begin{pmatrix} \nu_k & 0 \\ 0 & \rho_k \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)^{-1} \\ &= \frac{\rho_k}{\sigma_o^2 + \rho_k} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.\end{aligned}\quad (2.17)$$

On en déduit l'estimation des coordonnées de la position du canot

$$\begin{aligned} \begin{pmatrix} u_k^a \\ v_k^a \end{pmatrix} &= \begin{pmatrix} u_k^f \\ v_k^f \end{pmatrix} + \mathbf{K}_k^* \left(y_k - (0, 1) \begin{pmatrix} u_k^f \\ v_k^f \end{pmatrix} \right) \\ &= \begin{pmatrix} u_k^f \\ v_k^f \end{pmatrix} + \mathbf{K}_k^* (y_k - v_k^f) \\ &= \begin{pmatrix} u_k^f \\ v_k^f + \frac{\rho_k}{\sigma_o^2 + \rho_k} (y_k - v_k^f) \end{pmatrix} \end{aligned} \quad (2.18)$$

On déduit ensuite de $\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f$ que

$$\lambda_k = \nu_k \quad \text{et} \quad \frac{1}{\mu_k} = \frac{1}{\sigma_o^2} + \frac{1}{\rho_k}. \quad (2.19)$$

Prévision

On passe maintenant à l'étape de prévision. Le modèle laisse alors inchangé l'estimateur de la position

$$\begin{pmatrix} u_{k+1}^f \\ v_{k+1}^f \end{pmatrix} = \begin{pmatrix} u_k^a \\ v_k^a \end{pmatrix}. \quad (2.20)$$

La matrice de covariance d'erreur de prévision s'obtient comme $\mathbf{P}_{k+1}^f = \mathbf{P}_k^a + \sigma_m^2 \mathbf{I}$, ce qui s'explique en

$$\nu_{k+1} = \lambda_k + \sigma_m^2, \quad \rho_{k+1} = \mu_k + \sigma_m^2. \quad (2.21)$$

Il résulte donc de la succession de l'analyse et de la prévision que

$$u_{k+1}^f = u_k^f, \quad v_{k+1}^f = v_k^f + \frac{\rho_k}{\sigma_o^2 + \rho_k} (y_k - v_k^f), \quad (2.22)$$

et

$$\nu_{k+1} = \nu_k + \sigma_m^2, \quad \frac{1}{\rho_{k+1} - \sigma_m^2} = \frac{1}{\sigma_o^2} + \frac{1}{\rho_k}. \quad (2.23)$$

Comme $u_0^f = 0$, il est clair que $u_k^f = 0$ ultérieurement : faute de mesure, on n'apprend rien de plus sur la coordonnée parallèle à la côte. Cependant, l'incertitude croît linéairement avec le temps puisque

$$\nu_k = k \sigma_m^2. \quad (2.24)$$

De l'équation sur l'incertitude de la coordonnée v (2.23), on tire que $\rho_k \geq \sigma_m^2$. On peut chercher un point fixe ρ_* de l'équation de récurrence satisfaisant cette contrainte. Il est solution de

$$\rho_*^2 - \sigma_m^2 \rho_* - \sigma_o^2 \sigma_m^2 = 0, \quad (2.25)$$

ce qui conduit à

$$\rho_* = \frac{\sigma_m^2}{2} \left(1 + \sqrt{1 + 4 \frac{\sigma_o^2}{\sigma_m^2}} \right), \quad (2.26)$$

C'est donc le résultat asymptotique d'un compromis entre la réduction de l'incertitude par l'assimilation des observations et l'accroissement de cette même incertitude due à la dérive non-contrôlée de l'embarcation. On vérifie aisément que $\rho_* \geq \sigma_m^2$.

2.2.4 Deuxième exemple : assimilation d'un oscillateur

On définit le modèle discret direct suivant :

$$x_0 = 0, \quad x_1 = 1 \quad \text{et pour } 1 \leq k \leq N \quad x_{k+1} - 2x_k + x_{k-1} = -\omega^2 x_k. \quad (2.27)$$

Il s'agit bien sûr d'un schéma numérique de l'équation de l'oscillateur harmonique unidimensionnel $\ddot{x} + \Omega^2 x = 0$. C'est une équation discrète du second ordre. Un vecteur d'état comprend donc deux positions

$$u_k = \begin{pmatrix} x_k \\ x_{k-1} \end{pmatrix} \quad (2.28)$$

La matrice de transition est

$$\mathbf{M}_k = \begin{pmatrix} 2 - \omega^2 & -1 \\ 1 & 0 \end{pmatrix}. \quad (2.29)$$

L'opérateur d'observation est $\mathbf{H}_k = (1, 0)$. Le système d'observation est alors

$$y_k = \mathbf{H}_k u_k + \xi_k, \quad (2.30)$$

avec un bruit ξ_k blanc et gaussien, de variance g supposée connue. On ne dispose pas nécessairement d'observations pour tous les temps t_k .

On met en œuvre sur cet exemple le filtre de Kalman linéaire. On donne sur la figure Fig. 2.2 l'exemple d'une assimilation correspondant à un jeu d'observations.

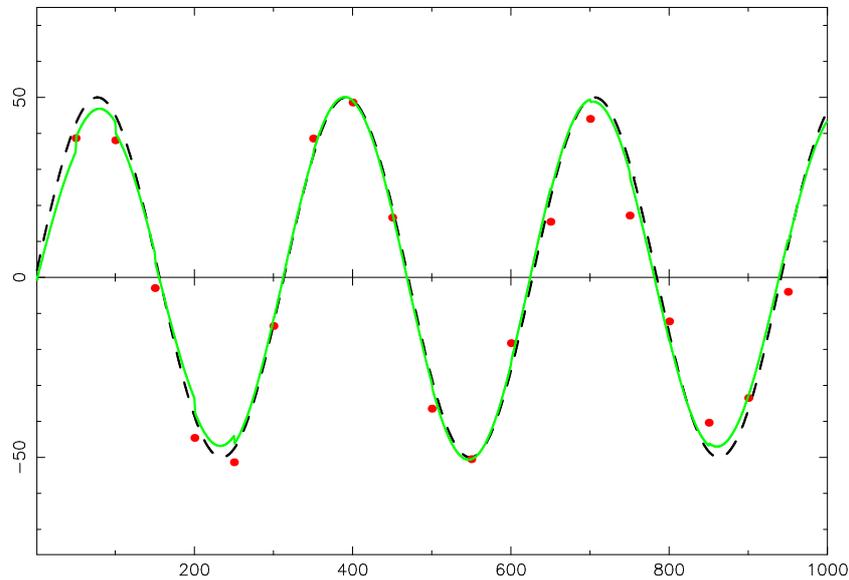


FIGURE 2.2 – La courbe tiretée représente l'évolution du modèle direct. En sus des conditions initiales retenues dans le texte, on a choisi $N = 1000$ unités de temps, $\omega = 0.02 \text{ s}^{-1}$. Les points représentent les observations, de variance d'erreur $g = 7$. Les observations sont acquises tous les $\Delta = 50$ unités de temps. La courbe pleine représente la prévision de $H_k u_k^f$ donnée par le filtre de Kalman linéaire.

2.3 Le filtre de Kalman étendu

Jusqu'à présent, on a supposé que l'hypothèse de linéarité était valide. Or il se peut que

- la fonction d'observation H_k soit non-linéaire. C'est parfois le cas lorsque des données satellitaires sont assimilées. L'opérateur d'observation qui relie des données optiques (radiances) aux variables d'état du système peut faire intervenir un modèle de transfert radiatif. Lorsque la diffusion est prise en compte, ce modèle peut être fortement non-linéaire.
- le modèle M_k soit non-linéaire. C'est le cas des équations primitives (atmosphère et océan), qui donnent notamment naissance au chaos. C'est aussi le cas en chimie atmosphérique dès que les réactions chimiques sont du deuxième ordre. On est donc très souvent dans ce cas !

Pour autant, comme on va le voir, cela n'implique pas d'abandonner le filtre de Kalman, mais davantage de l'adapter.

2.3.1 Linéarisations de la prévision et de l'analyse

L'estimé de prévision est naturellement donné par

$$\mathbf{x}_{k+1}^f = M_{k+1}[\mathbf{x}_k^a]. \quad (2.31)$$

On montrera dans le chapitre 4 que cet estimateur n'est précis qu'au premier ordre des perturbations (dans un sens à préciser) et qu'il peut être amélioré.

Pour déterminer la matrice de covariance d'erreur de prévision, on va utiliser le fait que l'on espère \mathbf{x}_k^a suffisamment proche de \mathbf{x}_k . On notera \mathbf{M}_k la matrice linéaire tangente à M_k en \mathbf{x}_k . Reprenons d'abord le raisonnement mené pour l'étape de prévision. On a

$$\begin{aligned} \mathbf{e}_{k+1}^f &= \mathbf{x}_{k+1}^f - \mathbf{x}_{k+1} = M_{k+1}[\mathbf{x}_k^a] - \mathbf{x}_{k+1} \\ &= M_{k+1} \left[\mathbf{x}_k + (\mathbf{x}_k^a - \mathbf{x}_k) \right] - \mathbf{x}_{k+1} \\ &\simeq M_{k+1}[\mathbf{x}_k] + \mathbf{M}_{k+1}(\mathbf{x}_k^a - \mathbf{x}_k) - \mathbf{x}_{k+1} \\ &\simeq \mathbf{M}_{k+1}\mathbf{e}_k^a - \mathbf{w}_k. \end{aligned} \quad (2.32)$$

Il en résulte que

$$\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1}\mathbf{P}_k^a\mathbf{M}_{k+1}^\top + \mathbf{Q}_k. \quad (2.33)$$

L'estimateur de l'étape d'analyse reste donné par une analyse de la forme

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k (\mathbf{y}_k - H_k[\mathbf{x}_k^f]). \quad (2.34)$$

On en déduit que

$$\begin{aligned} \mathbf{e}_k^a &= \mathbf{x}_k^a - \mathbf{x}_k = \mathbf{x}_k^f - \mathbf{x}_k + \mathbf{K}_k (\mathbf{y}_k - H_k[\mathbf{x}_k] + H_k[\mathbf{x}_k] - H_k[\mathbf{x}_k^f]) \\ &= \mathbf{x}_k^f - \mathbf{x}_k + \mathbf{K}_k (\mathbf{y}_k - H_k[\mathbf{x}_k] + H_k[\mathbf{x}_k] - H_k[\mathbf{x}_k^f - \mathbf{x}_k + \mathbf{x}_k]) \\ &\simeq \mathbf{e}_k^f + \mathbf{K}_k (\mathbf{e}_k^o - \mathbf{H}_k \mathbf{e}_k^f). \end{aligned} \quad (2.35)$$

Et de là

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)\mathbf{P}_k^f(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k)^\top + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^\top. \quad (2.36)$$

Le gain de Kalman est alors :

$$\mathbf{K}_k^* = \mathbf{P}_k^f \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k)^{-1}, \quad (2.37)$$

et dans ce cas, on a

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f. \quad (2.38)$$

L'estimateur optimal est donc

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* (\mathbf{y}_k - H_k[\mathbf{x}_k^f]). \quad (2.39)$$

On notera que le linéaire tangent de H_k intervient dans \mathbf{K}_k^* , mais pas dans le vecteur innovation $\mathbf{y}_k - H_k[\mathbf{x}_k^f]$.

2.3.2 Résumé

Filtre de Kalman étendu

1. Initialisation
 - État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .
2. Pour $t_k = 1, 2, \dots$
 - (a) Analyse
 - Calcul du gain $\mathbf{K}_k^* = \mathbf{P}_k^f (\mathbf{H}_k)^T (\mathbf{H}_k \mathbf{P}_k^f (\mathbf{H}_k)^T + \mathbf{R}_k)^{-1}$
 - Calcul de l'estimé de l'analyse
$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* (\mathbf{y}_k - H_k[\mathbf{x}_k^f])$$
 - Calcul de la matrice de covariance
$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f$$
 - (b) Prévion
 - Calcul de l'estimé de prévision $\mathbf{x}_{k+1}^f = M_{k+1}[\mathbf{x}_k^a]$
 - Calcul de la matrice de covariance
$$\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T + \mathbf{Q}_k$$

2.3.3 Assimilation des observations sur un oscillateur non-linéaire

On définit le modèle discret direct suivant :

$$x_0 = 0, \quad x_1 = 1 \quad \text{et pour } 1 \leq k \leq N \quad x_{k+1} - 2x_k + x_{k-1} = \omega^2 x_k - \lambda^2 x_k^3. \quad (2.40)$$

Il s'agit bien sûr d'un schéma numérique de l'équation de l'oscillateur anharmonique unidimensionnel $\ddot{x} - \Omega^2 x + \Lambda^2 x^3 = 0$. Le potentiel associé est $-\frac{1}{2}\Omega^2 x^2 + \frac{1}{4}\Lambda^2 x^4$. Le second terme stabilise l'oscillateur et joue comme force de rappel, alors que le premier terme destabilise

le point $x = 0$, créant deux puits de potentiel. C'est une équation discrète du second ordre. Un vecteur d'état comprend donc deux positions

$$u_k = \begin{pmatrix} x_k \\ x_{k-1} \end{pmatrix} \quad (2.41)$$

La matrice de transition est

$$M_k = \begin{pmatrix} 2 + \omega^2 - \lambda^2 x_k^2 & -1 \\ 1 & 0 \end{pmatrix}. \quad (2.42)$$

L'opérateur d'observation est $\mathbf{H}_k = (1, 0)$. Le système d'observation est alors

$$y_k = \mathbf{H}_k u_k + \xi_k, \quad (2.43)$$

avec un bruit ξ_k blanc et gaussien, de variance g supposée connue. On ne dispose pas nécessairement d'observation pour tous les temps t_k .

On met en œuvre sur cet exemple le filtre de Kalman étendu. On donne sur la figure Fig.2.3 l'exemple d'une assimilation correspondant à un jeu d'observations.

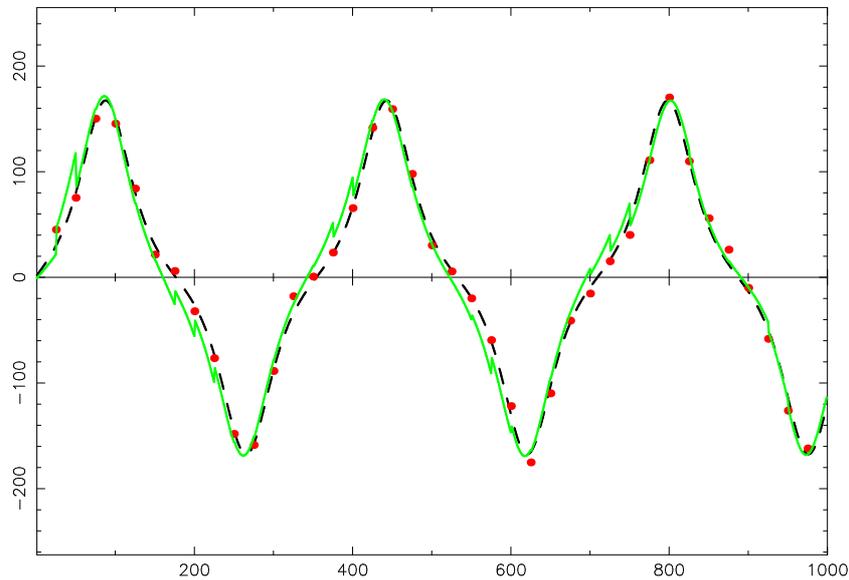


FIGURE 2.3 – La courbe tiretée représente l'évolution du modèle direct. En sus des conditions initiales retenues dans le texte, on a choisi $N = 1000$ unités de temps, $\omega = 0.035 \text{ s}^{-1}$, et $\lambda = 0.003 \text{ s}^{-1}$. Les points représentent les observations, de variance d'erreur $g = 7$. Les observations sont acquises tous les $\Delta = 25$ unités de temps. La courbe pleine représente la prévision de $\mathbf{H}_k u_k^f$ donnée par le filtre de Kalman linéaire.

2.4 Compléments :

2.4.1 Théorème d'indépendance des innovations

De même que pour l'interpolation non-séquentielle, on peut montrer que les vecteurs innovations (que l'on note $\mathbf{d}_k = \mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f$) ont un biais nul (sont d'espérance nulle).

$$\begin{aligned} \mathbf{d}_k &= \mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f \\ &= \mathbf{H}_k \mathbf{x}_k - \mathbf{H}_k \mathbf{x}_k^f + \mathbf{v}_k \\ &= \mathbf{H}_k \mathbf{e}_k^f + \mathbf{v}_k \end{aligned} \quad (2.44)$$

D'où l'on tire que $E(\mathbf{d}_k) = \mathbf{0}$ puisque \mathbf{e}_k^f et \mathbf{v}_k sont sans biais. Cette propriété découle donc de l'hypothèse d'absence de biais des sources d'erreur.

On va maintenant montrer que dans le cas d'un filtrage optimal, la suite des vecteurs innovations est blanche (deux éléments distincts de la suite sont statistiquement indépendants). C'est un résultat bien plus fort que le résultat précédent puisqu'il requiert l'optimalité.

On suppose donc que l'analyse n'est pas nécessairement optimale et on cherche à calculer la fonction $C_{k,l} = E(\mathbf{d}_k \mathbf{d}_l^T)$, qui est la matrice de covariance des vecteurs innovations puisque leur espérance est nulle. On a donc

$$\begin{aligned} C_{k,l} &= E \left[(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k^f) (\mathbf{y}_l - \mathbf{H}_l \mathbf{x}_l^f)^T \right] = E \left[(\mathbf{v}_k - \mathbf{H}_k \mathbf{e}_k^f) (\mathbf{v}_l - \mathbf{H}_l \mathbf{e}_l^f)^T \right] \\ &= E \left[\mathbf{v}_k \mathbf{v}_l^T \right] + \mathbf{H}_k E \left[\mathbf{e}_k^f (\mathbf{e}_l^f)^T \right] \mathbf{H}_l^T - \mathbf{H}_k E \left[\mathbf{e}_k^f \mathbf{v}_l^T \right] - E \left[\mathbf{v}_k (\mathbf{e}_l^f)^T \right] \mathbf{H}_l^T \end{aligned} \quad (2.45)$$

Dans le cas où $l = k$, alors on a immédiatement

$$C_{k,k} = \mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T. \quad (2.46)$$

On considère maintenant le cas $k-1 \geq l \geq 0$. Par respect de la causalité et parce que la suite des erreurs d'observation est blanche, il reste

$$C_{k,l} = \mathbf{H}_k E \left[\mathbf{e}_k^f (\mathbf{e}_l^f)^T \right] \mathbf{H}_l^T - \mathbf{H}_k E \left[\mathbf{e}_k^f \mathbf{v}_l^T \right]. \quad (2.47)$$

Il s'agit donc de calculer cette matrice de covariance. \mathbf{e}_k^f est a priori statistiquement lié à \mathbf{e}_l^f . On cherche donc à donner de \mathbf{e}_k^f la dépendance explicite en \mathbf{e}_l^f . Pour cela, on utilise les relations déjà obtenues :

$$\mathbf{e}_k^f = \mathbf{M}_k \mathbf{e}_{k-1}^a - \mathbf{w}_{k-1} \quad \text{et} \quad \mathbf{e}_{k-1}^a = (\mathbf{I} - \mathbf{K}_{k-1} \mathbf{H}_{k-1}) \mathbf{e}_{k-1}^f + \mathbf{K}_{k-1} \mathbf{v}_{k-1}. \quad (2.48)$$

D'où l'on tire que

$$\mathbf{e}_k^f = \mathbf{M}_k (\mathbf{I} - \mathbf{K}_{k-1} \mathbf{H}_{k-1}) \mathbf{e}_{k-1}^f + \mathbf{M}_k \mathbf{K}_{k-1} \mathbf{v}_{k-1} - \mathbf{w}_{k-1}, \quad (2.49)$$

ce qui permet d'amorcer une récurrence descendant jusqu'à l . Il est commode de noter $\Gamma_{k,k-1} = \mathbf{M}_k (\mathbf{I} - \mathbf{K}_{k-1} \mathbf{H}_{k-1})$ et plus généralement

$$\Gamma_{k,l} = \mathbf{M}_k (\mathbf{I} - \mathbf{K}_{k-1} \mathbf{H}_{k-1}) \mathbf{M}_{k-1} (\mathbf{I} - \mathbf{K}_{k-2} \mathbf{H}_{k-2}) \times \cdots \times \mathbf{M}_{l+1} (\mathbf{I} - \mathbf{K}_l \mathbf{H}_l), \quad (2.50)$$

pour $k - 1 \leq l \leq 0$ et $\Gamma_{k,k} = \mathbf{I}$ sinon. On en déduit par induction que

$$\mathbf{e}_k^f = \Gamma_{k,l} \mathbf{e}_l^f + \sum_{j=l}^{j=k-1} \Gamma_{k,j+1} (M_{j+1} \mathbf{K}_j \mathbf{v}_j - \mathbf{w}_j). \quad (2.51)$$

On a alors en tenant compte de la causalité

$$\begin{aligned} C_{k,l} &= \mathbf{H}_k E \left[\Gamma_{k,l} \mathbf{e}_l^f (\mathbf{e}_l^f)^\top \right] \mathbf{H}_l^\top - \mathbf{H}_k E \left[\Gamma_{k,l+1} \mathbf{M}_{l+1} \mathbf{K}_l \mathbf{v}_l \mathbf{v}_l^\top \right] \\ &= \mathbf{H}_k \Gamma_{k,l} \mathbf{P}_l^f \mathbf{H}_l^\top - \mathbf{H}_k \Gamma_{k,l+1} \mathbf{M}_{l+1} \mathbf{K}_l \mathbf{R}_l \\ &= \mathbf{H}_k \Gamma_{k,l+1} (\Gamma_{l+1,l} \mathbf{P}_l^f \mathbf{H}_l^\top - \mathbf{M}_l \mathbf{K}_l \mathbf{R}_l) \\ &= \mathbf{H}_k \Gamma_{k,l+1} \mathbf{M}_{l+1} ((\mathbf{I} - \mathbf{K}_l \mathbf{H}_l) \mathbf{P}_l^f \mathbf{H}_l^\top - \mathbf{K}_l \mathbf{R}_l). \end{aligned} \quad (2.52)$$

Or on sait qu'avec un filtre optimal on a $-(\mathbf{I} - \mathbf{K}^* \mathbf{H}) \mathbf{P}^f \mathbf{H}^\top + \mathbf{K}^* \mathbf{R} = 0$, donc

$$C_{k,l} = \mathbf{H}_k \Gamma_{k,l+1} \mathbf{M}_{l+1} ((\mathbf{I} - \mathbf{K}_l \mathbf{H}_l) \mathbf{P}_l^f \mathbf{H}_l^\top - \mathbf{K}_l \mathbf{R}_l - (\mathbf{I} - \mathbf{K}_l^* \mathbf{H}_l) \mathbf{P}_l^f \mathbf{H}_l^\top + \mathbf{K}_l^* \mathbf{R}_l), \quad (2.53)$$

et en simplifiant

$$C_{k,l} = \mathbf{H}_k \Gamma_{k,l+1} \mathbf{M}_{l+1} (\mathbf{K}_l^* - \mathbf{K}_l) (\mathbf{H}_l \mathbf{P}_l^f \mathbf{H}_l^\top + \mathbf{R}_l). \quad (2.54)$$

Par conséquent, dans le cas où l'analyse est complètement optimale

$$\forall k, l \quad C_{k,l} = (\mathbf{R}_k + \mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^\top) \delta_{k,l}. \quad (2.55)$$

Cette propriété procure un moyen pertinent et pratique de tester l'optimalité d'un filtre de Kalman, en particulier de vérifier l'absence de biais.

2.4.2 Orthogonalité entre l'estimé de l'analyse et l'erreur d'analyse

De façon très similaire à la propriété de l'interpolation optimale mise en évidence dans la section (1.3.3), on cherche à calculer la matrice de covariance dans le cas non-nécessairement optimal

$$C_k = E \left[\mathbf{x}_k^a (\mathbf{e}_k^a)^\top \right]. \quad (2.56)$$

On va procéder par induction sur le temps t_k , $k = 1, \dots, n$. \mathbf{x}_1^f et \mathbf{e}_1^f constituent l'ébauche et son erreur. Le calcul de C_0 est donc rigoureusement identique à celui effectué dans la section(1.3.3) dans le cadre de l'interpolation optimale. À savoir

$$C_0 = 0, \quad (2.57)$$

à la condition que la propriété soit vérifiée pour la toute première ébauche, soit :

$$E \left[\mathbf{x}_0^f (\mathbf{e}_0^f)^\top \right] = 0. \quad (2.58)$$

Supposons maintenant que la propriété est satisfaite au temps t_{k-1} : $C_{k-1} = 0$. Alors on a

$$\begin{aligned}
C_k &= E \left[\mathbf{x}_k^a (\mathbf{e}_k^a)^T \right] = E \left[(\mathbf{x}_k^f + \mathbf{K}_k (-\mathbf{H}_k \mathbf{e}_k^f + \mathbf{v}_k)) ((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{e}_k^f + \mathbf{K}_k \mathbf{v}_k)^T \right] \\
&= E \left[(\mathbf{M}_k \mathbf{x}_{k-1}^a) ((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{e}_k^f + \mathbf{K}_k \mathbf{v}_k)^T \right] \\
&\quad + E \left[(\mathbf{K}_k (-\mathbf{H}_k \mathbf{e}_k^f + \mathbf{v}_k)) ((\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{e}_k^f + \mathbf{K}_k \mathbf{v}_k)^T \right] \\
&= \mathbf{M}_k E \left[\mathbf{x}_{k-1}^a (\mathbf{M}_k \mathbf{e}_{k-1}^a - \mathbf{w}_{k-1})^T \right] (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{M}_k E \left[\mathbf{x}_{k-1}^a \mathbf{v}_k^T \right] \mathbf{K}_k^T \\
&\quad + \mathbf{K}_k \left[-\mathbf{H}_k \mathbf{P}_k^f (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{R}_k \mathbf{K}_k^T \right] \\
&= \mathbf{M}_k E \left[\mathbf{x}_{k-1}^a (\mathbf{e}_{k-1}^a)^T \right] \mathbf{M}_k^T (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T - \mathbf{M}_k E \left[\mathbf{x}_{k-1}^a \mathbf{w}_k^T \right] (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T \\
&\quad + \mathbf{K}_k \left[-\mathbf{H}_k \mathbf{P}_k^f (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{R}_k \mathbf{K}_k^T \right] \\
&= \mathbf{M}_k C_{k-1} \mathbf{M}_k^T (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \left[-\mathbf{H}_k \mathbf{P}_k^f (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{R}_k \mathbf{K}_k^T \right]. \quad (2.59)
\end{aligned}$$

Alors compte tenu de l'hypothèse de récurrence $C_{k-1} = 0$ et si l'analyse au temps t_k est optimale $\mathbf{K}_k = \mathbf{K}_k^*$, on a $C_k = 0$. Ce qui complète la preuve de la propriété annoncée.

2.4.3 Propagation de l'erreur d'analyse

On montre que l'on a :

$$\mathbf{P}_k^f = \Gamma_{k,l} \mathbf{P}_l^f (\mathbf{M}_{l,k})^T \quad (2.60)$$

Chapitre 3

Formalisme variationnel généralisé : le 4D-Var

On a vu dans le chapitre précédent que l'interpolation optimale pouvait se généraliser à une interpolation séquentielle lorsque

- le système à assimiler possède une dimension temporelle.
- et lorsque son évolution temporelle est décrite par un modèle (parfait ou non, déterministe ou bruité).

On aboutit alors au formalisme du filtre de Kalman. D'autre part, dans le chapitre 1 on a fait correspondre à la technique de l'interpolation optimale une approche variationnelle (le 3D-Var), qui possède l'intérêt

- d'offrir une alternative efficace en temps de calcul dans la résolution numérique du problème.
- de permettre une généralisation naturelle à des problèmes non-linéaires.

Sous certaines conditions, on peut faire de même avec le filtre de Kalman tel que vu dans le chapitre 2 en lui associant un formalisme variationnel appelé *4D-Var*.

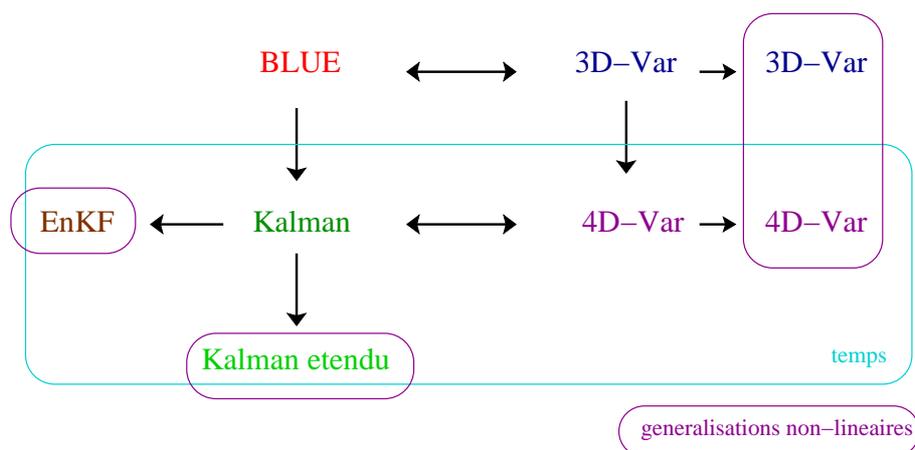


FIGURE 3.1 – Les principaux schémas d'assimilation de données et leur relation. Le filtre EnKF n'a pas été étudié à ce stade, et le 4D-Var est l'objet de ce chapitre.

Une correspondance rigoureuse entre le 4D-Var et le filtre de Kalman ne s'envisage que lorsque la fonction d'observation H et la dynamique M sont linéaires. Elles peuvent néanmoins dépendre du temps (ce que rappellent les notations H_k et M_k). C'est ce qu'on suppose donc dans cette section. De plus, on va également supposer dans la majeure partie de ce chapitre que le modèle *direct* rend compte totalement de l'évolution déterministe du système, c'est-à-dire qu'on suppose l'erreur modèle nulle. On dit que le modèle direct est *parfait* (on envisagera ultérieurement le cas où cette restriction est levée). On dit alors que le problème variationnel est sous *contrainte forte*.

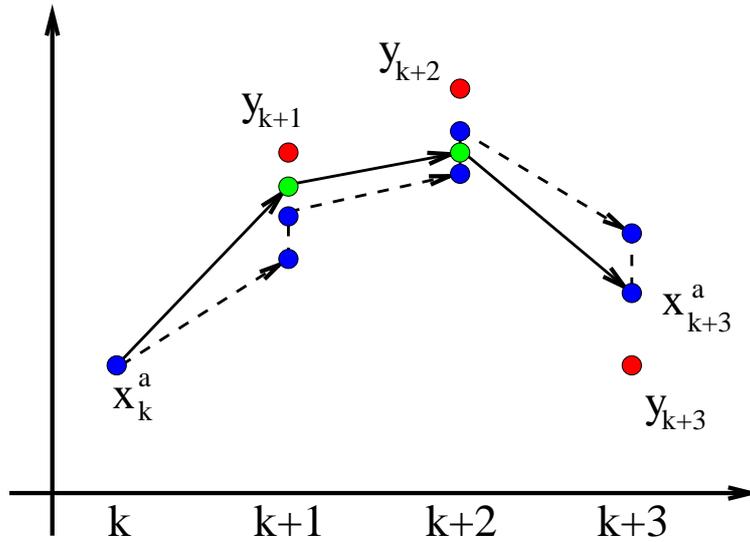


FIGURE 3.2 – L'assimilation 4D-Var comparée à l'assimilation du filtre de Kalman.

3.1 Fonction de coût et calcul de son gradient

On suppose dans ce chapitre que l'état du système est donné par les $K + 1$ vecteurs \mathbf{x}_k pour chacun des $K + 1$ instants t_0, t_1, \dots, t_K . Chacun de ces vecteurs appartient à \mathbb{R}^n . On suppose de plus qu'à chacun de ces instants arrivent des données paramétrisées par $\mathbf{y}_k \in \mathbb{R}^p$. Formellement, la généralisation temporelle de la fonctionnelle 3D-Var (1.34) est

$$J(\mathbf{x}) = \frac{1}{2} \left(\mathbf{x}_0 - \mathbf{x}_0^b, (\mathbf{P}_0^b)^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) \right) + \frac{1}{2} \sum_{k=0}^K (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k, \mathbf{R}_k^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k)) . \quad (3.1)$$

Cette fonctionnelle est définie sous la contrainte que les vecteurs d'état \mathbf{x}_k (dont est fonction J) correspondent à une évolution admissible du système, c'est-à-dire que

$$\forall k \geq 0, \quad \mathbf{x}_{k+1} = M_{k+1}(\mathbf{x}_k) . \quad (3.2)$$

On suppose pour le moment que $M_k = \mathbf{M}_k$ est linéaire. La donnée d'une ébauche complète (pour tout t_k) $\{\mathbf{x}_k^b\}$ est beaucoup plus rare dans le cadre du 4D-Var. Cependant, il est fréquent qu'une ébauche de l'état initial soit donnée, d'espérance \mathbf{x}_0^b et de matrice de covariance \mathbf{P}_0^b .

Un moyen d'inclure directement la contrainte de modèle dans le lagrangien est de recourir à des multiplicateurs de Lagrange $\{\Lambda_k\}$. La fonctionnelle prend alors la forme

$$L(\mathbf{x}, \Lambda) = J(\mathbf{x}) + \sum_{k=0}^{K-1} \Lambda_k^T (\mathbf{x}_{k+1} - \mathbf{M}_{k+1} \mathbf{x}_k). \quad (3.3)$$

C'est un procédé commode dans le cadre de la version continue du formalisme variationnel (pas de temps $t_{k+1} - t_k \rightarrow 0$), mais sans intérêt déterminant dans le cas discret, où un raisonnement direct (sans utiliser Λ_k) est tout aussi efficace. En effet, dans la version discrète, la variable d'état à l'instant t_k , \mathbf{x}_k est donnée, en fonction de la valeur initiale \mathbf{x}_0 de la variable, comme le produit des matrices de transition appliqué à l'état initial

$$\mathbf{x}_{k+1} = \mathbf{M}_{k+1} \mathbf{M}_k \dots \mathbf{M}_1 \mathbf{x}_0. \quad (3.4)$$

Dans un modèle continu la variable d'état serait solution d'un système d'équations différentielles, de manipulation moins commode. La fonctionnelle (3.1) sous contrainte est par conséquent une fonction de \mathbf{x}_0 sans contrainte. La fonction de coût, qui apparait clairement désormais comme une fonction effective de \mathbf{x}_0 peut maintenant être minimisée directement par rapport à \mathbf{x}_0 . Pour cela, on souhaite calculer le gradient de J par rapport au vecteur initial \mathbf{x}_0 . On note

$$\mathbf{d}_k = \mathbf{y}_k - \mathbf{H}_k [\mathbf{M}_k \mathbf{M}_{k-1} \dots \mathbf{M}_1] \mathbf{x}_0 \quad \text{et} \quad \Delta_k = \mathbf{R}_k^{-1} \mathbf{d}_k. \quad (3.5)$$

\mathbf{d}_k est un *vecteur innovation* tandis que Δ_k est un *vecteur innovation normalisé*. D'autre part, on oublie temporairement le terme d'ébauche dans $J(\mathbf{x})$ car un tel terme a déjà été étudié dans le cadre de l'interpolation optimale. Alors

$$\begin{aligned} \delta J(\mathbf{x}_0) &= \frac{1}{2} \sum_{k=0}^K \delta(\mathbf{d}_k, (\mathbf{R}_k)^{-1} \mathbf{d}_k) \\ &= \frac{1}{2} \sum_{k=0}^K (\delta \mathbf{d}_k, (\mathbf{R}_k)^{-1} \mathbf{d}_k) + \frac{1}{2} \sum_{k=0}^K (\mathbf{d}_k, (\mathbf{R}_k)^{-1} \delta \mathbf{d}_k) \\ &= \sum_{k=0}^K (\delta \mathbf{d}_k, (\mathbf{R}_k)^{-1} \mathbf{d}_k) \\ &= - \sum_{k=0}^K (\mathbf{H}_k [\mathbf{M}_k \mathbf{M}_{k-1} \dots \mathbf{M}_1] \delta \mathbf{x}_0, \Delta_k) \\ &= - \sum_{k=0}^K (\delta \mathbf{x}_0, [\mathbf{M}_1^T \mathbf{M}_2^T \dots \mathbf{M}_{k-1}^T \mathbf{M}_k^T] \mathbf{H}_k^T \Delta_k). \end{aligned} \quad (3.6)$$

On en déduit donc le gradient de la fonctionnelle par rapport à \mathbf{x}_0 :

$$\nabla_{\mathbf{x}_0} J = - \sum_{k=0}^K [\mathbf{M}_1^T \mathbf{M}_2^T \dots \mathbf{M}_{k-1}^T \mathbf{M}_k^T] \mathbf{H}_k^T \Delta_k \quad (3.7)$$

$$= - (\mathbf{H}_0^T \Delta_0 + \mathbf{M}_1^T [\mathbf{H}_1^T \Delta_1 + \mathbf{M}_2^T [\mathbf{H}_2^T \Delta_2 + \dots + [\mathbf{M}_n^T] \mathbf{H}_n^T \Delta_n] \dots]) \quad (3.8)$$

Cette dernière écriture du gradient (sous la forme d'une factorisation de Horner) fournit une façon de le calculer à moindre coût :

1. On commence par calculer les \mathbf{x}_k par récurrence à l'aide de la matrice de transition \mathbf{M}_k et de la variable d'état initiale \mathbf{x}_0 .
2. On calcule alors les vecteurs innovation normalisés $\Delta_k = \mathbf{R}_k^{-1}(\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k)$ que l'on enregistre.
3. On définit alors une **variable d'état adjointe** $\tilde{\mathbf{x}}_k$. Elle est définie à la date finale par $\tilde{\mathbf{x}}_K = \mathbf{H}_K^T \Delta_K$. Connaissant $\tilde{\mathbf{x}}_{k+1}$, on obtient $\tilde{\mathbf{x}}_k$ par

$$\tilde{\mathbf{x}}_k = \mathbf{H}_k^T \Delta_k + \mathbf{M}_{k+1}^T \tilde{\mathbf{x}}_{k+1}. \quad (3.9)$$

Par cette récurrence, on remonte à $\tilde{\mathbf{x}}_0$.

4. On a alors $\nabla_{\mathbf{x}_0} J = -\tilde{\mathbf{x}}_0$.

Le calcul de l'évolution à rebours de la variable adjointe nécessite en particulier le calcul de l'opérateur adjoint des matrices de transition \mathbf{M}_k . Cela semble une opération triviale lorsque \mathbf{M}_k est une matrice, même de grande taille. Cependant, en pratique, l'évolution d'un système est calculé à partir d'un code numérique de plusieurs milliers de lignes. C'est un objet plus complexe qu'une simple matrice, dont on doit calculer l'objet adjoint (mathématiquement bien défini). L'**adjointisation** présente donc une difficulté technique. Aujourd'hui, on a recours à des programmes d'**adjointisation automatique**[8] qui, à partir du code associé à M_k délivrent un code numérique associé à \mathbf{M}_k^T . Des exemples de tels programmes sont : TAPENADE de l'INRIA (le code à adjointiser peut être soumis en ligne et le code adjoint retourné), TAF (programme commercial très réputé), mais aussi OpenAD, ADIFOR, ADIC, ADOL-C, etc.

3.1.1 Gradient avec un terme d'ébauche

Tenir compte du terme d'ébauche n'apporte qu'une complication mineure. Le gradient devient

$$\nabla_{\mathbf{x}_0} J = (\mathbf{P}_0^b)^{-1}(\mathbf{x}_0 - \mathbf{x}_0^b) - \sum_{k=0}^K \mathbf{M}_{k,1}^T \mathbf{H}_k^T \mathbf{R}_k^{-1}(\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,1} \mathbf{x}_0), \quad (3.10)$$

où pour $k > l \geq 0$, on a noté $\mathbf{M}_{k,l} = \mathbf{M}_k \mathbf{M}_{k-1} \cdots \mathbf{M}_{l+1}$. Conventionnellement, on pose $\mathbf{M}_{k,k} = I$. Par extension, on note pour $l > k \geq 0$ que $\mathbf{M}_{k,l} = \mathbf{M}_{l,k}^{-1} = \mathbf{M}_{k+1}^{-1} \mathbf{M}_{k+2}^{-1} \cdots \mathbf{M}_l^{-1}$. Grâce à cette convention, on a

$$\forall k, l \quad n \leq k, l \leq 0 \quad \mathbf{x}_k = \mathbf{M}_{k,l} \mathbf{x}_l. \quad (3.11)$$

3.1.2 Cas où le modèle direct et/ou l'observation sont non-linéaires

On peut procéder comment on l'a fait dans la section 1.4.3 dans le cadre du formalisme 3D-Var. Comme précédemment, on note \mathbf{M} la matrice jacobienne dérivée de M par rapport à \mathbf{x} . Éventuellement, on peut préciser en quel point est évaluée cette dérivée : $\mathbf{M}(\mathbf{x}_0)$ par exemple lorsqu'on l'évalue en \mathbf{x}_0 . Le calcul du gradient est affecté par la complication apportée par ces non-linéarités. Principalement, le calcul de $\delta \mathbf{d}_k$ est modifié en

$$\delta \mathbf{d}_k = -\mathbf{H}_k(\mathbf{x}_{k-1}) \mathbf{M}_k(\mathbf{x}_{k-2}) \mathbf{M}_{k-1}(\mathbf{x}_{k-3}) \cdots \mathbf{M}_1(\mathbf{x}_0) \delta \mathbf{x}_0. \quad (3.12)$$

Le gradient s'écrit alors

$$\nabla_{\mathbf{x}_0} J = - \sum_{k=0}^K \left[\mathbf{M}_1^T(\mathbf{x}_0) \mathbf{M}_2^T(\mathbf{x}_1) \cdots \mathbf{M}_{k-1}^T(\mathbf{x}_{k-2}) \mathbf{M}_k^T(\mathbf{x}_{k-1}) \right] \mathbf{H}_k^T(\mathbf{x}_k) \Delta_k. \quad (3.13)$$

Dans le cadre d'une application effective de cette généralisation non-linéaire, il est donc non seulement nécessaire de savoir calculer l'adjoint d'un code numérique mais aussi la différentielle d'un code. C'est ce qu'on appelle la *différentiation automatique*.

3.2 Solution du problème variationnel

Dans un premier temps, on va chercher à minimiser la fonctionnelle $J(\mathbf{x})$. C'est ensuite seulement que l'on montrera en quoi et quand cette démarche est bien équivalente à une interpolation optimale séquentielle du type filtre de Kalman.

3.2.1 Solution par rapport à \mathbf{x}_0

De l'expression du gradient obtenu, il est facile d'obtenir une solution au problème de minimisation de J par rapport à \mathbf{x}_0 en résolvant $\nabla_{\mathbf{x}_0} J = 0$. Avant cela, il est toutefois plus commode d'introduire le Hessien de la fonction J , noté $\mathcal{H}_{K,0}$, pour une assimilation sur l'intervalle de temps $[t_0, t_K]$. Dans le cas d'une fonctionnelle quadratique comme c'est ici le cas, la matrice Hessienne qui ne dépend plus de \mathbf{x}_0 s'obtient aisément à partir du gradient

$$\mathcal{H}_{K,0} = (\mathbf{P}_0^b)^{-1} + \sum_{k=0}^K \mathbf{M}_{k,0}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{M}_{k,0}. \quad (3.14)$$

Avec cette notation, la solution \mathbf{x}_0^a s'obtient très simplement de (3.10)

$$\mathbf{x}_0^a = \mathbf{x}_0^b - \mathcal{H}_{K,0}^{-1} \nabla_{\mathbf{x}_0} J(\mathbf{x}_0^b). \quad (3.15)$$

On reconnaît la solution de Newton au problème de minimisation d'une fonctionnelle quadratique.

L'erreur d'analyse commise avec cette estimation est donnée par la matrice de covariance

$$\mathbf{P}_0^a = E \left[(\mathbf{x}_0^a - \mathbf{x}_0) (\mathbf{x}_0^a - \mathbf{x}_0)^T \right]. \quad (3.16)$$

Et l'on a

$$(\mathbf{P}_0^a)^{-1} = \mathcal{H}_{K,0} = (\mathbf{P}_0^b)^{-1} + \sum_{k=0}^K \mathbf{M}_{k,0}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{M}_{k,0}. \quad (3.17)$$

3.2.2 Solution par rapport à \mathbf{x}_k

Dans le cas du modèle direct parfait, les valeurs (vraies) de l'état du système \mathbf{x}_k sont toutes reliées entre elles, grâce à la matrice de transition $\mathbf{M}_{k,l}$. Réaliser une analyse sur la condition initiale \mathbf{x}_0 est donc équivalent à le faire sur un quelconque \mathbf{x}_j , $j \in [0, K]$. Dans cet esprit, on peut récrire la fonctionnelle de coût (3.1) en fonction de \mathbf{x}_j ($K \geq j \geq 0$) et non plus de \mathbf{x}_0 selon

$$\begin{aligned} J(\mathbf{x}) &= \frac{1}{2} \left(\mathbf{M}_{0,j} \left(\mathbf{x}_j - \mathbf{x}_j^b \right), (\mathbf{P}_0^b)^{-1} \mathbf{M}_{0,j} \left(\mathbf{x}_j - \mathbf{x}_j^b \right) \right) \\ &\quad + \frac{1}{2} \sum_{k=0}^K \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,j} \mathbf{x}_j, (\mathbf{R}_k)^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,j} \mathbf{x}_j) \right), \end{aligned} \quad (3.18)$$

grâce à la convention (3.11). On remarque que $\mathbf{M}_{0,j}^\top (\mathbf{P}_0^b)^{-1} \mathbf{M}_{0,j} = \left(\mathbf{M}_{j,0} \mathbf{P}_0^b \mathbf{M}_{j,0}^\top \right)^{-1} \equiv (\mathbf{P}_j^b)^{-1}$. Le calcul du gradient par rapport à \mathbf{x}_j s'en déduit

$$-\nabla_{\mathbf{x}_j} J = \left(\mathbf{P}_j^b \right)^{-1} \left(\mathbf{x}_j^b - \mathbf{x}_j \right) + \sum_{k=0}^K \mathbf{M}_{k,j}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \left(\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,j} \mathbf{x}_j \right). \quad (3.19)$$

On en tire l'analyse donnée par $\nabla_{\mathbf{x}_j} J(\mathbf{x}) = 0$

$$\mathbf{x}_j^a = \mathbf{x}_j^b - \mathcal{H}_{K,j}^{-1} \nabla_{\mathbf{x}_j} J(\mathbf{x}_j^b), \quad (3.20)$$

où l'on a noté

$$\mathcal{H}_{K,j} = \left(\mathbf{P}_j^b \right)^{-1} + \sum_{k=0}^K \mathbf{M}_{k,j}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{M}_{k,j}. \quad (3.21)$$

L'erreur d'analyse commise avec cette estimation est donnée par la matrice de covariance

$$\mathbf{P}^a_k = E \left[\left(\mathbf{x}_j^a - \mathbf{x}_j \right) \left(\mathbf{x}_j^a - \mathbf{x}_j \right)^\top \right]. \quad (3.22)$$

Si, à ce stade, on se contente de raisonner par analogie avec le résultat (1.38) obtenu par l'approche variationnelle associée à l'interpolation optimale, on a

$$\left(\mathbf{P}^a_j \right)^{-1} = \mathcal{H}_{K,j} = \left(\mathbf{P}_j^b \right)^{-1} + \sum_{k=0}^K \mathbf{M}_{k,j}^\top \mathbf{H}_k^\top \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{M}_{k,j}. \quad (3.23)$$

Ce qui signifie que la matrice de confiance sur \mathbf{x}_0^a est la somme de la confiance sur l'ébauche et des confiances sur les mesures propagées par l'inverse du modèle direct jusqu'à l'instant t_0 .

3.3 Propriétés

3.3.1 Propagation de l'erreur d'analyse

On vérifie aisément que l'on a

$$\mathbf{M}_{j,l}^\top \mathcal{H}_{K,j} \mathbf{M}_{j,l} = \mathcal{H}_{K,l}. \quad (3.24)$$

Compte tenu de $\mathcal{H}_{K,k} = \left(\mathbf{P}^a_k \right)^{-1}$, cela conduit à

$$\mathbf{P}^a_l = \mathbf{M}_{l,j} \mathbf{P}^a_j \mathbf{M}_{l,j}^\top, \quad (3.25)$$

qui constitue la relation décrivant la propagation de l'erreur d'analyse dans le formalisme 4D-Var (sous contrainte forte). De toute évidence, ce résultat diffère de celui de la propagation de l'erreur dans le cadre du filtre de Kalman (2.60). Cela tient à ce que l'analyse de l'erreur dans le filtre de Kalman est *causale* (dépend des observations antérieures à la date de l'analyse), alors que cela n'est pas le cas pour l'analyse 4D-Var.

C'est ce qui constitue le point faible du 4D-Var : sa difficulté à estimer l'incertitude.

3.3.2 Transférabilité de l'optimalité

On dispose maintenant des résultats nécessaires pour démontrer une propriété essentielle attachée à la minimisation de la fonctionnelle. On considère un intervalle de temps d'assimilation $[t_0, t_K]$. Et on subdivise cet intervalle en deux périodes $[t_0, t_m]$ et $[t_m, t_K]$ avec m un entier entre 0 et n . On peut opérer la minimisation de deux manières

1. On minimise la fonctionnelle complète pour le problème défini sur l'intervalle $[t_0, t_K]$.
2. On procède en deux étapes. On commence par minimiser la fonctionnelle associée au problème défini sur l'intervalle $[t_0, t_m]$. On en tire une solution \mathbf{x}_m^a au temps t_m , avec une incertitude \mathbf{P}_m^a . On réalise alors une seconde analyse sur l'intervalle $[t_m, t_K]$, **mais** en incorporant les résultats \mathbf{x}_m^a et \mathbf{P}_m^a de la première étape dans un terme d'ébauche dans la fonctionnelle.

On cherche à montrer que ces deux routes sont équivalentes du point de vue d'une minimisation par rapport à \mathbf{x}_0 , mais on pourrait aussi choisir un vecteur d'état à toute autre date que t_0 .

D'après ce qui précède, le résultat de la voie (1) est le résultat de l'optimisation de la fonctionnelle (3.1), que l'on choisit ici d'effectuer par rapport à \mathbf{x}_0

$$J(\mathbf{x}_0) = \frac{1}{2} \left(\mathbf{x}_0 - \mathbf{x}_0^b, (\mathbf{P}_0^b)^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) \right) + \frac{1}{2} \sum_{k=0}^K (\mathbf{x}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0, (\mathbf{R}_k)^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0)) . \quad (3.26)$$

On peut décomposer la même fonctionnelle en

$$J_{m,0}(\mathbf{x}_0) = \frac{1}{2} \left(\mathbf{x}_0 - \mathbf{x}_0^b, (\mathbf{P}_0^b)^{-1} (\mathbf{x}_0 - \mathbf{x}_0^b) \right) + \frac{1}{2} \sum_{k=0}^m (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0, (\mathbf{R}_k)^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0)) , \quad (3.27)$$

et

$$J_{n,m}(x_0) = \frac{1}{2} \sum_{k=m+1}^K (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0, (\mathbf{R}_k)^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{M}_{k,0} \mathbf{x}_0)) , \quad (3.28)$$

de sorte que

$$J(\mathbf{x}_0) = J_{m,0}(\mathbf{x}_0) + J_{n,m}(\mathbf{x}_0) . \quad (3.29)$$

On remarque que la fonctionnelle $J_{m,0}(\mathbf{x}_0)$ est la fonctionnelle associée à la première étape de la voie (2). Soit donc \mathbf{x}_0^a et \mathbf{P}_0^a le résultat de son optimisation. Compte tenu des hypothèses de linéarité, $J_{m,0}(\mathbf{x}_0)$ est quadratique et s'identifie donc à son développement de Taylor à l'ordre 2 en tout point, et en particulier en \mathbf{x}_0^a . On a

$$J_{m,0}(\mathbf{x}_0) = J_{m,0}(\mathbf{x}_0^a) + (\nabla_{\mathbf{x}_0} J_{m,0}(\mathbf{x}_0^a), \mathbf{x}_0 - \mathbf{x}_0^a) + \frac{1}{2} (\mathbf{x}_0 - \mathbf{x}_0^a, \mathcal{H}_{m,0}(\mathbf{x}_0 - \mathbf{x}_0^a)) . \quad (3.30)$$

Or, par définition, $\nabla_{\mathbf{x}_0} J_{m,0}(\mathbf{x}_0^a) = 0$. Il en résulte que minimiser $J_{m,0}$ est équivalent à minimiser le carré

$$C(\mathbf{x}_0) = \frac{1}{2} \left(\mathbf{x}_0 - \mathbf{x}_0^a, (\mathbf{P}_0^a)^{-1} (\mathbf{x}_0 - \mathbf{x}_0^a) \right) , \quad (3.31)$$

avec

$$\begin{aligned} \mathbf{x}_0^a &= \mathbf{x}_0^b - \mathcal{H}_{m,0}^{-1} \nabla_{\mathbf{x}_0} J(\mathbf{x}_0^b) \\ (\mathbf{P}_0^a)^{-1} &= \mathcal{H}_{m,0} = (\mathbf{P}_0^b)^{-1} + \sum_{k=0}^m \mathbf{M}_{k,0}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{M}_{k,0}, \end{aligned} \quad (3.32)$$

puisque $J_{m,0}(\mathbf{x}_0) = J_{m,0}(\mathbf{x}_0^a) + C(\mathbf{x}_0)$. Donc

$$J(\mathbf{x}_0) = J_{m,0}(\mathbf{x}_0^a) + C(\mathbf{x}_0) + J_{K,m}(\mathbf{x}_0). \quad (3.33)$$

Or $C(\mathbf{x}_0)$ est précisément le terme d'ébauche de la seconde étape de la voie (2). Le résultat de la minimisation globale de $J(\mathbf{x}_0)$ sur $[t_0, t_K]$ selon \mathbf{x}_0 est donc égale au résultat de la minimisation de la procédure en deux étapes, avec un terme d'ébauche dans la seconde étape qui résume les résultats de la première étape. Notons que l'on aurait pu tout aussi bien minimiser sur $\mathbf{x}_K, \mathbf{x}_m$ ou n'importe quel autre \mathbf{x}_k . Cette propriété est appelée **transférabilité de l'optimalité**. On prendra toutefois garde au fait qu'en général il n'y pas identité entre

- le résultat de l'analyse par la voie (1) portant sur le vecteur d'état du système au temps t_m, \mathbf{x}_m^a et
- le résultat de l'analyse *intermédiaire* au temps t_m déduit de la première étape de la voie (2).

3.3.3 Équivalence entre le 4D-Var et le filtre de Kalman

On justifie maintenant a posteriori l'intérêt d'avoir utilisé la fonction de coût (3.1), en montrant que le résultat final x_K^a au temps t_K est celui que fournit l'analyse optimale séquentielle du filtre de Kalman au temps t_K .

On a démontré précédemment la transférabilité de l'optimalité du formalisme variationnel en décomposant l'intervalle d'analyse en deux étapes. On peut subdiviser ces sous-intervalles et invoquer à nouveau la propriété : l'optimalité est toujours assurée. À l'extrême, on peut subdiviser $[t_0, t_K]$ en les K segments $[t_k, t_{k+1}]$ avec $0 \leq k \leq K-1$. Or la minimisation de la fonctionnelle $J_{k,k+1}$ est identique à l'approche 3D-Var, qui, on l'a vu, est équivalent à l'analyse de l'interpolation optimale. Mais c'est précisément l'interpolation optimale qui est utilisée à chaque étape du filtre de Kalman ! Il en résulte que résoudre le problème en subdivisant $[t_0, t_K]$ en les K segments $[t_k, t_{k+1}]$ avec $0 \leq k \leq K-1$, est équivalent à l'analyse du filtre de Kalman. Cette démarche n'est toutefois pas complètement équivalente à l'optimisation de la fonctionnelle J sur $[t_0, t_K]$. On peut cependant affirmer d'après ce qui précède, que le résultat de l'analyse au temps final t_K par l'optimisation globale est le même que celui de l'analyse par subdivision et donc que celui du filtre de Kalman.

3.4 Calculs de sensibilité

Un des atouts majeurs de la formulation variationnelle est de permettre un calcul efficace des **sensibilités**. Il s'agit de déterminer quantitativement l'influence qu'exerce un paramètre sur l'état du système ou sur la solution du problème d'assimilation.

3.4.1 Sensibilité du premier ordre

Hors contexte assimilation de données, on peut vouloir déterminer la variation d'une variable d'état δx_i du modèle direct lorsqu'on fait varier un paramètre p_j de δp_j . On a bien sûr

$$\delta x_i = \chi_{ij} \delta p_j \quad \text{avec} \quad \chi_{ij} = \frac{\partial x_i}{\partial p_j}. \quad (3.34)$$

Si le calcul de cette sensibilité χ_{ij} n'est pas en soi lié à l'assimilation de données, on en utilise toutefois l'arsenal technologique puisque, pour un modèle direct imposant, le calcul de la dérivée partielle doit passer par la différentiation automatique. Remarquons que si le modèle est non-linéaire en les paramètres, χ_{ij} dépend des p_j .

3.4.2 Sensibilité du second ordre

Dans le cadre d'une analyse du second ordre de la sensibilité, on souhaite connaître la sensibilité de l'estimation optimale du problème d'assimilation à un paramètre ou à une observation. C'est donc un problème lié directement à l'assimilation de données. Et son calcul par une méthode variationnelle requiert l'utilisation du modèle direct et du modèle adjoint.

Supposons que le problème dépend de paramètres p_j avec $j = 1, \dots, q$. Il peut s'agir par exemple de coefficients paramétrisant les matrices de covariance d'erreurs, ou de paramètres du modèle physique (coefficients de diffusion, de friction, de dépôts, constantes cinétiques, etc.). On s'intéresse à la sensibilité de la variable d'état x_i^a par rapport au paramètre p_j

$$\chi_{ij} = \frac{\partial x_i^a}{\partial p_j}. \quad (3.35)$$

Or en cas d'unicité de la solution (dans le cas linéaire en particulier), celle-ci est solution implicite de

$$(\nabla_x J)(x^a(p), p) = 0, \quad (3.36)$$

la solution de l'analyse x^a étant elle-même fonction des paramètres p_j . A fortiori, pour tout i, j dans $[1, q]^2$:

$$\frac{d}{dp_j} (\partial_{x_i} J)(x^a(p), p) = 0, \quad (3.37)$$

soit

$$\partial_{x_i} \partial_{p_j} J(x^a(p), p) + \sum_{k=1}^q \frac{\partial x_k^a(p)}{\partial p_j} \partial_{x_k} \partial_{x_i} J(x^a(p), p) = 0. \quad (3.38)$$

Et donc

$$\partial_{x_i} \partial_{p_j} J + \sum_{k=1}^q (\partial_{x_i} \partial_{x_k} J) \frac{\partial x_k^a(p)}{\partial p_j} = 0, \quad (3.39)$$

qui se récrit

$$\chi_{ij} = \frac{\partial x_i^a(p)}{\partial p_j} = - \sum_{k=1}^q (H^{-1})_{ik} \partial_{x_k} \partial_{p_j} J, \quad (3.40)$$

où H est la matrice Hessienne au point de la solution de l'analyse. Les problèmes de calcul de sensibilité sont donc des opérations du second ordre sur la fonction de coût [12]. Ils sont par conséquent lourds à mettre en œuvre, étant donnée la taille des matrices (tenseurs) en jeu.

3.5 Algorithmes de minimisation de la fonction de coût

Abordons maintenant l'aspect pratique mais essentiel des méthodes variationnelles : la minimisation des fonctionnelles. On se donne donc une fonction de coût $J(\mathbf{x})$ que l'on cherche à minimiser. Si J est quadratique strictement convexe, il existe un unique minimum dans \mathbb{R}^n . Plus généralement, J peut posséder plusieurs minima locaux (mais si le problème est physiquement bien posé, il doit exister au moins un minimum). La détermination de la liste exhaustive de ces minima est difficile. On se concentre ici sur la détermination du minimum ou d'un de ces minima. On notera $\mathbf{g}(\mathbf{x})$ le gradient de $J(\mathbf{x})$, et $\mathbf{H}(\mathbf{x})$ la matrice hessienne associée. Un moyen de tester les algorithmes d'optimisation est de les appliquer au cas de la fonction de coût quadratique

$$J_{\mathbf{Q}}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, \mathbf{Q}\mathbf{x}) + (\mathbf{b}, \mathbf{x}), \quad (3.41)$$

où \mathbf{Q} est une matrice symétrique définie positive (donc J est strictement convexe).

3.5.1 Algorithmes de descente

Les algorithmes de minimisation commencent par un point \mathbf{x}_0 et construisent une suite de points \mathbf{x}_k qui, on l'espère, approche un minimum local. \mathbf{x}_0 doit se trouver dans le bassin d'attraction (algorithme-dépendant) du minimum local recherché. À l'étape k de la construction, on détermine une direction \mathbf{d}_k caractéristique de la méthode. Elle sert à définir le point suivant qui est donné par

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \lambda_k \mathbf{d}_k, \quad (3.42)$$

où $\lambda_k \in \mathbb{R}$. λ_k est obtenu par une minimisation unidimensionnelle ("*line search*"), le long de la direction affine définie par \mathbf{x}_k et \mathbf{d}_k . Le réel λ_k est une solution résultant de la minimisation de

$$\varphi_k(\lambda) \equiv J(\mathbf{x}_k - \lambda \mathbf{d}_k), \quad (3.43)$$

Comme il s'agit d'un problème de minimisation, la direction \mathbf{d}_k est en règle générale une direction de descente $(\mathbf{g}(\mathbf{x}_k), \mathbf{d}_k) > 0$. Il en résulte

$$\varphi'_k(0) = -(\mathbf{g}(\mathbf{x}_k), \mathbf{d}_k) < 0. \quad (3.44)$$

De sorte que l'on choisit d'optimiser en restreignant λ à \mathbb{R}^+ .

Méthodes de la plus forte pente, du gradient conjugué, et de Newton-Raphson

On peut choisir pour \mathbf{d}_k la direction de la plus forte pente (*steepest descent method*), c'est-à-dire $\mathbf{d}_k = \mathbf{g}(\mathbf{x}_k)$, puisque c'est ce qu'indique le gradient $\mathbf{g}(\mathbf{x})$ de la fonction de coût. On peut aussi exiger des vecteurs \mathbf{d}_k qu'ils explorent plus efficacement l'espace de minimisation que ne le fait l'algorithme de descente de plus forte pente, dans lequel on n'a aucune liberté sur le choix de \mathbf{d}_k .

Si l'on raisonne sur le cas quadratique Eq. (3.41), avant toute extrapolation au cas non-linéaire, on peut imposer que les $\{\mathbf{d}_j\}_{0 \leq j \leq k}$ soient Q -conjugués, c'est-à-dire qu'ils soient mutuellement orthogonaux par rapport à la forme définie par Q . Cette construction, analogue à l'orthogonalisation de Gram-Schmidt est possible et assure que la solution puisse

être obtenue en au plus $n + 1$ étapes (où n est la taille du système) si l'erreur numérique est nulle (précision machine infinie). De plus, il est clair que dans le cas quadratique, on obtient une solution explicite pour α_k , donnée ci-dessous. Cette méthode est appelée **gradient conjugué**.

Elle s'applique également au cas non-linéaire (la conjugaison n'est alors qu'approximativement satisfaite et l'approche de la solution nécessite en général plus de $n + 1$ itérations). Les différentes variantes (Hestenes-Stiefel, Polak-Ribière et Fletcher-Reeves) de cette généralisation diffèrent par la formule explicite retenue pour λ_k . Le point essentiel est que dans tous les cas, ces formules sont construites de telle sorte qu'elles ne fassent pas intervenir le calcul du Hessien.

En dépit de la sophistication du gradient conjugué et de ses variantes, il s'avère néanmoins souvent plus efficace d'utiliser la méthode de Newton-Raphson. Son avantage réside dans sa vitesse de convergence. L'idée qui soutend la méthode est la suivante. On souhaite approcher localement dans un voisinage de \mathbf{x}_k la fonction J de coût par une fonction quadratique.

$$J(\mathbf{x}) = \underbrace{J(\mathbf{x}_k) + (\mathbf{g}(\mathbf{x}_k), \mathbf{x} - \mathbf{x}_k) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_k, \mathbf{H}(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k))}_{q(\mathbf{x})} + o(\|\mathbf{x} - \mathbf{x}_k\|^2). \quad (3.45)$$

Une estimation du minimum local est alors donnée par le minimum de l'approximation quadratique $q(\mathbf{x})$, noté \mathbf{x}_{k+1} , que l'on obtient en résolvant $\nabla q(\mathbf{x}) \simeq 0$:

$$0 \simeq \nabla q(\mathbf{x}) = \mathbf{g}(\mathbf{x}_k) + \mathbf{H}(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k). \quad (3.46)$$

Si le Hessien est inversible, il en résulte qu'une meilleure idée du point de minimum $\bar{\mathbf{x}}$ est donné par $\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{H}(\mathbf{x}_k)^{-1}\mathbf{g}(\mathbf{x}_k)$. Autrement dit

$$\mathbf{d}_k = -\mathbf{H}(\mathbf{x}_k)^{-1}\mathbf{g}(\mathbf{x}_k). \quad (3.47)$$

On notera que cela correspond à faire un choix pour \mathbf{d}_k , mais aussi pour $\lambda_k \in \mathbb{R}^+$ qui est pris égal à 1. On peut montrer qu'on aboutit à une méthode convergente. Appliquée à une fonction de coût quadratique, cette méthode est dite du second ordre car

$$\mathbf{x}_{k+1} - \bar{\mathbf{x}} = O(\|\mathbf{x}_k - \bar{\mathbf{x}}\|^2). \quad (3.48)$$

Cette méthode est très efficace. Elle présente néanmoins un inconvénient majeur, puisqu'elle requiert le calcul du Hessien et de son inverse. Pour un problème non-linéaire, le calcul doit être effectué à chaque itération car le Hessien change. Pour un problème de grande taille, comme ceux qui nous intéressent, ce calcul est réputé prohibitif.

3.5.2 Algorithmes quasi-Newton

Une alternative consiste à utiliser à l'étape k une matrice \mathbf{H}_k qui soit aisément calculable et se comporte de façon similaire à l'inverse du Hessien sur le sous-espace engendré par le gradient $\mathbf{g}(\mathbf{x}_k)$ également noté \mathbf{g}_k , de sorte que l'on puisse écrire $\mathbf{d}_k = -\mathbf{H}_k\mathbf{g}_k$. On prendra garde au changement de notation abrupt : \mathbf{H}_k désigne maintenant une approximation de l'inverse du Hessien.

L'idée consiste à construire \mathbf{H}_k au fur et à mesure des itérations. Les directions de \mathbb{R}^n qui sont explorées lorsque l'on calcule \mathbf{x}_{k+1} connaissant \mathbf{x}_k se réduisent à $\Delta\mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$,

et $\Delta \mathbf{g}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$, puisque l'on ne dispose que des \mathbf{x}_k et des gradients $\Delta \mathbf{g}_k$. Or une simple intégration montre que $\Delta \mathbf{x}_k$ et $\Delta \mathbf{g}_k$ sont reliés exactement par

$$\left\{ \int_0^1 d\lambda \nabla^2 J(\mathbf{x} + \lambda \mathbf{d}_k) \right\} \Delta \mathbf{x}_k = \Delta \mathbf{g}_k. \quad (3.49)$$

Et $\int_0^1 d\lambda \nabla^2 J(\mathbf{x} + \lambda \mathbf{d}_k)$ est la moyenne prise par le Hessien sur le trajet $\Delta \mathbf{x}_k$. La conclusion de cet argument est que l'on peut extraire de $\Delta \mathbf{g}_k$ et $\Delta \mathbf{x}_k$ une information sur le Hessien. Comme cette information est de rang faible, l'inverse du Hessien sera construit progressivement itération après itération.

À la lumière de cet argument heuristique, on souhaite que la suite \mathbf{H}_k satisfasse

$$\forall k \geq 0, \quad \mathbf{H}_k (\mathbf{g}_{k+1} - \mathbf{g}_k) = \mathbf{x}_{k+1} - \mathbf{x}_k. \quad (3.50)$$

Cette condition fait de l'algorithme une méthode dite *quasi-Newton*. On veut aussi que \mathbf{H}_k soit positive. En effet on a

$$\varphi'(0) = -(\mathbf{g}_k, \mathbf{d}_k) = -(\mathbf{g}_k, \mathbf{H}_k \mathbf{g}_k), \quad (3.51)$$

de sorte que la positivité de \mathbf{H}_k suffit à assurer la condition de descente.

Les méthodes quasi-Newton remportent les suffrages des numériciens, puisqu'elles combinent vitesse de convergence (sophistication) tout en réduisant la complexité algorithmique (pas de calcul explicite du Hessien à chaque itération). L'algorithme générique d'une méthode quasi-Newton prend la forme donnée en Fig. 3.5.2.

Algorithme générique de Quasi-Newton

1. On part de $k = 0$ et on choisit \mathbf{x}_0 , point de départ, ainsi qu'une matrice symétrique \mathbf{H}_0 .
2. Pour $k = 0, 1, 2, \dots$,
 - Si $\mathbf{g}_k = 0$, l'algorithme se termine sinon

$$\mathbf{d}_k = -\mathbf{H}_k \mathbf{g}_k$$

- On calcule

$$\lambda_k = \arg \min J(\mathbf{x}_k - \lambda \mathbf{d}_k) \quad \text{pour } \lambda \geq 0$$

puis

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \lambda_k \mathbf{d}_k$$

- Construire \mathbf{H}_{k+1} à l'aide de \mathbf{H}_k et

$$\Delta \mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k \quad \text{et} \quad \Delta \mathbf{g}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$$

La suite \mathbf{H}_k doit satisfaire à la condition quasi-Newton, mais cela ne suffit pas à la définir complètement. C'est la prescription complète de cette suite qui différencie les variantes des méthodes de type quasi-Newton.

Application à une fonction quadratique

Voyons plus en détails comment une méthode quasi-Newton fonctionne dans le cas linéaire, c'est-à-dire lorsque la fonction de coût est quadratique $J_{\mathbf{Q}}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}, \mathbf{Q}\mathbf{x}) + (\mathbf{b}, \mathbf{x})$. L'algorithme générique d'une méthode quasi-Newton dans le cas linéaire prend la forme donnée en Fig. 3.5.2.

Algorithme générique de Quasi-Newton (cas linéaire)

1. On part de $k = 0$ et on choisit \mathbf{x}_0 , point de départ, ainsi qu'une matrice symétrique \mathbf{H}_0 .
2. Pour $k = 0, 1, 2, \dots$,
 - Si $\mathbf{g}_k = 0$, l'algorithme se termine sinon

$$\mathbf{d}_k = -\mathbf{H}_k \mathbf{g}_k$$

- On calcule

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \lambda_k \mathbf{d}_k \quad \text{où} \quad \lambda_k = \frac{(\mathbf{g}_k, \mathbf{d}_k)}{(\mathbf{d}_k, \mathbf{Q}\mathbf{d}_k)}$$

- Construire \mathbf{H}_{k+1} à l'aide de \mathbf{H}_k et

$$\Delta \mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k \quad \text{et} \quad \Delta \mathbf{g}_k = \mathbf{g}_{k+1} - \mathbf{g}_k$$

La valeur de λ_k est le résultat de la minimisation de

$$J(\mathbf{x}_k + \lambda \mathbf{d}_k) = \frac{1}{2}(\mathbf{x}_k - \lambda \mathbf{d}_k, \mathbf{Q}(\mathbf{x}_k - \lambda \mathbf{d}_k)) + (\mathbf{b}, \mathbf{x}_k - \lambda \mathbf{d}_k), \quad (3.52)$$

de solution immédiate

$$\lambda_k = \frac{(\mathbf{g}_k, \mathbf{d}_k)}{(\mathbf{d}_k, \mathbf{Q}\mathbf{d}_k)},$$

Q-Conjugaison

Une propriété remarquable d'un algorithme de quasi-Newton générique *appliqué à la fonctionnelle quadratique* est que les directions \mathbf{d}_k sont mutuellement *Q*-conjuguées.

Complétude

3.5.3 Une méthode quasi-Newton : l'algorithme BFGS

La méthode quasi-Newton considérée en pratique comme la plus efficace a été proposée en 1970 par Broyden, Fletcher, Goldfarb et Shanno, d'où son nom de méthode BFGS. La

suite des matrices \mathbf{H}_k est définie par la formule matricielle de récurrence

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \left(1 + \frac{(\Delta \mathbf{g}_k, \mathbf{H}_k \Delta \mathbf{g}_k)}{(\Delta \mathbf{g}_k, \Delta \mathbf{x}_k)} \right) \frac{\Delta \mathbf{g}_k (\Delta \mathbf{g}_k)^\top}{(\Delta \mathbf{x}_k, \Delta \mathbf{g}_k)} - \frac{\mathbf{H}_k \Delta \mathbf{g}_k (\Delta \mathbf{x}_k)^\top + (\mathbf{H}_k \Delta \mathbf{g}_k (\Delta \mathbf{x}_k)^\top)^\top}{(\Delta \mathbf{g}_k, \Delta \mathbf{x}_k)}. \quad (3.53)$$

$$\mathbf{H}_{k+1} = \mathbf{H}_k + \left(1 - \frac{\Delta \mathbf{x}_k \Delta \mathbf{g}_k^\top}{(\Delta \mathbf{g}_k, \Delta \mathbf{x}_k)} \right) \mathbf{H}_k \left(1 - \frac{\Delta \mathbf{g}_k (\Delta \mathbf{x}_k)^\top}{(\Delta \mathbf{g}_k, \Delta \mathbf{x}_k)} \right) + \frac{\Delta \mathbf{x}_k (\Delta \mathbf{x}_k)^\top}{(\Delta \mathbf{g}_k, \Delta \mathbf{x}_k)}. \quad (3.54)$$

C'est une méthode dite de rang deux car

$$\text{rang}(\mathbf{H}_{k+1} - \mathbf{H}_k) \leq 2. \quad (3.55)$$

On vérifie sans peine que c'est bien une méthode de quasi-Newton, et que de plus si \mathbf{H}_k est définie positive, alors \mathbf{H}_{k+1} l'est aussi.

Il a été démontré que, sous réserve que le point de départ \mathbf{x}_0 ne soit pas trop éloigné de l'argument $\bar{\mathbf{x}}$ du minimum de la fonction de coût, la convergence de l'algorithme BFGS, dans le cas non-linéaire, est superlinéaire, c'est-à-dire

$$\lim_{k \rightarrow \infty} \frac{\|\mathbf{x}_k - \bar{\mathbf{x}}\|}{\|\mathbf{x}_{k+1} - \bar{\mathbf{x}}\|} = 0. \quad (3.56)$$

Sous des contraintes plus fortes, on montre même que c'est toute une classe d'algorithmes de quasi-Newton qui convergent de façon quadratique, même dans le cas non-linéaire.

3.6 Contrainte de modèle faible

L'analogie variationnel du filtre de Kalman comportant de l'erreur modèle est donné par la formulation dite *faible* du 4D-Var. La fonction coût devient :

$$J(\mathbf{x}) = \frac{1}{2} \sum_{k=0}^n (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k, (\mathbf{R}_k)^{-1} (\mathbf{y}_k - \mathbf{H}_k \mathbf{x}_k)) \quad (3.57)$$

$$+ \frac{1}{2} \sum_{k=0}^n (\mathbf{x}_{k+1} - \mathbf{M}_{k+1} \mathbf{x}_k, (\mathbf{Q}_k)^{-1} (\mathbf{x}_{k+1} - \mathbf{M}_{k+1} \mathbf{x}_k)). \quad (3.58)$$

3.7 Méthode incrémentale

Deuxième partie

Assimilation de données avancée

Dans cette partie, on décrit ce que devrait être le cadre formel idéal de l'assimilation de données séquentielle (formule de Bayes). Son application bute sur deux problèmes majeurs. D'une part, la non-linéarité du modèle physique n'est représenté qu'approximativement et plusieurs filtres d'assimilation ont été proposés pour remédier à ce problème : filtre d'ensemble, filtre "unscented". D'autre part, en géophysique, les systèmes sont de très grande dimensionalité et la représentation des matrices de covariances d'erreur, sur lesquelles se fonde le filtre de Kalman, devient impossible. Aussi plusieurs filtres mettant en oeuvre des techniques de réduction et de représentation des erreurs ont été proposés : filtres RRSQRT, SEEK, SEIK.

Chapitre 4

Assimilation de données non-linéaire

On illustrera la performance des algorithmes décrits dans ce chapitre sur un oscillateur non-linéaire comportant une dose d'erreur modèle.

Pour cela, nous reprenons l'exemple de l'oscillateur non-linéaire du chapitre 2, auquel nous adjoignons de l'erreur modèle. Cette dernière accentue le caractère imprévisible du système :

$$x_0 = 0, \quad x_1 = 1 \quad \text{et pour } 1 \leq k \leq N \quad x_{k+1} - 2x_k + x_{k-1} = \omega^2 x_k + \lambda^2 x_k^3 + \xi_k. \quad (4.1)$$

Les paramètres retenus ici sont : $\omega = 0.035$, $\lambda = 3.10^{-5}$. La variance de l'erreur modèle est de 0.0025, alors que la variance de l'erreur d'observation est de 49. La condition initiale est $x_0 = 0$ et $x_1 = 1$, puis le système est itéré 10000 fois.

4.1 Les limites du filtre de Kalman étendu

Voyons maintenant le comportement caractéristique d'un filtre de Kalman étendu appliqué à un système fortement non-linéaire.

Prenons donc l'exemple de l'oscillateur anaharmonique stochastique décrit ci-dessus. Dans un premier temps, on suppose que la fréquence des observations est d'une mesure toutes les 50 itérations. Cela conduit au résultat représenté en figure 4.1. qui atteste de la performance du filtre sur cette période de temps.

Dans un second temps, on suppose que la fréquence des observations est de une mesure toutes les 58 itérations. Cela conduit au résultat représenté en figure 4.2. On constate que le filtre finit par diverger, car il est insuffisamment renseigné sur l'état du système.

Il existe des filtres de Kalman étendus d'ordre supérieur (en particulier d'ordre 2) qui permettent de mieux rendre compte des non-linéarités du modèle. Cependant ceux-ci font apparaître naturellement le tangent linéaire du modèle, mais aussi le hessien de celui-ci, ce qui est très coûteux à mettre en œuvre (en mémoire et temps CPU). Cela limite l'intérêt de ces extensions pour des systèmes de grande taille.

Pour comprendre et tenter d'améliorer la performance des filtres sur des systèmes dynamiques très non-linéaires, il est nécessaire de définir ce que devrait être le système d'assimilation qui rende compte de tous les moments des distributions statistiques. En effet une dynamique non-linéaire propage les moments de façon non-triviale à la différence d'une dynamique linéaire.

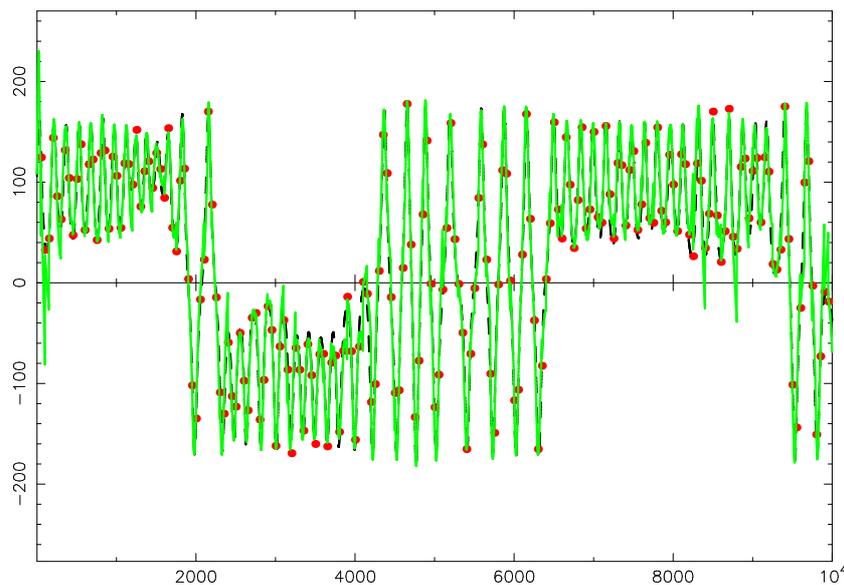


FIGURE 4.1 – Assimilation d’un oscillateur anharmonique perturbé, à l’aide d’un filtre de Kalman étendu. La fréquence des observations est de 50 itérations pour une observation.

4.2 Interprétation probabiliste et interpolation optimale

Jusqu’à présent on s’est posé le problème suivant : compte tenu d’une connaissance a priori mais imparfaite du système et d’un jeu plus récent d’observations, quel est, au sens des moindres carrés, l’état du système le plus probable et quelle est la variance de l’erreur commise ?

On peut cependant se poser un tel problème en termes plus généraux.

Le même problème peut en effet être formulé en termes de densité de probabilité. Autrement dit, on ne se contente plus de travailler sur l’espérance et les covariances des distributions de probabilité des variables mais (a priori) sur tous les moments [22].

Connaissant la densité de probabilité de l’ébauche et de celle des observations, quelle est la densité de probabilité pour le système d’être dans tel ou tel état ? Contrairement à l’approche poursuivie dans les précédents chapitres, il n’y a pas de critère fourni a priori (les moindres carrés auparavant) dont on puisse tirer un estimateur de l’état du système et de l’erreur commise. On se donne a posteriori un critère définissant la notion d’optimalité, dont on déduit un estimateur. Cet estimateur varie selon le critère retenu.

4.2.1 Assimilation d’une observation

On considère l’équation de la mesure la plus générale

$$\mathbf{y} = H(\mathbf{x}, \mathbf{v}), \quad (4.2)$$

où l’opérateur d’observation H n’est donc pas nécessairement linéaire. Ce que l’on suppose connaître a priori, c’est la distribution de probabilité de \mathbf{x} , c’est-à-dire la densité de probabilité de son ébauche, que l’on note $p_X(\mathbf{x})$. On suppose aussi connue la distribution

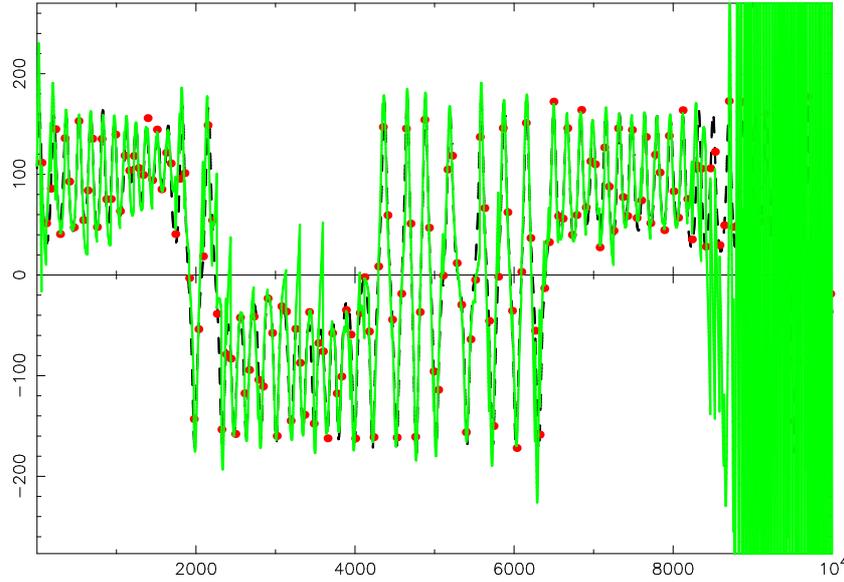


FIGURE 4.2 – Assimilation d’un oscillateur anharmonique perturbé, à l’aide d’un filtre de Kalman étendu. La fréquence d’assimilation des observations est insuffisante (58 itérations pour une observation) et le filtre diverge.

de probabilité du bruit, notée $p_V(\mathbf{v})$. Ce qu’il est important d’obtenir à partir de ces distributions, c’est la densité de probabilité de \mathbf{x} , connaissant la valeur de l’observation \mathbf{y} , c’est-à-dire la probabilité conditionnelle de \mathbf{x} , connaissant \mathbf{y} , notée $p_{X|Y}(\mathbf{x}|\mathbf{y})$. Utilisant la règle de Bayes, on a la relation

$$p_{X|Y}(\mathbf{x}|\mathbf{y}) = p_{Y|X}(\mathbf{y}|\mathbf{x}) \frac{p_X(\mathbf{x})}{p_Y(\mathbf{y})}. \quad (4.3)$$

p_X , on l’a vu, est connu. D’autre part, il est équivalent de connaître $p_{Y|X}$ ou de connaître p_V , les deux distributions étant directement reliée par la relation d’observation $\mathbf{y} = H(\mathbf{x}, \mathbf{v})$. Reste à obtenir p_Y . Or on a

$$p_Y(\mathbf{y}) = \int d\mathbf{x} p_{Y|X}(\mathbf{y}|\mathbf{x}) p_X(\mathbf{x}), \quad (4.4)$$

ce qui détermine p_Y . On remarque donc que le calcul de ce dernier terme correspond à une normalisation de la distribution de probabilité par rapport à \mathbf{x} . En pratique, il n’est pas toujours nécessaire de le calculer, car il ne dépend pas de \mathbf{x} .

4.2.2 Théorie de l’estimation et analyse BLUE

Pour nous convaincre de l’intérêt de ce point de vue probabiliste, commençons par vérifier qu’il corrobore l’analyse de l’interpolation optimale linéaire (BLUE). Pour cela, l’équation d’observation est simplifiée en

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{v}, \quad (4.5)$$

où l'opérateur d'observation \mathbf{H} est ici linéaire. On suppose que \mathbf{x} est distribuée selon une loi normale $N(\bar{\mathbf{x}}, \mathbf{P})$, c'est-à-dire selon la densité de probabilité (normalisée donc)

$$p_X(\mathbf{x}) = \frac{1}{(2\pi)^{n/2} |\mathbf{P}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{P}^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right), \quad (4.6)$$

avec $|\mathbf{P}|$ le déterminant de la matrice carrée \mathbf{P} . Cela constitue donc une ébauche de la densité de probabilité sur \mathbf{x} . Sachant que \mathbf{v} est distribuée selon la loi normale $N(\mathbf{0}, \mathbf{R})$, indépendante de celle de \mathbf{x} , la probabilité conditionnelle de \mathbf{y} , connaissant \mathbf{x} est

$$p_{Y|X}(\mathbf{y}|\mathbf{x}) = \frac{1}{(2\pi)^{p/2} |\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x})\right). \quad (4.7)$$

Le calcul de p_Y (normalisation du produit des deux densités précédentes) correspond à la composition de deux densités gaussiennes. La densité obtenue est donc aussi gaussienne en \mathbf{y} . Il n'est pas difficile de calculer son espérance $\bar{\mathbf{x}}_y = E[\mathbf{y}]$ et sa variance \mathbf{P}_y . En effet

$$\begin{aligned} \bar{\mathbf{x}}_y &= E[\mathbf{H}\mathbf{x} + \mathbf{v}] = \mathbf{H}\bar{\mathbf{x}}, \\ \mathbf{P}_y &= E[(\mathbf{y} - \bar{\mathbf{x}}_y)(\mathbf{y} - \bar{\mathbf{x}}_y)^T] = E[(\mathbf{H}(\mathbf{x} - \bar{\mathbf{x}}) + \mathbf{v})(\mathbf{H}(\mathbf{x} - \bar{\mathbf{x}}) + \mathbf{v})^T] \\ &= \mathbf{H}E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T] \mathbf{H}^T + E[\mathbf{v}\mathbf{v}^T] = \mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}. \end{aligned} \quad (4.8)$$

Cela caractérise donc complètement la distribution de \mathbf{y}

$$p_Y(\mathbf{y}) = \frac{1}{(2\pi)^{p/2} |\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}})^T (\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}})\right). \quad (4.9)$$

En utilisant (4.9), (4.7) et la règle de Bayes (4.3), on obtient la probabilité conditionnelle recherchée

$$\begin{aligned} p_{X|Y}(\mathbf{x}|\mathbf{y}) &= \frac{|\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}|^{1/2}}{(2\pi)^{n/2} |\mathbf{P}|^{1/2} (2\pi)^{p/2} |\mathbf{R}|^{1/2}} \times \\ &\exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}) - \frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^T \mathbf{P}^{-1}(\mathbf{x} - \bar{\mathbf{x}}) \right. \\ &\left. + \frac{1}{2}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}})^T (\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\bar{\mathbf{x}})\right). \end{aligned} \quad (4.10)$$

En factorisant l'argument de l'exponentielle selon \mathbf{x} , on obtient

$$p_{X|Y}(\mathbf{x}|\mathbf{y}) = \frac{|\mathbf{H}\mathbf{P}\mathbf{H}^T + \mathbf{R}|^{1/2}}{(2\pi)^{n/2} |\mathbf{P}|^{1/2} (2\pi)^{p/2} |\mathbf{R}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^T \mathbf{P}_*^{-1}(\mathbf{x} - \mathbf{x}^*)\right), \quad (4.11)$$

avec

$$\mathbf{P}_*^{-1} = \mathbf{P}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad \text{et} \quad \mathbf{x}^* = \mathbf{P}_x(\mathbf{H}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{P}^{-1} \bar{\mathbf{x}}). \quad (4.12)$$

4.2.3 Choix d'un estimateur

On peut alors choisir un *estimateur*. C'est-à-dire, donner un sens rigoureux à la notion subjective de valeur *la plus vraisemblable* d'une analyse portant sur l'état du système \mathbf{x} , supposées connues les observations \mathbf{y} . Cet estimateur est à définir à partir de la distribution de probabilité conditionnelle $p_{X|Y}$. On le souhaite en général sans biais. Dans le

cas présent, parce que les distributions sont gaussiennes, tous les choix raisonnables conduisent au même estimateur. Toutefois sur une distribution plus générale, on distinguerait l'estimateur de *variance minimum* et l'estimateur *a posteriori de probabilité maximum* (correspondant donc au maximum de la distribution $p_{X|Y}$). Mais il y en a d'autres possibles.

Dans le cas présent, les deux se confondent en l'estimateur $\hat{\mathbf{x}} = E[\mathbf{x}|y]$, qu'on lit simplement sur la distribution (4.11). Si l'on utilise maintenant la valeur du gain \mathbf{K}^* de l'analyse optimale, on obtient que

$$\begin{aligned}\hat{\mathbf{x}} &= (\mathbf{P}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{y} + \mathbf{P}^{-1} \bar{\mathbf{x}}) \\ &= \mathbf{K}^* \mathbf{y} + (\mathbf{P}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{P}^{-1} \bar{\mathbf{x}} \\ &= \mathbf{K}^* \mathbf{y} + (\mathbf{I} - \mathbf{K}^* \mathbf{H}) \bar{\mathbf{x}} \\ &= \bar{\mathbf{x}} + \mathbf{K}^* (\mathbf{y} - \mathbf{H} \bar{\mathbf{x}}),\end{aligned}\tag{4.13}$$

où on a utilisé les relations (1.22) et (1.20). Ce résultat est à comparer à (1.14). Le résultat de ce calcul probabiliste coïncide donc avec le résultat de l'analyse BLUE, avec l'identification $\mathbf{x}_b = \bar{\mathbf{x}}$.

4.3 Interprétation probabiliste et assimilation séquentielle

On se place maintenant dans le cadre d'une analyse séquentielle. Et on souhaite mener une étude probabiliste similaire à la précédente. Il s'agit du problème du *filtrage bayésien*. On va noter par \mathbf{Z}_k , l'ensemble des observations passées depuis t_0 jusqu'à t_k , c'est-à-dire

$$\mathbf{Z}_k = \{\mathbf{z}_k, \mathbf{z}_{k-1}, \dots, \mathbf{z}_0\}.\tag{4.14}$$

La densité que l'on souhaite calculer est la probabilité conditionnelle de \mathbf{x}_k connaissant \mathbf{Z}_k , notée $p(\mathbf{x}_k | \mathbf{Z}_k)$. Pour plus de concision, on abandonne ici l'indice descriptif des distributions, qui sont donc toutes notées p , et qui ne se distinguent plus que par leurs arguments.

4.3.1 Étape de la prévision

Dans l'étape de prévision, on cherche à calculer la distribution de probabilité $p(\mathbf{x}_{k+1} | \mathbf{Z}_k)$, connaissant la densité $p(\mathbf{x}_k | \mathbf{Z}_k)$. Sans hypothèse particulière, on a donc

$$p(\mathbf{x}_{k+1} | \mathbf{Z}_k) = \int d\mathbf{x}_k p(\mathbf{x}_{k+1} | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{Z}_k).\tag{4.15}$$

Une équation très générale pour la modélisation stochastique de la dynamique du système est

$$\mathbf{x}_{k+1} = F(\mathbf{x}_k, \mathbf{w}_k),\tag{4.16}$$

où \mathbf{w}_k est supposée être un bruit blanc (décorrélé en temps) non-nécessairement gaussien de densité p_W . On a donc

$$p(\mathbf{x}_{k+1} | \mathbf{x}_k) = \int d\mathbf{w}_k p_W(\mathbf{w}_k) \delta(\mathbf{x}_{k+1} - F(\mathbf{x}_k, \mathbf{w}_k)),\tag{4.17}$$

où δ est la distribution de Dirac. L'évolution des densités entre les instants t_k et t_{k-1} est alors donnée par

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \int d\mathbf{x}_k d\mathbf{w}_k p_W(\mathbf{w}_k) p(\mathbf{x}_k|\mathbf{Z}_k) \delta(\mathbf{x}_{k+1} - F(\mathbf{x}_k, \mathbf{w}_k)). \quad (4.18)$$

Lorsque le bruit est additif, c'est-à-dire si (4.16) se réduit à

$$\mathbf{x}_{k+1} = M(\mathbf{x}_k) + \mathbf{w}_k, \quad (4.19)$$

la convolution des densités peut se simplifier en

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \int d\mathbf{x}_k p_W(\mathbf{x}_{k+1} - M(\mathbf{x}_k)) p(\mathbf{x}_k|\mathbf{Z}_k). \quad (4.20)$$

Dans le cas particulier où l'erreur modèle est nulle, on peut poursuivre puisqu'alors

$$p_W(\mathbf{x}_{k+1} - M(\mathbf{x}_k)) = \delta(\mathbf{x}_{k+1} - M(\mathbf{x}_k)). \quad (4.21)$$

Dans ce cas

$$p(\mathbf{x}_{k+1}|\mathbf{Z}_k) = \frac{p((M_k)^{-1}(\mathbf{x}_{k+1})|\mathbf{Z}_k)}{|(M_k') \circ M_k^{-1}(\mathbf{x}_{k+1})|}, \quad (4.22)$$

où les barres $|\mathcal{O}|$ désignent le déterminant de l'opérateur \mathcal{O} .

4.3.2 Étape de l'analyse

Une analyse est effectuée dès qu'une nouvelle observation est acquise. De façon générale, l'équation d'observation est

$$\mathbf{z}_k = H(\mathbf{x}_k, \mathbf{v}_k), \quad (4.23)$$

où \mathbf{v}_k est un bruit de distribution quelconque. L'utilisation de la règle de Bayes conduit à

$$p(\mathbf{x}_k|\mathbf{Z}_k) = p(\mathbf{x}_k|\mathbf{z}_k, \mathbf{Z}_{k-1}) = p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{k-1}) \frac{p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{p(\mathbf{z}_k|\mathbf{Z}_{k-1})}. \quad (4.24)$$

C'est une version légèrement plus élaborée de la formule (4.3). $p(\mathbf{x}_k|\mathbf{Z}_{k-1})$ est supposée connue puisqu'il s'agit de la distribution issue du résultat de la prévision de t_{k-1} à t_k . On remarque que l'on a

$$p(\mathbf{z}_k|\mathbf{Z}_{k-1}) = \int d\mathbf{x}_k p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{k-1}) p(\mathbf{x}_k|\mathbf{Z}_{k-1}), \quad (4.25)$$

qui s'apparente donc à une normalisation de la distribution de probabilité recherchée par rapport à \mathbf{x}_k . En résumé

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{k-1}) p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{\int d\mathbf{x}_k p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{k-1}) p(\mathbf{x}_k|\mathbf{Z}_{k-1})}. \quad (4.26)$$

Si l'on suppose que le bruit \mathbf{v}_k est blanc alors $p(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{k-1}) = p(\mathbf{z}_k|\mathbf{x}_k)$, puisque la connaissance de $p(\mathbf{z}_k|\mathbf{x}_k)$ est équivalent à celle du bruit blanc au temps t_k , qui est donc indépendant de \mathbf{Z}_{k-1} . Dans ce cas, la formule précédente se simplifie en

$$p(\mathbf{x}_k|\mathbf{Z}_k) = \frac{p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Z}_{k-1})}{\int d\mathbf{x}_k p(\mathbf{z}_k|\mathbf{x}_k) p(\mathbf{x}_k|\mathbf{Z}_{k-1})}. \quad (4.27)$$

4.3.3 Théorie de l'estimation et filtre de Kalman linéaire

On se place dans le cadre des hypothèses formulées pour le filtre de Kalman *linéaire*. On cherche à montrer que dans ces conditions, la théorie de l'estimation fournit effectivement les résultats du filtre de Kalman linéaire.

Étape de la prévision : L'équation du modèle (transition de t_k à t_{k+1}) est

$$\mathbf{x}_{k+1} = \mathbf{M}_k \mathbf{x}_k + \mathbf{w}_k, \quad (4.28)$$

où \mathbf{w}_k représente un bruit blanc gaussien de covariance \mathbf{Q}_k . On a par hypothèse

$$p_W(\mathbf{x}_{k+1} - \mathbf{M}_k \mathbf{x}_k) = \frac{1}{(2\pi)^{n/2} |\mathbf{Q}_k|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}_{k+1} - \mathbf{M}_k \mathbf{x}_k)^T (\mathbf{Q}_k)^{-1} (\mathbf{x}_{k+1} - \mathbf{M}_k \mathbf{x}_k) \right), \quad (4.29)$$

d'une part et

$$p(\mathbf{x}_k | \mathbf{Z}_k) = \frac{1}{(2\pi)^{n/2} |\mathbf{P}_k^a|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}_k - \mathbf{x}_k^a)^T (\mathbf{P}_k^a)^{-1} (\mathbf{x}_k - \mathbf{x}_k^a) \right), \quad (4.30)$$

d'autre part, avec \mathbf{x}_k^a et \mathbf{P}_k^a le résultat de l'analyse au temps t_k . Il reste à calculer l'intégrale (4.20). On est à nouveau dans la situation où l'on doit effectuer la convolution de deux gaussiennes. Le calcul est formellement identique à (4.6), (4.7) et (4.9). Ce qui conduit à

$$p(\mathbf{x}_{k+1} | \mathbf{Z}_k) = \frac{1}{(2\pi)^{n/2} |\mathbf{P}_{k+1}^f|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{x}_{k+1} - \mathbf{x}_{k+1}^f)^T (\mathbf{P}_{k+1}^f)^{-1} (\mathbf{x}_{k+1} - \mathbf{x}_{k+1}^f) \right), \quad (4.31)$$

avec

$$\mathbf{x}_{k+1}^f = \mathbf{M}_k \mathbf{x}_k^a \quad \text{et} \quad \mathbf{P}_{k+1}^f = \mathbf{M}_k^T \mathbf{P}_k^a \mathbf{M}_k + \mathbf{Q}_k. \quad (4.32)$$

Étape de l'analyse : Il n'y a rien de plus à prouver pour l'étape de l'analyse, puisque cette partie de la preuve est identique à l'obtention de l'analyse BLUE par la théorie de l'estimation, ce qui a été décrit dans la section 4.2.2.

On retrouve donc bien, dans le cas linéaire et de statistique de bruits gaussienne, que la théorie de l'estimation redonne le filtre de Kalman linéaire. Un estimateur possible est celui du maximum de probabilité a posteriori.

4.4 Interprétation probabiliste et assimilation variationnelle

On souhaite, en utilisant l'interprétation probabiliste, montrer que le système d'assimilation non-linéaire

$$\begin{aligned} \mathbf{x}_{k+1} &= M_{k+1}(\mathbf{x}_k) + \mathbf{w}_k \\ \mathbf{z}_k &= H_k(\mathbf{x}_k) + \mathbf{v}_k, \end{aligned} \quad (4.33)$$

sur l'intervalle de temps $[t_0, t_N]$ accepte une formulation variationnelle 4D-Var. Les opérateurs M_k et H_k ne sont pas nécessairement linéaires. Enfin le bruit de type erreur-modèle

\mathbf{w}_k et les erreurs d'observation \mathbf{v}_k sont supposés blanc et gaussiens d'espérances nulles et de matrices de covariance \mathbf{Q}_k et \mathbf{R}_k respectivement. Leur distribution est donc dans l'ordre

$$p_W(\mathbf{w}_k) = N(\mathbf{0}, \mathbf{Q}_k) \quad \text{et} \quad p_V(\mathbf{v}_k) = N(\mathbf{0}, \mathbf{R}_k). \quad (4.34)$$

La distribution de probabilité pertinente n'est pas celle utilisée pour l'assimilation séquentielle mais plutôt $p(\mathbf{X}_k | \mathbf{Z}_k)$ (problème du *lissage bayésien*), avec

$$\mathbf{X}_k = \{\mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0\}. \quad (4.35)$$

Cela nécessite de reprendre partiellement les calculs probabilistes de l'approche séquentielle. Tout d'abord la loi de Bayes donne

$$p(\mathbf{X}_k | \mathbf{Z}_k) = p(\mathbf{Z}_k | \mathbf{X}_k) \frac{p(\mathbf{X}_k)}{p(\mathbf{Z}_k)}. \quad (4.36)$$

Puisque seule la dépendance de $p(\mathbf{X}_k | \mathbf{Z}_k)$ en \mathbf{X}_k nous intéresse, on se contentera dans la suite de $p(\mathbf{X}_k | \mathbf{Z}_k) \propto p(\mathbf{Z}_k | \mathbf{X}_k) p(\mathbf{X}_k)$. Il est clair que

$$p(\mathbf{X}_k) = p(\mathbf{x}_k, \mathbf{X}_{k-1}) = p(\mathbf{x}_k | \mathbf{X}_{k-1}) p(\mathbf{X}_{k-1}) = p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{X}_{k-1}), \quad (4.37)$$

où l'on a fait appel dans la dernière identité au fait que la suite de variables aléatoires \mathbf{X}_k est markovienne. Mais on peut également réduire $p(\mathbf{Z}_k | \mathbf{X}_k)$:

$$\begin{aligned} p(\mathbf{Z}_k | \mathbf{X}_k) &= p(\mathbf{z}_k, \mathbf{Z}_{k-1} | \mathbf{X}_k) = p(\mathbf{z}_k | \mathbf{X}_k) p(\mathbf{Z}_{k-1} | \mathbf{X}_k) \\ &= p(\mathbf{z}_k | \mathbf{x}_k, \mathbf{X}_{k-1}) p(\mathbf{Z}_{k-1} | \mathbf{x}_k, \mathbf{X}_{k-1}) \\ &= p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{Z}_{k-1} | \mathbf{X}_{k-1}). \end{aligned} \quad (4.38)$$

On en déduit que

$$\begin{aligned} p(\mathbf{X}_k | \mathbf{Z}_k) &\propto p(\mathbf{Z}_k | \mathbf{X}_k) p(\mathbf{X}_k) = p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{Z}_{k-1} | \mathbf{X}_{k-1}) p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{X}_{k-1}) \\ &\propto p(\mathbf{z}_k | \mathbf{x}_k) p(\mathbf{x}_k | \mathbf{x}_{k-1}) p(\mathbf{Z}_{k-1} | \mathbf{X}_{k-1}) p(\mathbf{X}_{k-1}) \\ &\propto p_V(\mathbf{z}_k - H(\mathbf{x}_k)) p_W(\mathbf{x}_k - M_k(\mathbf{x}_{k-1})) p(\mathbf{Z}_{k-1} | \mathbf{X}_{k-1}) p(\mathbf{X}_{k-1}). \end{aligned} \quad (4.39)$$

Cette relation de récurrence est suffisante pour déterminer la densité $p(\mathbf{X}_k | \mathbf{Z}_k)$ compte tenu des hypothèses sur les bruits.

$$\begin{aligned} p(\mathbf{X}_k | \mathbf{Z}_k) &\propto \left[\prod_{j=0}^k p_V(\mathbf{z}_j - H_j(\mathbf{x}_j)) \prod_{j=0}^{k-1} p_W(\mathbf{x}_{j+1} - M_{j+1}(\mathbf{x}_j)) \right] p(\mathbf{x}_0) \\ &\propto \exp \left[-\frac{1}{2} \sum_{j=0}^k (\mathbf{z}_j - H_j(\mathbf{x}_j))^T \mathbf{R}_j^{-1} (\mathbf{z}_j - H_j(\mathbf{x}_j)) \right. \\ &\quad \left. - \frac{1}{2} \sum_{j=0}^{k-1} (\mathbf{x}_{j+1} - M_{j+1}(\mathbf{x}_j))^T \mathbf{Q}_j^{-1} (\mathbf{x}_{j+1} - M_{j+1}(\mathbf{x}_j)) \right] p(\mathbf{x}_0). \end{aligned} \quad (4.40)$$

La densité de probabilité $p(\mathbf{x}_0)$ définit l'ébauche de l'état du système à l'instant t_0 . Alors, si l'on choisit pour estimateur le maximum de probabilité a posteriori, on doit minimiser l'argument de l'exponentielle. Cet argument n'est autre que la fonction de coût du formalisme 4D-Var. Cela justifie donc l'utilisation de cette fonctionnelle dans le formalisme 4D-Var dans un cadre d'une assimilation non-linéaire et étend les résultats obtenus dans le chapitre 3. De plus, cela constitue une manière différente de prouver l'équivalence du 4D-Var dans le cadre linéaire avec le filtre de Kalman linéaire.

4.5 Filtres particuliers (Monte Carlo)

Il est possible de construire des solutions numériques qui convergent vers la solution du problème de *filtrage bayésien*, c'est à dire de la solution des équations de filtrage exactes présentées en début de chapitre. Ces approches numériques s'inspirent des méthodes de Monte-Carlo. C'est-à-dire qu'une densité de probabilité est approximativement représentée par un échantillon discret tiré de cette densité. Plutôt que de calculer exactement la solution des équations de filtrage bayésien, on applique les transformations de ce filtrage (formule de Bayes pour l'analyse, et propagation par le modèle lors de l'étape de prévision) aux membres de l'échantillon. Les propriétés statistiques (moments) de l'échantillon sont alors censées être celles de la densité de probabilité filtrée. Naturellement cette approche par échantillonnage n'est exacte que dans la limite où le nombre de membres (ou particules) tend vers l'infini.

L'algorithme le plus populaire de type Monte-Carlo permettant de résoudre les équations du filtrage bayésien, est appelé *filtre particulière bootstrap* [7]. Il se décompose ainsi.

Échantillonnage Considérons un échantillon de particules $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$. La densité de probabilité à un instant t_k associée est $p_k(\mathbf{x}) : p_k(\mathbf{x}) \simeq \sum_{i=1}^M w_i^k \delta(\mathbf{x} - \mathbf{x}_k^i)$, qui se veut une approximation de la densité exacte qu'approche l'échantillon. À ce stade, on suppose les poids w_i^k uniformes $w_i^k = 1/M$.

Prévision À l'étape de prévision, les particules sont propagées de façon exacte par le modèle : $p_{k+1}(\mathbf{x}) \simeq \sum_{i=1}^M w_i^k \delta(\mathbf{x} - \mathbf{x}_{k+1}^i)$, avec $\mathbf{x}_{k+1}^i = M_{k+1}(\mathbf{x}_k)$. w_i^k est un scalaire qui pondère l'importance de la particule i dans l'ensemble.

Analyse L'analyse est remarquablement simple et élégante. L'application rigoureuse de la formule de Bayes conduit à associer à chaque particule un poids statistique correspondant à la vraisemblance de la particule relativement aux données. Ainsi pour chaque poids de chaque particule

$$w_{k+1,+}^i \propto w_{k+1,-}^i p(\mathbf{y}_{k+1} | \mathbf{x}_{k+1}^i). \quad (4.41)$$

Re-échantillonnage Malheureusement, ces poids statistiques ont potentiellement une très forte amplitude. Pire, au fur et à mesure du filtrage, une particule (trajectoire du modèle) est sélectionnée au détriment des autres. Si bien qu'un seul poids domine ($w_i \lesssim 1$) tandis que tous les autres sont pratiquement nuls. Le filtre particulière devient alors inefficace dans son estimation statistique. Ce phénomène est appelé *dégénérescence*. Un exemple d'une telle dégénérescence est donné sur la figure Fig.4.3, où la statistique du maximum des poids est étudiée sur le cas d'un modèle jouet météorologique à 40 et 80 variables. Dans un cas de dégénérescence, le maximum atteindra fréquemment 1, alors que dans un cas équilibré (propre à l'estimation), les valeurs élevées des poids seront plus rares.

Une manière de pallier partiellement ce défaut est d'effectuer un re-échantillonnage des particules en retirant un échantillon avec des poids homogènes à partir de la distribution en dégénérescence. Après re-échantillonnage, les particules ont toutes le même poids : $w_i^k = 1/M$.

Le schéma de principe du filtre bootstrap est représenté sur la figure Fig.4.4.

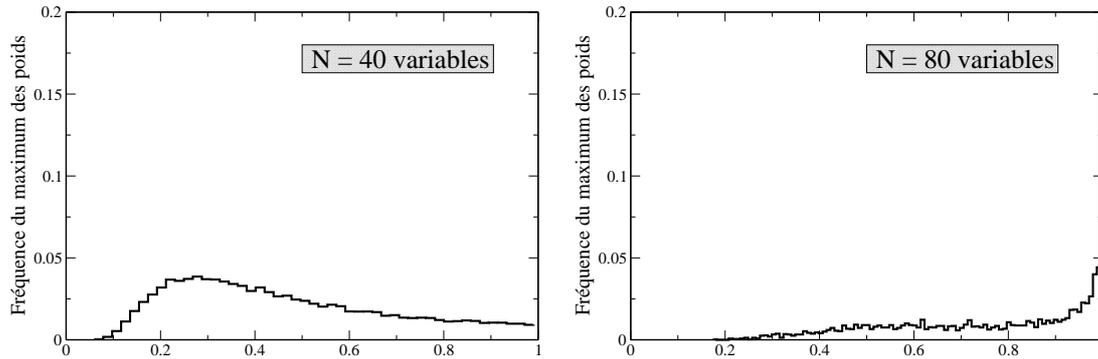


FIGURE 4.3 – Sur la gauche : distribution statistique des poids d’un filtre bootstrap dans un cas équilibré. Le système physique est un modèle de Lorenz 1996 à 40 variables [13]. Sur la droite : le même filtre particulière est appliqué à un modèle de Lorenz 1996, mais à 80 variables. Les poids dégénèrent.

Le filtre particulaire est très efficace sur des modèles fortement non-linéaires en faible dimension. Malheureusement, il est inadapté aux systèmes de grande taille dès que la dimension du système dépasse 8. On est donc très loin d’une situation confortable sur les systèmes géophysiques de grande taille. On montre ainsi qu’une absence de dégénérescence requiert un très grand nombre de particules. Ce nombre croît exponentiellement avec la dimension de l’espace des états !

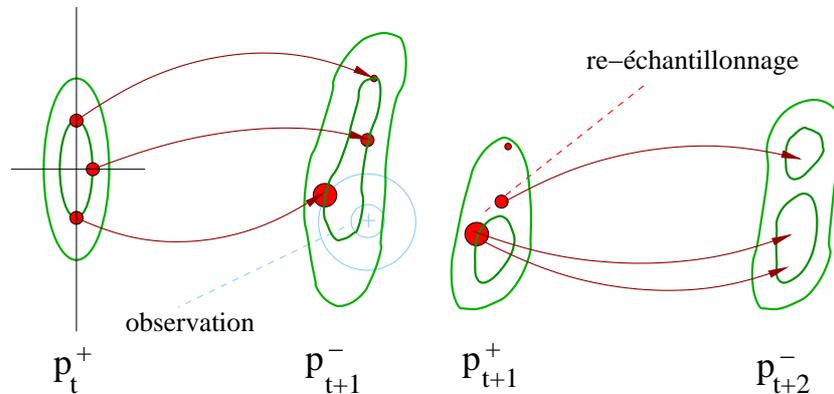


FIGURE 4.4 – Schéma de principe du filtre particulaire bootstrap.

Dans l’étape de prévision, il est en réalité crucial d’introduire une perturbation stochastique des états. En effet au fur et à mesure des re-échantillonnage l’ensemble des particules va s’appauvrir. Pour l’enrichir, il est donc nécessaire de perturber stochastiquement (ou d’une autre façon raisonnable) les états des particules.

Enfin, s’il semble difficile d’appliquer en l’état le filtre particulaire à des systèmes géophysiques de grande taille, des tests ont été effectués en océanographie [24] qui sont prometteurs pour le long terme.

Chapitre 5

Filtres : mieux prendre en compte la non-linéarité

Comme on l'a mentionné, pour améliorer la performance des filtres utilisés pour l'assimilation séquentielle, on peut :

- réduire la dimensionalité du système physique.
- mieux prendre en compte les non-linéarités du modèle physique.

Dans un cadre variationnel les non-linéarités du modèle sont prises en compte naturellement. Cela ne peut être le cas dans l'approche séquentielle, car il est nécessaire de transporter (par le modèle d'évolution) les statistiques d'erreur de façon approchée. Dans le cas des filtres dérivés de Kalman, il s'agit de propager les matrices de covariances d'erreur. On a vu que la stratégie du filtre étendu consiste à appliquer aux matrices de covariance l'opérateur de transfert linéarisé.

Récemment, des réponses plus évoluées ont été non seulement proposées, mais aussi implémentées opérationnellement. Les applications les plus nombreuses ont été entreprises pour l'océanographie, avant que les météorologues n'y concentrent davantage leurs efforts.

5.1 Le filtre de Kalman d'ensemble

Le filtre de Kalman d'ensemble a été proposé par Geir Evensen en 1994, puis corrigé en 1998 (voir par exemple [6]). Sa justification est semi-empirique et parfois déroutante. Néanmoins, il a fait preuve d'une grande efficacité sur des problèmes réels.

Le filtre d'ensemble est d'abord un filtre de Kalman, c'est-à-dire qu'il gère les statistiques d'erreur jusqu'à l'ordre 2, mais pas au delà. Le filtre de Kalman d'ensemble n'est donc pas un filtre particulière comme vu dans le chapitre 4. C'est un filtre gaussien¹.

Cependant il emprunte la notion de particule au filtre particulière (comme d'autres filtres : unscented, SEIK). Plutôt que de propager une matrice de covariance, les erreurs sont représentés statistiquement par un nuage de points. Ces points sont propagés grâce au modèle, sans linéarisation aucune.

L'étape d'analyse est celle d'un filtre de Kalman standard.

1. On peut montrer que dans la limite d'un grand nombre de particules, le filtre de Kalman d'ensemble ne résoud pas le problème du filtrage Bayésien tel qu'exposé au chapitre 4, contrairement au filtre particulière.

Il existe toutefois une modification un peu mystérieuse (mais justifiée [2]) qui consiste à brouter les observations pour chacun des membres de l'ensemble, donc de façon non-uniforme. Comme un échantillonnage statistique a tendance à s'appauvrir (par coalescence des points), on peut comprendre cela comme l'adjonction d'une partie stochastique permettant d'enrichir l'échantillon. L'algorithme est intuitif.

Filtre de Kalman d'ensemble

1. Initialisation

- État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .

2. Pour $t_k = 1, 2, \dots$

(a) Observation

- Création d'un jeu d'observations pour $j = 1, \dots, N$:

$$\mathbf{z}_j = \mathbf{z} + \mathbf{u}_j \quad \sum_{j=1}^N \mathbf{u}_j = \mathbf{0}$$

- Calcul de la matrice de covariance d'erreur associée

$$\mathbf{R} = \frac{1}{N-1} \sum_{j=1}^N (\mathbf{u}_j) (\mathbf{u}_j)^T$$

(b) Analyse

- Calcul du gain $\mathbf{K}^* = \mathbf{P}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R})^{-1}$
- Calcul de l'estimé de l'analyse pour $j = 1, \dots, N$

$$\mathbf{x}_j^a = \mathbf{x}_j^f + \mathbf{K}^* (\mathbf{z}_j - H(\mathbf{x}_j^f))$$

et de leur moyenne

$$\bar{\mathbf{x}}^a = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j^a$$

- Calcul de la matrice de covariance

$$\mathbf{P}^a = \frac{1}{N-1} \sum_{j=1}^N (\mathbf{x}_j^a - \bar{\mathbf{x}}^a) (\mathbf{x}_j^a - \bar{\mathbf{x}}^a)^T$$

(c) Prévision

- Calcul de l'estimé de prévision pour $j = 1, \dots, N$ $\mathbf{x}_j^f = M(\mathbf{x}_j^a)$
et de leur moyenne

$$\bar{\mathbf{x}}^f = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j^f$$

- Calcul de la matrice de covariance

$$\mathbf{P}^f = \frac{1}{N-1} \sum_{j=1}^N (\mathbf{x}_j^f - \bar{\mathbf{x}}^f) (\mathbf{x}_j^f - \bar{\mathbf{x}}^f)^T$$

Considérons à nouveau le cas de l'oscillateur non-linéaire. La performance du filtre de Kalman d'ensemble se visualise alors sur la figure 5.1, pour une évolution lente du système.

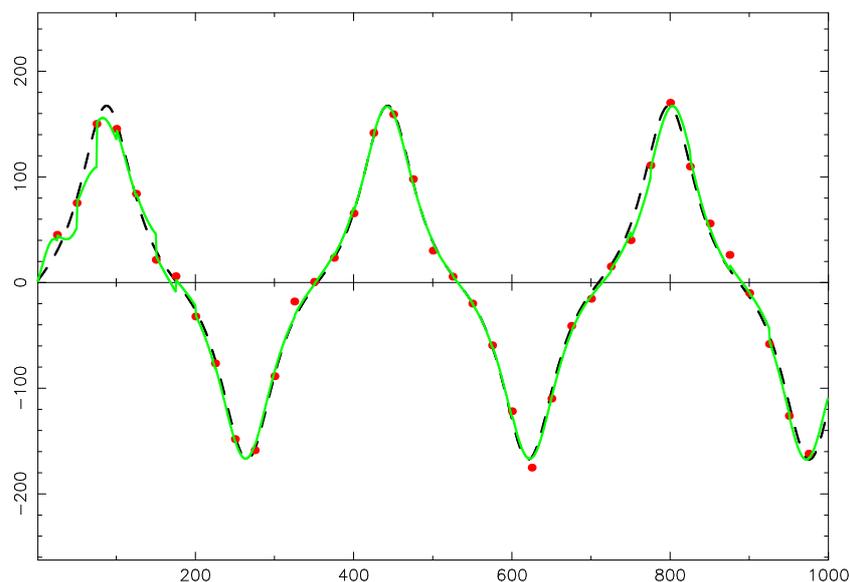


FIGURE 5.1 – Filtre de Kalman d’ensemble appliqué à l’oscillateur anharmonique, pour une dynamique lente.

Plaçons dans les conditions plus difficiles de l’oscillateur anharmonique perturbé comme introduit au chapitre 4 qui montrait la divergence du filtre de Kalman étendu en l’absence d’observations fréquentes. On se place dans les mêmes conditions, mais en n’assimilant qu’une observation pour 50, puis 100 itérations. Le résultat est représenté sur la figure 5.2 et 5.3.

Contrairement au filtre étendu, le filtre d’ensemble supporte une fréquence moindre. Le défaut principal de ce filtre est qu’il demande un nombre important de membres. Cette contrainte n’est pas décelable sur l’exemple de l’oscillateur qui ne comporte que deux variables. Néanmoins avec le succès récent rencontré par les architectures de calcul parallèles à bas coût, cet inconvénient est de moins en moins sensible.

Il existe également une variante déterministe de la partie analyse du filtre de Kalman d’ensemble qui permet d’éviter de bruyter les observations. Cette version déterministe se décline en une implémentation de type *racine carrée*, qui améliore le conditionnement numérique, et dont les détails seront discutés au chapitre 6 dans le cadre du filtre RRSQRT.

5.2 Le filtre de Kalman “unscented”

Le filtre de Kalman “unscented” est ce qu’aurait pu être le filtre de Kalman d’ensemble avec une rigueur mathématique plus poussée (une bonne revue est fournie par [26]). Il n’est pas ou peu connu en géophysique. C’est pourquoi on détaille ici sa formulation.

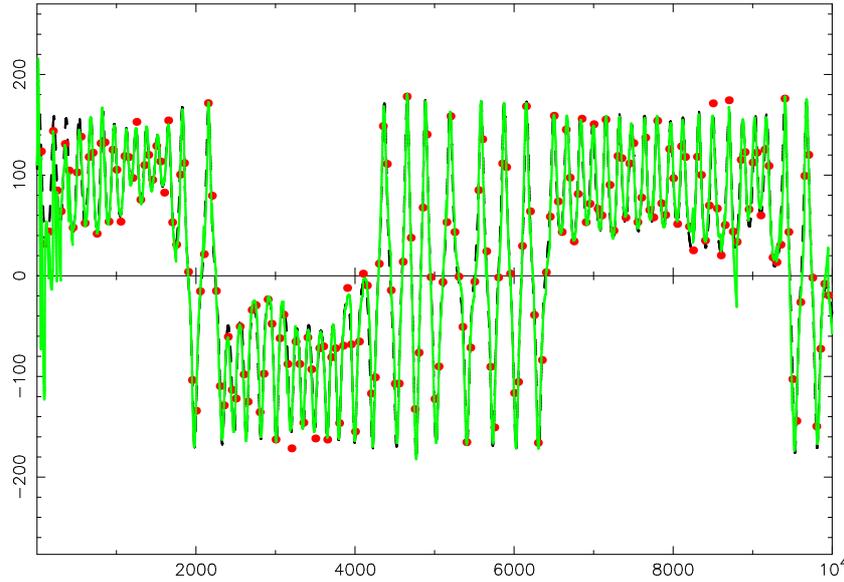


FIGURE 5.2 – Assimilation d’un oscillateur anharmonique perturbé, à l’aide d’un filtre de Kalman d’ensemble. La fréquence d’assimilation des observations est de 50 itérations pour une observation.

5.2.1 La transformation unscented

Propager les moments

Le problème est le suivant : connaissant l’espérance $\bar{\mathbf{x}}$ et la matrice de covariance d’une distribution P_x du vecteur aléatoire \mathbf{x} , quelle est l’espérance $\bar{\mathbf{y}}$ et la matrice de covariance P_y de la distribution du vecteur aléatoire \mathbf{y} image de \mathbf{x} par une transformation (en général non-linéaire) M ? Ce problème est lié à celui de la *propagation non-linéaire des incertitudes*. On suppose que le vecteur aléatoire x est assez proche d’être gaussien, sans l’être nécessairement.

L’étape de prédiction du filtre de Kalman étendu utilise la solution suivante à ce problème :

$$\begin{cases} \bar{\mathbf{y}} = M(\bar{\mathbf{x}}) \\ P_y = M' P_x (M')^T, \end{cases}$$

où M' (notée ∇M dans la suite de cette section) est la matrice jacobienne de M en $\bar{\mathbf{x}}$. Ce résultat est exact au premier ordre des perturbations autour de l’écart de l’opérateur M à la linéarité, au voisinage de $\bar{\mathbf{x}}$. Pour mieux comprendre la nature de l’erreur commise, on peut procéder au début d’un développement systématique de $\bar{\mathbf{y}}$ et P_y . M se développe autour de $\bar{\mathbf{x}}$ en (développement de Taylor des composantes de M en ses variables $x_i, i = 1, \dots, n$.)

$$\begin{aligned} M^l(\mathbf{x}) &= M^l(\bar{\mathbf{x}}) + \left(\partial_{x_i} M^l \right) (\mathbf{x} - \bar{\mathbf{x}})_i + \frac{1}{2} \left(\partial_{x_i} \partial_{x_j} M^l \right) (\mathbf{x} - \bar{\mathbf{x}})_i (\mathbf{x} - \bar{\mathbf{x}})_j \\ &\quad + \frac{1}{6} \left(\partial_{x_i} \partial_{x_j} \partial_{x_k} M^l \right) (\mathbf{x} - \bar{\mathbf{x}})_i (\mathbf{x} - \bar{\mathbf{x}})_j (\mathbf{x} - \bar{\mathbf{x}})_k + \dots, \end{aligned} \quad (5.1)$$

où la sommation sur les indices répétés est sous-entendue. Pour condenser les notations, on adopte des notations tensorielles, aussi bien dans l’espace de départ que celui d’arrivée.

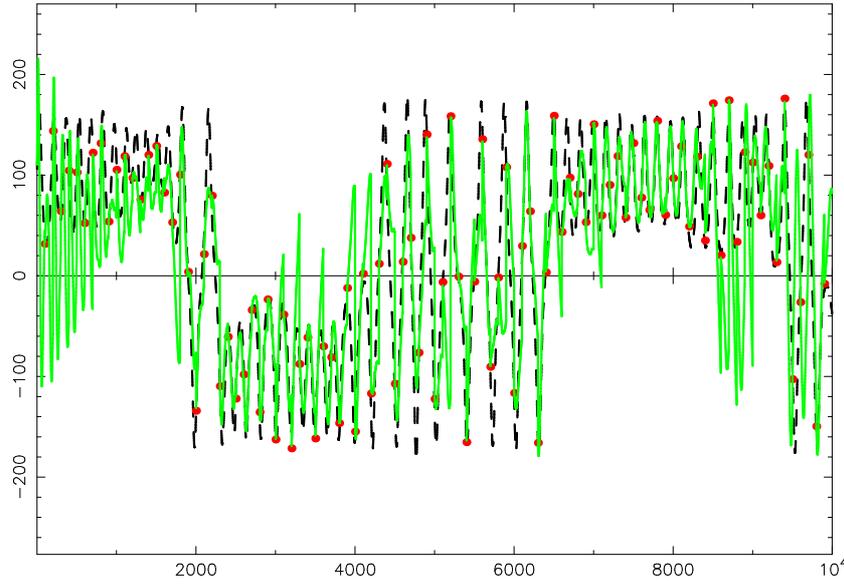


FIGURE 5.3 – Assimilation d'un oscillateur anharmonique perturbé, à l'aide d'un filtre de Kalman d'ensemble. La fréquence d'assimilation des observations est de 100 itérations pour une observation. Le filtre reste stable et les performances bonnes visuellement.

Le symbole \cdot dénote une contraction simple, double ou triple, alors que $(A)^n$ (version courte de $A^{\otimes n}$) représente le tenseur $A_{i_1}A_{i_2}\cdots A_{i_n}$. On notera $\nabla^n M$ le tenseur symétrique $(\partial_{x_{i_1}}\cdots\partial_{x_{i_n}})M$. Les indices de l'espace d'arrivée sont entièrement portés par les instances de M , alors que les indices de l'espace de départ sont tous contractés. On a alors :

$$M(\mathbf{x}) = M(\bar{\mathbf{x}}) + \nabla M \cdot (\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^2 + \frac{1}{6}\nabla^3 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^3 + \cdots$$

De plus, on définit le tenseur d'ordre 3 (*skew*) $S_x = E_x[(\mathbf{x} - \bar{\mathbf{x}})_i(\mathbf{x} - \bar{\mathbf{x}})_j(\mathbf{x} - \bar{\mathbf{x}})_k]$ et le tenseur d'ordre 4 (*kurtosis*) $K_x = E_x[(\mathbf{x} - \bar{\mathbf{x}})_i(\mathbf{x} - \bar{\mathbf{x}})_j(\mathbf{x} - \bar{\mathbf{x}})_k(\mathbf{x} - \bar{\mathbf{x}})_l]$. Dans le cas particulier d'une distribution gaussienne, le moment centré d'ordre 3 est nul $S_x = 0$, alors que le moment centré d'ordre 4, comme tous les moments pairs d'ordre supérieurs s'exprime en fonction du moment P_x d'ordre 2 par

$$K_x = (P_x)_{ij}(P_x)_{kl} + (P_x)_{ik}(P_x)_{jl} + (P_x)_{il}(P_x)_{jk}.$$

Avec ces notations, on a

$$\begin{aligned} \bar{\mathbf{y}} &= E_y(\mathbf{y}) = E_x(M(\mathbf{x})) \\ &= E_x\left(M(\bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^2 + \frac{1}{6}\nabla^3 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^3\right) + \cdots \\ &= M(\bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot P_x + \frac{1}{6}\nabla^3 M \cdot S_x + \cdots, \end{aligned} \quad (5.2)$$

On notera donc que la modélisation de l'étape de prédiction du filtre de Kalman étendu

est incorrecte à l’ordre deux puisqu’elle omet le terme $\frac{1}{2}\nabla^2 M \cdot P_x$! On a aussi

$$\begin{aligned}
P_y &\simeq E_y [(\mathbf{y} - \bar{\mathbf{y}})^2] = E_x [(M(\mathbf{x}) - \bar{\mathbf{y}})^2] \\
&= E_x \left[\left\{ M(\bar{\mathbf{x}}) + \nabla M \cdot (\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^2 + \frac{1}{6}\nabla^3 M \cdot (\mathbf{x} - \bar{\mathbf{x}})^3 \right. \right. \\
&\quad \left. \left. - \left(M(\bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot P_x + \frac{1}{6}\nabla^3 M \cdot S_x \right) \right\}^2 \right] \\
&= E_x \left[\left\{ \nabla M \cdot (\mathbf{x} - \bar{\mathbf{x}}) + \frac{1}{2}\nabla^2 M \cdot ((\mathbf{x} - \bar{\mathbf{x}})^2 - P_x) + \frac{1}{6}\nabla^3 M \cdot ((\mathbf{x} - \bar{\mathbf{x}})^3 - S_x) \right\}^2 \right] \\
&= \nabla M \cdot P_x \cdot (\nabla M) + \frac{1}{2}\nabla^2 M \cdot S_x \cdot \nabla M + \frac{1}{2}\nabla M \cdot S_x \cdot \nabla^2 M \\
&\quad + \frac{1}{4}\nabla^2 M \cdot (K_x - P_x^2) \cdot \nabla^2 M + \frac{1}{6}\nabla^3 M \cdot K_x \cdot \nabla M + \frac{1}{6}\nabla M \cdot K_x \cdot \nabla^3 M \quad (5.3)
\end{aligned}$$

Cette fois-ci, le terme dominant est bien celui qui est retenu dans l’algorithme du filtre de Kalman étendu.

Filtre étendu du second ordre

Au vu des développements précédents, on peut envisager de perfectionner le filtre de Kalman étendu en proposant l’étape de propagation

$$\begin{cases} \bar{\mathbf{y}} = M(\bar{\mathbf{x}}) + \frac{1}{2}M'' \cdot P_x \\ P_y = M'P_x(M')^T, \end{cases}$$

où $M'' = \nabla^2 M$ est le Hessian de la transformation M au point $\bar{\mathbf{x}}$. Le filtre obtenu porte le nom de *filtre de Kalman étendu du second ordre*. Il présente l’inconvénient majeur d’être numériquement inefficace dès que le système est de taille significative car il requiert le calcul du Hessian de la transformation (de taille L^2) !

La transformation unscented

Il existe une technique récente permettant de calculer numériquement $\bar{\mathbf{y}}$ et P_y et qui tire avantage de ces résultats. Elle est bien distincte de l’estimation utilisée par le filtre de Kalman étendu parce que

- les estimations de $\bar{\mathbf{y}}$ et P_y qu’elle fournit sont justes à l’ordre deux, comme le filtre de Kalman étendu du second ordre. Plus finement, elle permet la réduction des erreurs commises aux ordres supérieurs à deux.
- elle ne nécessite ni le calcul de la matrice jacobienne de M comme requis par le filtre de Kalman étendu standard, ni le calcul prohibitif et gourmand du Hessian, comme requis par le filtre étendu du second ordre. Ceci présente un intérêt pour les systèmes de grande taille. Toutefois, elle nécessite une décomposition de Choleski dont le coût de calcul (en $O(L^3)$) constitue son seul véritable défaut.

Cette méthode s’appelle la *transformation “unscented”*, et dans ses derniers raffinements, la *transformation “unscented” affine* (scaled unscented transformation ou *SUT*).

Représentation statistique par un ensemble de points

La transformation SUT consiste (essentiellement) à calculer l'image par M d'un nuage de point (dit *points sigma*) d'espérance $\bar{\mathbf{x}}$ et de covariance P_x et de calculer l'espérance et la covariance de ce nuage image, qu'on espère proches des valeurs exactes $\bar{\mathbf{y}}$, et P_y .

Si n est la dimension de l'espace des états, le nuage de points sigma est défini par

$$\begin{aligned} X_0 &= \bar{\mathbf{x}} \\ X_i &= \bar{\mathbf{x}} + \left(\sqrt{(n+\lambda)P_x} \right)_i \quad i = 1, \dots, n \\ X_i &= \bar{\mathbf{x}} - \left(\sqrt{(n+\lambda)P_x} \right)_i \quad i = n+1, \dots, 2n, \end{aligned} \quad (5.4)$$

où

$$\lambda = \alpha^2(n + \kappa) - n.$$

$(A)_i$ désigne la colonne i de la matrice A . Le nuage est donc constitué de $2n+1$ points. Le choix des paramètres α et κ sera discuté ultérieurement. À chacun des points, on affecte un poids. Ceux-ci sont définis par

$$\begin{aligned} W_0^{(m)} &= \frac{\lambda}{n+\lambda} \\ W_0^{(c)} &= \frac{\lambda}{n+\lambda} + (1 - \alpha^2 + \beta) \\ W_i^{(m)} &= W_i^{(c)} = \frac{1}{2(n+\lambda)} \quad i = 1, \dots, 2n, \end{aligned} \quad (5.5)$$

où β est un troisième paramètre, après α et κ , à préciser ultérieurement. On notera que $\sum_{i=0}^{2n} W^{(m)} = 1$, alors que $\sum_{i=0}^{2n} W_i^{(c)} = 1 + \beta + (1 - \alpha^2)$. On obtient de suite que

$$\sum_{i=0}^{2n} W_i^{(m)} X_i = \bar{\mathbf{x}} \quad (5.6)$$

$$\sum_{i=0}^{2n} W^{(c)} (X_i - \bar{\mathbf{x}})(X_i - \bar{\mathbf{x}})^T = \sum_{i=1}^n W^{(c)} (\sqrt{P_x})_i (\sqrt{P_x})_i^T = P_x, \quad (5.7)$$

ce qui est une condition nécessaire à la bonne construction du nuage sigma. Une fois défini le nuage sigma, on en calcule l'image par M :

$$Y_i = M(X_i) \quad i = 0, \dots, 2n.$$

Une fois obtenu le nuage image, on peut en extraire l'espérance et la matrice de covariance

$$\begin{aligned} \bar{\mathbf{y}} &= \sum_{i=0}^{2n} W_i^{(m)} Y_i \\ P_y &= \sum_{i=0}^{2n} W^{(c)} (Y_i - \bar{\mathbf{y}})(Y_i - \bar{\mathbf{y}})^T. \end{aligned} \quad (5.8)$$

On vérifie maintenant que cette transformation permet d’approcher correctement les valeurs de $\bar{\mathbf{y}}$ et P_y . On a en effet

$$\begin{aligned}
\bar{\mathbf{y}} &= \sum_{i=0}^{2n} W_i^{(m)} Y_i = \sum_{i=0}^{2n} W_i^{(m)} M(X_i) \\
&\simeq \sum_{i=0}^{2n} W_i^{(m)} \left\{ M(\bar{\mathbf{x}}) + \nabla M \cdot (X_i - \bar{\mathbf{x}}) + \frac{1}{2} \nabla^2 M \cdot (X_i - \bar{\mathbf{x}})^2 + \frac{1}{6} \nabla^3 M \cdot (X_i - \bar{\mathbf{x}})^3 \right\} \\
&\simeq M(\bar{\mathbf{x}}) + \nabla M \cdot (\bar{\mathbf{x}} - \bar{\mathbf{x}}) + \frac{1}{2} \nabla^2 M \cdot \sum_{i=0}^{2n} W_i^{(m)} (X_i - \bar{\mathbf{x}})^2 + \frac{1}{6} \nabla^3 M \cdot \sum_{i=0}^{2n} W_i^{(m)} (X_i - \bar{\mathbf{x}})^3 \\
&\simeq M(\bar{\mathbf{x}}) + \frac{1}{2} \nabla^2 M \cdot P_x
\end{aligned} \tag{5.9}$$

Donc la propagation par le nuage sigma est juste à l’ordre 2. On note qu’elle ne rend pas compte du moment d’ordre 3 (qu’elle sous-estime donc), et plus généralement de tous les moments d’ordre impair. On a aussi

$$\begin{aligned}
P_y &= \sum_{i=0}^{2n} W_i^{(c)} (Y_i - \bar{\mathbf{y}})^2 \\
&\simeq \sum_{i=0}^{2n} W_i^{(c)} \left(M(\bar{\mathbf{x}}) + \nabla M \cdot (X_i - \bar{\mathbf{x}}) + \frac{1}{2} \nabla^2 M \cdot (X_i - \bar{\mathbf{x}})^2 + \frac{1}{6} \nabla^3 M \cdot (X_i - \bar{\mathbf{x}})^3 \right. \\
&\quad \left. - M(\bar{\mathbf{x}}) - \frac{1}{2} \nabla^2 \cdot P_x \right)^2 \\
&\simeq \sum_{i=0}^{2n} W_i^{(c)} \left(\nabla M \cdot (X_i - \bar{\mathbf{x}}) + \frac{1}{2} \nabla^2 M \cdot ((X_i - \bar{\mathbf{x}})^2 - P_x) + \frac{1}{6} \nabla^3 M \cdot (X_i - \bar{\mathbf{x}})^3 \right)^2 \\
&\simeq \nabla M \cdot \left(\sum_{i=0}^{2n} W_i^{(c)} (X_i - \bar{\mathbf{x}})^2 \right) \cdot \nabla M + \frac{1}{2} \left(\sum_{i=0}^{2n} W_i^{(c)} \right) \nabla^2 M \cdot P_x^2 \cdot \nabla^2 M + \dots \\
&\simeq \nabla M \cdot P_x \cdot \nabla M + \frac{1}{2} (2 - \alpha^2 + \beta) \nabla^2 M \cdot P_x^2 \cdot \nabla^2 M + \dots
\end{aligned} \tag{5.10}$$

Donc la propagation par le nuage sigma est juste à l’ordre 2 également pour les moments d’ordre deux des distributions de probabilité. On a retenu dans le développement un seul des termes du développement à l’ordre 4. Le préfacteur de ce terme peut être ajusté par le paramètre β car ce paramètre n’apparaît pas dans les autres termes du développement à l’ordre 4. Ce qui justifie a posteriori l’introduction de ce paramètre (il n’intervient pas avant l’ordre 4).

Ajustement des paramètres de la transformation

On choisit $\kappa \geq 0$ de façon à maintenir la positivité de la matrice de covariance. On choisit souvent $\kappa = 0$, car son influence est limitée. α contrôle la taille du nuage (facteur d’échelle). Il est donc important qu’il ne soit pas trop grand, de manière à ne pas prendre en compte des effets non-locaux engendrés par une forte non-linéarité. On choisit souvent $0 \leq \alpha \leq 1$. Dans le cas où la distribution a priori est exactement gaussienne, alors un choix optimum pour β est possible : $\beta = 2$.

5.2.2 Le filtre de Kalman “unscented”

Un fois présentées les vertus de la transformation unscented, on souhaite reprendre le canevas du filtre de Kalman étendu et substituer aux calculs récursifs de matrice de covariance d’erreur, des transformations unscented.

On rappelle à quel type de système d’assimilation on souhaite appliquer le filtre. Il s’agit du système de variable d’état \mathbf{x}_k , de fonction de transition M_k , et de fonction d’observation H_k , gouverné par

$$\begin{aligned} \text{Évolution du système : } \mathbf{x}_{k+1} &= M_k(\mathbf{x}_k) + \mathbf{w}_k & \mathbf{E}(\mathbf{w}_k) &= 0 \quad \text{et} \quad \mathbf{R}(\mathbf{w}_k \mathbf{w}_l) = \mathbf{Q}_k \delta_{kl} \\ \text{Observation : } \mathbf{z}_k &= \mathbf{H}_k(\mathbf{x}_k) + \mathbf{v}_k & \langle \mathbf{v}_k \rangle &= 0 \quad \text{et} \quad \langle \mathbf{v}_k \mathbf{v}_l \rangle = \mathbf{R}_k \delta_{kl} \end{aligned} \quad (5.11)$$

On suppose ici que M_k et H_k sont éventuellement non-linéaires. Reprenons les étapes principales du filtre de Kalman :

Étape d’analyse

Dans l’algorithme du filtre de Kalman étendu, on commence l’étape d’analyse par un calcul brutal du gain de Kalman \mathbf{K}_k^*

$$\mathbf{K}^* = \mathbf{P}^f \mathbf{H}^T \left(\mathbf{H} \mathbf{P}^f \mathbf{H}^T + \mathbf{R} \right)^{-1} \quad (5.12)$$

or celui-ci fait intervenir H'_k . On doit donc substituer à cette linéarisation, une transformation SUT. Pour cela, on construit le nuage sigma correspondant à $(\mathbf{x}_k^f, \mathbf{P}_k^f)$:

$$X_k = \left\{ \mathbf{x}_k^f \quad \mathbf{x}_k^f + \gamma \sqrt{\mathbf{P}_k^f} \quad \mathbf{x}_k^f - \gamma \sqrt{\mathbf{P}_k^f} \right\},$$

et on calcule l’image de cet ensemble par la fonction d’observation $Y_k^f = H(X_k)$ qui conduit à une espérance estimée à

$$\mathbf{y}_k^f = \sum_{j=0}^{2L} W_j^{(m)} [Y_k^f]_j$$

où les poids W sont ceux la transformation SUT. Deux moments du second ordre utiles peuvent être construits à partir du nuage image

$$\begin{aligned} P_k^{YY} &= \sum_{j=0}^{2L} W_j^{(c)} \left([Y_k^f]_j - y_k^f \right) \left([Y_k^f]_j - y_k^f \right) \\ P_k^{XY} &= \sum_{j=0}^{2L} W_j^{(c)} \left([X_k^f]_j - x_k^f \right) \left([Y_k^f]_j - y_k^f \right) \end{aligned}$$

Il est clair que la traduction de l’expression gain de Kalman en fonction de ceux deux matrices de covariances est

$$K_k^* = P_k^{XY} \left(P_k^{YY} + R_k \right)^{-1}$$

Dans la suite, on absorbe R_k dans la définition de P_k^{YY} de sorte que $K_k^* = P_k^{XY} \left(P_k^{YY} \right)^{-1}$. L’estimé de l’analyse s’en déduit simplement

$$x_k^a = x_k^f + K_k^* \left(y_k - y_k^f \right),$$

tandis que la matrice de covariance d’erreur d’analyse s’écrit

$$P_k^a = P_k^f - K_k^* (P_k^{XY})^T,$$

qui est la traduction sans surprise de la formule utilisée dans le filtre de Kalman étendu

$$P_k^a = (I - K_k^* H_k') P_k^f.$$

Étape de propagation

Dans l’étape de propagation, c’est la non-linéarité de M_k qui est un obstacle. La formule $x_{k+1}^f = M_k(x_k^a)$ doit être remplacée une transformation SUT. On construit donc un nuage de points sigma associés à (x_k^a, P_k^a)

$$X_k^a = \{x_k^a \quad x_k^a + \gamma\sqrt{P_k^a} \quad x_k^a - \gamma\sqrt{P_k^a}\},$$

que l’on propage à travers M_k , $X_k^f = M(X_k^a)$. On peut donc en déduire un l’estimé de prévision

$$x_k^f = \sum_{j=0}^{2L} W_j^{(m)} [Y_k]_j,$$

et la matrice de covariance d’erreur de prévision

$$P_k^f = \sum_{j=0}^{2L} W_j^{(c)} \left([X_k^f]_j - x_k^f \right) \left([X_k^f]_j - x_k^f \right) + Q_k.$$

Complexité de l’algorithme

Résumé

Le filtre de Kalman unscented est décrit synthétiquement dans la figure Fig. 5.2.2.

Filtre de Kalman "unscented" (UKF)

1. Initialisation

- État du système x_0^f et covariance d'erreur P_0^f .

2. Pour $t_k = 1, 2, \dots$

(a) Analyse

- Calcul d'un nuage de points sigma

$$X_k = \left\{ x_k^f \quad x_k^f + \gamma \sqrt{P_k^f} \quad x_k^f - \gamma \sqrt{P_k^f} \right\}$$

- Calcul de l'image du nuage par l'observation $Y_k^f = H(X_k)$ et de son espérance

$$y_k^f = \sum_{j=0}^{2L} W_j^{(m)} [Y_k^f]_j$$

- Calcul du gain $K_k^* = P_k^{XY} (P_k^{YY})^{-1}$ avec

$$P_k^{YY} = \sum_{j=0}^{2L} W_j^{(c)} \left([Y_k^f]_j - y_k^f \right) \left([Y_k^f]_j - y_k^f \right) + R_k$$

$$P_k^{XY} = \sum_{j=0}^{2L} W_j^{(c)} \left([X_k^f]_j - x_k^f \right) \left([Y_k^f]_j - y_k^f \right)$$

- Calcul de l'estimé de l'analyse $x_k^a = x_k^f + K_k^* (y_k - y_k^f)$
- Calcul de la matrice de covariance d'erreur d'analyse

$$P_k^a = P_k^f - K_k^* P_k^{YY} (K_k^*)^T = P_k^f - K_k^* (P_k^{XY})^T$$

(b) Prévision

- Calcul d'un nouveau nuage de points sigma

$$X_k^a = \left\{ x_k^a \quad x_k^a + \gamma \sqrt{P_k^a} \quad x_k^a - \gamma \sqrt{P_k^a} \right\}$$

- Propagation du nuage $X_k^f = M(X_k^a)$
- Calcul de l'estimé de prévision

$$x_k^f = \sum_{j=0}^{2L} W_j^{(m)} [Y_k]_j$$

- Calcul de la matrice de covariance d'erreur de prévision

$$P_k^f = \sum_{j=0}^{2L} W_j^{(c)} \left([X_k^f]_j - x_k^f \right) \left([X_k^f]_j - x_k^f \right) + Q_k$$

FIGURE 5.4 – Le filtre de Kalman unscented.

Chapitre 6

Appréhender la grande taille des systèmes : filtres réduits

Lorsque la dimensionnalité de l'espace des états est très grande, il n'est plus possible d'appliquer les techniques séquentielles. La dimension d'un système géophysique peut être considérable. Dans le cas d'un système de prévision de la qualité de l'air, si l'on considère que le modèle numérique comporte 6400 mailles spatiales (80×80) et 60 espèces, on arrive à une dimension de 300.000. Il est alors difficile de manipuler des matrices de covariances d'erreur, encore plus de les propager. Cependant, à un instant donné, la physique du problème n'est contrôlée que par un nombre limité de variables (ou combinaisons de variables). Dans le cas d'un modèle de qualité de l'air, seules quelques espèces déterminent le comportement local du système à un instant donné. On peut alors faire l'hypothèse que les statistiques d'erreur significatives sont données par celles portant sur ces quelques degrés de liberté (les modes réduits). Encore faut-il les identifier.

On imagine que l'idée est transposable à un filtre de Kalman, à la condition que l'on sache l'appliquer à ces modes réduits.

Dans cette optique, il faut également être capable d'enrichir stochastiquement le système car la base des modes réduits doit pouvoir évoluer sans contraintes trop fortes. Le risque est que ces modes, s'ils dégèrent, ne sous-tendent plus la fraction de l'espace des états dans lequel évolue le système dynamique réel.

6.1 Le filtre RRSQRT

Le filtre RRSQRT est une réponse naturelle à ce problème. Outre le problème de la dimensionnalité, le problème du calcul et de la propagation des matrices de covariance d'erreur est contourné. En effet les directions principales des matrices d'erreur seront représentées par les modes réduits. De sorte que l'accent (propagation, enrichissement) est mis sur les modes et non les matrices.

On considère donc un espace des états de dimension n . L'état du système initial est \mathbf{x}_0^f , auquel on associe une matrice de covariance d'erreur \mathbf{P}_0^f . On suppose que l'on réalise une décomposition en modes principaux de $\mathbf{P}_0^f : \mathbf{P}_0^f \simeq \mathbf{S}_0^f (\mathbf{S}_0^f)^T$, où \mathbf{S}_0^f est une matrice $n \times m$ de vecteurs colonnes les m premiers modes principaux de \mathbf{P}_0^f . L'ébauche sur l'erreur a donc été réduite. Plutôt que de raisonner sur la matrice \mathbf{P}_0^f , nous raisonnons sur ses

modes \mathbf{S}_0^f , et également sur la transposition de ces modes dans l'espace des observations $\Psi = \left(\mathbf{H}_k \mathbf{S}_k^f\right)^T$. \mathbf{H}_k est soit l'opérateur d'observation lorsque celui-ci est linéaire, ou bien son linéaire tangent. Cette opérateur apparaît naturellement lorsqu'on considère le gain de Kalman que l'on doit calculer lors de l'étape d'analyse.

$$\begin{aligned} \mathbf{K}_k^* &= \mathbf{P}_k^f \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \\ &= \mathbf{S}_k^f \left(\mathbf{S}_k^f \right)^T \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{S}_k^f \left(\mathbf{S}_k^f \right)^T \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \\ &= \mathbf{S}_k^f \left(\mathbf{H}_k \mathbf{S}_k^f \right)^T \left(\left(\mathbf{H}_k \mathbf{S}_k^f \right) \left(\mathbf{H}_k \mathbf{S}_k^f \right)^T + \mathbf{R}_k \right)^{-1} \end{aligned} \quad (6.1)$$

Le gain de Kalman, calculé dans l'étape d'analyse s'exprime alors simplement à l'aide de la matrice Ψ^T de taille $p \times m$

$$\mathbf{K}_k^* = \mathbf{S}_k^f \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1}. \quad (6.2)$$

L'estimé de l'analyse \mathbf{x}_k^a s'obtient à l'aide de ce gain et de la formule normale.

Qu'advient-il de la formule permettant d'en déduire la matrice de covariance d'analyse \mathbf{P}_k^a ? On a :

$$\begin{aligned} \mathbf{P}_k^a &= \left(\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k \right) \mathbf{P}_k^f \\ &= \left(\mathbf{I} - \mathbf{S}_k^f \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1} \mathbf{H}_k \right) \mathbf{S}_k^f \left(\mathbf{S}_k^f \right)^T \\ &= \mathbf{S}_k^f \left(\mathbf{I} - \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1} \Psi^T \right) \left(\mathbf{S}_k^f \right)^T \end{aligned} \quad (6.3)$$

De cette dernière expression, on tire que la matrice (racine) de covariance d'analyse peut s'obtenir au moyen de

$$\mathbf{S}_k^a = \mathbf{S}_k^f \left(\mathbf{I} - \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1} \Psi^T \right)^{1/2}, \quad (6.4)$$

et non par un calcul direct sur les matrices de covariances. Le calcul de la racine peut sembler coûteux numériquement. Il n'en est rien puisque Ψ est de taille réduite $m \times p$, et que $\mathbf{I} - \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1} \Psi^T$ est de taille $m \times m$. De plus le conditionnement de la matrice racine est meilleur que celle de la matrice de départ (racine du conditionnement originel). Ceci assure une meilleure précision numérique.

Après l'analyse, on cherche à réduire la dimension du systèmes en diminuant le nombre de modes de m à $m - q$. Pour cela, on diagonalise $\mathbf{S}_k^a T \mathbf{S}_k^a = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T$. On ne retient alors de la matrice de passage V que les vecteurs associés au $m - q$ premiers modes ($m - q$ plus grandes valeurs propres de $\mathbf{\Lambda}$), rangés dans $\tilde{\mathbf{V}}$. Puis \mathbf{S}_k^a est réduit à $\mathbf{S}_k^a \tilde{\mathbf{V}}$.

À l'étape de prévision, l'estimé de prévision est propagé $\mathbf{x}_{k+1}^f = M_{k+1}(\mathbf{x}_k^a)$. La matrice racine \mathbf{S}_k^a est propagée à l'aide du modèle tangent. Étape cruciale, elle est ensuite élargie en ajoutant q modes que l'on impute à l'erreur modèle. La matrice des modes augmentée est alors de la forme

$$\mathbf{S}_{k+1}^f = [M'_{k+1} \tilde{\mathbf{S}}_k^a, \mathbf{T}_k], \quad (6.5)$$

et comporte m modes. On peut alors recommencer le cycle d'assimilation.

Ce type de filtre a en particulier été développé et utilisé en qualité de l'air ou le nombre d'espèces chimiques à suivre augmente significativement la taille du système (on pourra consulter la thèse d'Arjo Segers à ce sujet [17]).

6.2 Le filtre SEEK

Le *filtre SEEK* a été introduit par Dinh Tuan Pham en 1996 [15]. En deux mots, il s'agit d'un filtre réduit déduit du filtre de Kalman étendu. Par soucis de simplicité, on pourra donc se restreindre au cas du filtre de Kalman linéaire, le filtre étendu lui étant formellement analogue. Il repose sur la stagnation ou la décroissance du rang des matrices de covariances d'erreur, propriété avérée ou forcée selon le cas.

6.2.1 Filtre sans erreur modèle

Dans un premier temps, on suppose que le modèle utilisé est parfait. On rappelle qu'alors les équations d'évolution des matrices de covariance d'erreur s'écrivent

$$\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f \quad \mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^a \mathbf{M}_{k+1}^T. \quad (6.6)$$

Il résulte de ces formules que le rang de \mathbf{P}_k^f (noté $r = \text{rang}(\mathbf{P}_k^f)$) est une fonction décroissante de t_k , puisque, *in fine*, la récurrence est de la forme $\mathbf{P}_{k+1}^f = \mathbf{A}_k \mathbf{P}_k^f \mathbf{B}_k$. Par conséquent, si le rang de la matrice de covariance initiale est faible comparé à la dimension n de l'espace des états, il le restera. C'est l'idée de départ du filtre SEEK pour *Singular Evolutive Extended Kalman*. Cela tient à ce que l'erreur est portée par les r directions principales de \mathbf{P}_k^a , que l'analyse réduit les erreurs associées à ces seules directions et que la propagation ne fait que modifier ces directions sans en créer de nouvelles.

Comme toute matrice symétrique positive, \mathbf{P}_k^a peut se diagonaliser en $\mathbf{P}_k^a = \mathbf{V}_k \Lambda_k \mathbf{V}_k^T$, avec Λ_k une matrice diagonale à coefficient positifs ou nuls de taille $n \times n$, et \mathbf{V}_k une matrice orthogonale de taille $n \times n$, décrivant les directions principales d'erreur. Puisque $\text{rang}(\mathbf{P}_k^f) \leq r$, on peut écrire $\mathbf{V}_k = [\mathbf{L}_k \quad *]$ et $\Lambda_k = \begin{bmatrix} \mathbf{U}_k & 0 \\ 0 & 0 \end{bmatrix}$, avec \mathbf{L}_k une matrice de taille $n \times r$, et \mathbf{U}_k une matrice symétrique positive de taille $r \times r$, qui pourra, par la suite ne pas être diagonale. Sous cette forme, la décomposition $\mathbf{P}_k^a = \mathbf{V}_k \Lambda_k \mathbf{V}_k^T$ n'est donc pas nécessairement unique, mais elle existe dans tous les cas. On a

$$\mathbf{P}_k^f = \mathbf{V}_k \Lambda_k \mathbf{V}_k^T = [\mathbf{L}_k \quad *] \begin{bmatrix} \mathbf{U}_k & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{L}_k^T \\ * \end{bmatrix} = [\mathbf{L}_k \mathbf{U}_k \quad 0] \begin{bmatrix} \mathbf{L}_k^T \\ * \end{bmatrix} = \mathbf{L}_k \mathbf{U}_k \mathbf{L}_k^T. \quad (6.7)$$

Analyse

Il résulte de $\mathbf{P}_k^a = (\mathbf{I} - \mathbf{K}_k^* \mathbf{H}_k) \mathbf{P}_k^f$ et de $\mathbf{P}_k^f = \mathbf{L}_k \mathbf{U}_k \mathbf{L}_k^T$ que \mathbf{P}_k^a s'écrit

$$\begin{aligned} \mathbf{P}_k^a &= \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T - \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T \\ &= \mathbf{L}_k \left(\mathbf{U}_k^f - \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f \right) \mathbf{L}_k^T. \end{aligned} \quad (6.8)$$

Filtre de Kalman Réduit (RRSQRT)

1. Initialisation

- État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .
- Décomposition : $\mathbf{Q}_k = \mathbf{T}_k \mathbf{T}_k^T$ et $\mathbf{P}_0^f = \mathbf{S}_0^f \mathbf{S}_0^{fT}$.

2. Pour $t_k = 1, 2, \dots$

(a) Analyse

- Calcul du gain

$$\Psi = \left(\mathbf{H}_k \mathbf{S}_k^f \right)^T, \quad \mathbf{K}_k^* = \mathbf{S}_k^f \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1}$$

- Calcul de l'estimé de l'analyse

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* \left(\mathbf{z}_k - H_k(\mathbf{x}_k^f) \right)$$

- Calcul du nuage statistique

$$\mathbf{S}_k^a = \mathbf{S}_k^f \left(\mathbf{I} - \Psi \left(\Psi^T \Psi + \mathbf{R}_k \right)^{-1} \Psi^T \right)^{1/2}$$

- \mathbf{S}_k^a possède m modes.

(b) Réduction

- Diagonalisation $\mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = \mathbf{S}_k^{aT} \mathbf{S}_k^a$
- Réduction du nuage $\tilde{\mathbf{S}}_k^a = \mathbf{S}_k^a \tilde{\mathbf{V}}$
- $\tilde{\mathbf{S}}_k^a$ possède $m - q$ modes.

(c) Prévion

- Calcul de l'estimé de prévision $\mathbf{x}_{k+1}^f = M_{k+1}(\mathbf{x}_k^a)$
- Calcul du nuage statistique

$$\mathbf{S}_{k+1}^f = [M_{k+1}' \tilde{\mathbf{S}}_k^a, \mathbf{T}_k]$$

- \mathbf{S}_{k+1}^f possède m modes.

Par conséquent, un choix possible de décomposition de \mathbf{P}_k^a est $\mathbf{P}_k^a = \mathbf{L}_k \mathbf{U}_k^a \mathbf{L}_k^T$ avec

$$\mathbf{U}_k^a = \mathbf{U}_k^f - \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T \left[\mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f \mathbf{L}_k^T \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \mathbf{H}_k \mathbf{L}_k \mathbf{U}_k^f. \quad (6.9)$$

Par conséquent, l'analyse ne modifie pas l'espace vectoriel engendré par les directions principales des erreurs. Toutefois, les directions elles-même changent en général. Dans le cas où \mathbf{U}_k^f est inversible, ce qui est le cas sauf si le rang de \mathbf{P}_k a été réduit (auquel cas, on peut la décomposer à nouveau), la traduction de $(\mathbf{P}_k^a)^{-1} = \left(\mathbf{P}_k^f\right)^{-1} + \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k^T$ (adaptation de Eq.(1.25)) est

$$(\mathbf{U}_k^a)^{-1} = \left(\mathbf{U}_k^f\right)^{-1} + \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{L}_k, \quad (6.10)$$

qui traduit l'additivité des matrices de confiances dans une base réduite. La formule donnant l'estimé de l'analyse fait toujours intervenir le gain de Kalman \mathbf{K}_k^* ,

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* \left(\mathbf{z}_k - H_k(\mathbf{x}_k^f) \right). \quad (6.11)$$

Ce dernier s'obtient assez simplement puisque l'on a déjà obtenu \mathbf{P}_k^a :

$$\mathbf{K}_k^* = \mathbf{P}_k^a \mathbf{H}_k^T \mathbf{R}_k^{-1}. \quad (6.12)$$

Propagation

Il résulte de $\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^f \mathbf{M}_{k+1}^T$ et de $\mathbf{P}_k^a = \mathbf{L}_k \mathbf{U}_k^a \mathbf{L}_k^T$ que \mathbf{P}_{k+1}^f s'écrit

$$\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{L}_k \mathbf{U}_k^a \mathbf{L}_k^T \mathbf{M}_{k+1}^T, \quad (6.13)$$

qui suggère la décomposition $\mathbf{P}_{k+1}^f = \mathbf{L}_{k+1} \mathbf{P}_k^a \mathbf{L}_{k+1}^T$ avec

$$\mathbf{U}_{k+1}^f = \mathbf{U}_k^a, \quad (6.14)$$

et la propagation des directions de \mathbf{L}_k :

$$\mathbf{L}_{k+1} = \mathbf{M}_{k+1} \mathbf{L}_k. \quad (6.15)$$

Dans l'étape d'analyse, l'espace engendré par les r premières directions principales ne change pas, alors que la matrice de covariance réduite est modifiée. C'est tout le contraire dans l'étape de propagation.

6.2.2 Filtre avec erreur modèle

Revenons maintenant au cas plus général où le modèle est imparfait. Dans ce cas la propagation de la matrice de covariance d'erreur s'écrit $\mathbf{P}_{k+1}^f = \mathbf{M}_{k+1} \mathbf{P}_k^f \mathbf{M}_{k+1}^T + \mathbf{Q}_k$.

Il est donc clair que le rang de \mathbf{P}_{k+1}^f peut être supérieur à celui de \mathbf{P}_k^a . Et il ne plus y avoir réduction du filtre *sans approximation*. La solution la plus naturelle et la plus simple pour réduire le rang est de projeter \mathbf{Q}_k en $\Pi_k \mathbf{Q}_k \Pi_k$ où Π_k est la projection orthogonale sur l'espace engendré par les vecteurs-colonnes de \mathbf{L}_k . C'est-à-dire qu'on ne retient de l'erreur

modèle que sa composante agissant dans le sous-espace sur lequel agit P_k^a . Cette projection s'écrit

$$\Pi_k = \mathbf{L}_k (\mathbf{L}_k^T \mathbf{L}_k)^{-1} \mathbf{L}_k^T \quad (6.16)$$

où $\mathbf{L}_k^T \mathbf{L}_k$, matrice carrée de taille $r \times r$ est inversible car \mathbf{L}_k est de rang r (les vecteurs colonnes de \mathbf{L}_k sont indépendants). La propagation approchée de l'erreur s'écrit donc

$$\begin{aligned} \mathbf{P}_{k+1}^f &\simeq \mathbf{M}_{k+1} \mathbf{P}_k^f \mathbf{M}_{k+1}^T + \Pi_{k+1} \mathbf{Q}_k \Pi_{k+1} \\ &= \mathbf{M}_{k+1} \mathbf{L}_k (\mathbf{U}_k^a + (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1} \mathbf{L}_{k+1}^T \mathbf{Q}_k \mathbf{L}_{k+1} (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1}) \mathbf{L}_k^T \mathbf{M}_{k+1}^T \end{aligned} \quad (6.17)$$

et donc

$$\mathbf{U}_{k+1}^f = \mathbf{U}_k^a + (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1} \mathbf{L}_{k+1}^T \mathbf{Q}_k \mathbf{L}_{k+1} (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1}, \quad (6.18)$$

avec toujours

$$\mathbf{L}_{k+1} = \mathbf{M}_{k+1} \mathbf{L}_k. \quad (6.19)$$

Cette fois-ci, la matrice de covariance réduite est également affectée.

Pourquoi peut-on réduire ?

Le système dynamique (modèle direct), amplifie les erreurs associées à un sous-espace \mathcal{A}_k de l'espace tangent à l'espace modèle, tandis que les erreurs associées au complémentaire de cet espace sont atténuées. Si le rang de la matrice de covariance des erreurs est supérieur à la dimension de \mathcal{A}_k , alors on peut espérer que le système dynamique se chargera d'atténuer *toutes* les erreurs commises et non corrigées dans l'espace complémentaire à \mathcal{A}_k .

6.3 Le filtre SEIK

Le *filtre SEIK* a été introduit par Dinh Tuan Pham dans [14]. Il s'agit d'un filtre réduit du filtre de Kalman étendu du second ordre. On peut le voir comme une amélioration du filtre SEEK, en vue de l'appliquer à des systèmes significativement non-linéaires. Son squelette est celui du filtre SEEK. L'acronyme SEIK signifie *Singular evolutive interpolated Kalman*. Cela se comprend puisqu'il s'agit de substituer à la linéarisation opérée dans le filtre de Kalman étendu (et donc aussi dans le filtre SEEK), une interpolation sur un échantillon d'états bien choisis que l'on propage dans l'étape de prévision.

6.3.1 Filtre sans erreur modèle

Comme pour le filtre SEEK, on commence par envisager le cas, où l'erreur modèle est absente.

Tirage exact au second ordre

Le "tirage exact au second ordre" répond aux mêmes objectifs que la transformation unscented (confère section 5.2.1). Il s'agit de trouver un échantillon, composés de $r + 1$ vecteurs $\{X_i\}_{1 \leq i \leq r+1}$, dont la moyenne soit égale à un vecteur donné X et la matrice de covariance empirique soit égale à une matrice donnée P de rang r .

$$X = \frac{1}{r+1} \sum_{i=1}^{r+1} X_i \quad P = \frac{1}{r+1} \sum_{i=1}^{r+1} (X_i - X)(X_i - X)^T. \quad (6.20)$$

Filtre SEIK (Singular Evolutive Extended Kalman)

1. Initialisation

- État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .
- Diagonalisation $\mathbf{P}_0^f = \mathbf{L}_0 \mathbf{U}_0^f \mathbf{L}_0^T$.

2. Pour $t_k = 1, 2, \dots$

(a) Analyse

- Calcul de la matrice de covariance réduite

$$(\mathbf{U}_k^a)^{-1} = (\mathbf{U}_k^f)^{-1} + \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{L}_k$$

- Calcul du gain $\mathbf{K}_k^* = \mathbf{L}_k^T \mathbf{U}_k^a \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1}$
- Calcul de l'estimé de l'analyse

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* (\mathbf{z}_k - H_k(\mathbf{x}_k^f))$$

(b) Prévision

- Calcul de l'estimé de prévision $\mathbf{x}_{k+1}^f = M_{k+1}(\mathbf{x}_k^a)$
- Calcul des vecteurs engendrant l'espace des directions principales

$$\mathbf{L}_{k+1} = \mathbf{M}_{k+1} \mathbf{L}_k$$

- Calcul de la matrice de covariance réduite

$$\mathbf{U}_{k+1}^f = \mathbf{U}_k^a + (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1} \mathbf{L}_{k+1}^T \mathbf{Q}_k \mathbf{L}_{k+1} (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1}$$

Une telle construction nécessite au moins $r + 1$ vecteurs car $\text{rang}(P) = \dim \text{Vec}\{X_i\} - 1$. Contrairement à la version unscented, on souhaite que cette solution soit aléatoire, afin d'éviter de privilégier des directions.

On se donne donc X et P . On réalise d'abord une décomposition de Choleski de la matrice P , sous la forme $P = CC^T$, où C est une matrice triangulaire supérieure de taille $r \times r$. Puis on construit une matrice orthogonale de taille $(r + 1) \times (r + 1)$ de la forme

$$\tilde{\Omega} = [\omega \quad \Omega] \quad \text{avec} \quad \omega^T = \frac{1}{\sqrt{r+1}}(1, 1, \dots, 1), \quad (6.21)$$

et Ω est par conséquent une matrice de taille $(r + 1) \times r$ dont les vecteurs-colonnes sont mutuellement orthogonales et toutes orthogonales à ω . Le fait que $\tilde{\Omega}$ soit un opérateur orthogonal se traduit par

$$\mathbb{I}_{r+1} = \tilde{\Omega}^T \tilde{\Omega} = \begin{bmatrix} \omega^T \\ \Omega^T \end{bmatrix} [\omega \quad \Omega] = \begin{bmatrix} \omega^T \omega & \omega^T \Omega \\ \Omega^T \omega & \Omega^T \Omega \end{bmatrix} \quad (6.22)$$

On déduit que $\Omega^T \Omega = \mathbb{I}_r$ et que $\omega^T \Omega = \Omega^T \omega = 0$. Alors on a en particulier la factorisation

$$P = C\Omega^T \Omega C^T = (C\Omega^T)(C\Omega^T)^T. \quad (6.23)$$

qui suggère de définir les vecteurs pour $i = 1, \dots, r + 1$

$$X_i = X + \sqrt{r+1} C(\Omega_i)^T, \quad (6.24)$$

où Ω_i est ici le i -ème vecteur ligne de la matrice Ω . On a alors

$$\begin{aligned} \frac{1}{r+1} \sum_{i=1}^{r+1} X_i &= \frac{1}{r+1} \sum_{i=1}^{r+1} (X + \sqrt{r+1} C(\Omega_i)^T) = X + \frac{1}{\sqrt{r+1}} C \left(\sum_{i=1}^{r+1} \Omega_i \right)^T \\ &= X + \frac{1}{\sqrt{r+1}} C (\omega^T \Omega) = X. \end{aligned} \quad (6.25)$$

On a aussi

$$\begin{aligned} \frac{1}{r+1} \sum_{i=1}^{r+1} (X_i - X)(X_i - X)^T &= \sum_{i=1}^{r+1} (C(\Omega_i)^T) (C^T(\Omega_i)^T)^T \\ &= C \left(\sum_{i=1}^{r+1} (\Omega_i)^T \Omega_i \right) C^T \\ &= C\Omega^T \Omega C^T = P. \end{aligned} \quad (6.26)$$

Analyse

Comme pour le filtre SEEK l'estimé de l'analyse est donné par

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* \left(\mathbf{z}_k - H_k(\mathbf{x}_k^f) \right). \quad (6.27)$$

Afin d'éviter la linéarisation de l'opérateur d'observation H_k , le gain de Kalman n'est plus

$$\mathbf{K}_k^* = \mathbf{L}_k^T \mathbf{U}_k^a \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1}, \quad (6.28)$$

mais est remplacé par

$$\mathbf{K}_k^* = \mathbf{L}_k^T \mathbf{U}_k^a (\mathbf{H}_k \cdot \mathbf{L}_k)^T \mathbf{R}_k^{-1}, \quad (6.29)$$

avec

$$\mathbf{H}_k \cdot \mathbf{L}_k = [\mathbf{H}_k(\mathbf{L}_k)_1 \quad \cdots \quad \mathbf{H}_k(\mathbf{L}_k)_r]. \quad (6.30)$$

Il en va de même pour l'actualisation de la matrice de covariance d'erreur réduite qui doit prendre la forme

$$(\mathbf{U}_k^a)^{-1} = (\mathbf{U}_k^f)^{-1} + (\mathbf{H}_k \cdot \mathbf{L}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \cdot \mathbf{L}_k). \quad (6.31)$$

Propagation

L'étape de propagation commence par un ré-échantillonnage, faisant appel au tirage exact du second ordre (confère section 6.3.1). On dispose de la moyenne \mathbf{x}_k^a et de la matrice de covariance $\mathbf{P}_k^a = \mathbf{L}_k \mathbf{U}_k^a \mathbf{L}_k^T$, qui est de taille $n \times n$ mais de rang r . On utilise alors une décomposition (aléatoire) possible de \mathbf{U}_k^a qui s'écrit

$$\mathbf{U}_k^a = C_k \Omega_k^T \Omega_k C_k^T. \quad (6.32)$$

Alors les états interpolés sont de la forme

$$(\mathbf{x}_k^a)_i = \mathbf{x}_k^a + \sqrt{r+1} L_k C_{k-1} (\Omega_k)_i^T \quad (6.33)$$

avec $(\Omega_k)_i$ la i ème ligne de Ω_k . Il est nécessaire d'en construire au moins $r+1$, pour obtenir une matrice de covariance empirique de rang r . Les $r+1$ états $(\mathbf{x}_k^a)_i$ sont inclus dans l'espace affine $x_k^a + L_k$. Ce sont donc des combinaison linéaires (ou états "interpolés") des n précédents vecteurs (d'où l'appellation SEIK). Une fois construits, on peut les propager :

$$\left(\mathbf{x}_{k+1}^f \right)_i = M_{k+1} (\mathbf{x}_k^a)_i, \quad (6.34)$$

ce qui conduit à la moyenne

$$\mathbf{x}_{k+1}^f = \frac{1}{r+1} \sum_{i=1}^{r+1} \left(\mathbf{x}_{k+1}^f \right)_i, \quad (6.35)$$

et à la matrice de covariance

$$\mathbf{P}_{k+1}^f = \frac{1}{r+1} \sum_{i=1}^{r+1} \left[\left(\mathbf{x}_k^f \right)_i - \mathbf{x}_{k+1}^f \right] \left[\left(\mathbf{x}_{k+1}^f \right)_i - \mathbf{x}_{k+1}^f \right]^T \quad (6.36)$$

Il apparaîtra plus tard qu'il n'est pas nécessaire de calculer \mathbf{x}_{k+1}^f et \mathbf{P}_{k+1}^f . Des candidats naturels pour les colonnes de la matrice \mathbf{L}_k seraient donc les $\left(\mathbf{x}_k^f \right)_i - \mathbf{x}_k^f$. Cependant, ils sont au nombre de $r+1$, pour r colonnes. Il est donc nécessaire d'en éliminer un en adaptant la formule Eq.(6.36). L'astuce consiste à introduire la matrice constante T définie par

$$\mathbf{T} = \begin{bmatrix} \mathbb{I}_r \\ 0 \end{bmatrix} - \frac{1}{\sqrt{r+1}} [\omega \quad \cdots \quad \omega] \quad (6.37)$$

De sorte que

$$\mathbf{P}_{k+1}^f = \mathbf{L}_{k+1} [(r+1)\mathbf{T}^T\mathbf{T}]^{-1} \mathbf{L}_{k+1}^T, \quad (6.38)$$

avec

$$\mathbf{L}_k = \begin{bmatrix} (\mathbf{x}_k^f)_1 & \cdots & (\mathbf{x}_k^f)_{r+1} \end{bmatrix} \mathbf{T}, \quad (6.39)$$

qui est une matrice $n \times r$ constituée des r vecteur-colonnes recherchées. Une formule similaire à celle du filtre a donc été obtenue pour la matrice de covariance d'erreur de prédiction. La matrice de covariance d'erreur réduite doit donc être initialisée après la propagation à

$$\mathbf{U}_{k+1}^f = (r+1)\mathbf{T}^T\mathbf{T}. \quad (6.40)$$

6.3.2 Filtre avec erreur modèle

On procède avec le filtre SEIK comme on a fait pour le filtre SEEK. C'est-à-dire que l'on projette l'erreur modèle le long des directions principales

$$\mathbf{U}_{k+1}^f = (r+1)\mathbf{T}^T\mathbf{T} + (\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}\mathbf{L}_{k+1}^T\mathbf{Q}_k\mathbf{L}_{k+1}(\mathbf{L}_{k+1}^T\mathbf{L}_{k+1})^{-1}. \quad (6.41)$$

6.4 Le filtre UKF racine (SRUKF)

Filtre SEIK (Singular Evolutive Interpolated Kalman)

1. Initialisation
 - État du système \mathbf{x}_0^f et covariance d'erreur \mathbf{P}_0^f .
 - Diagonalisation $\mathbf{P}_0^f = \mathbf{L}_0 \mathbf{U}_0^f \mathbf{L}_0^T$.
2. Pour $t_k = 1, 2, \dots$

- (a) Analyse
 - Calcul de la matrice de covariance réduite

$$(\mathbf{U}_k^a)^{-1} = (\mathbf{U}_k^f)^{-1} + (\mathbf{H}_k \cdot \mathbf{L}_k)^T \mathbf{R}_k^{-1} (\mathbf{H}_k \cdot \mathbf{L}_k)$$

- Calcul du gain $\mathbf{K}_k^* = \mathbf{L}_k^T \mathbf{U}_k^a (\mathbf{H}_k^T \cdot \mathbf{L}_k) \mathbf{R}_k^{-1}$
- Calcul de l'estimé de l'analyse

$$\mathbf{x}_k^a = \mathbf{x}_k^f + \mathbf{K}_k^* (\mathbf{z}_k - H_k(\mathbf{x}_k^f))$$

- (b) Prévission
 - Tirage exact du second ordre à partir de x_k^a et U_k^a

$$(\mathbf{x}_k^a)_i = \mathbf{x}_k^a + \sqrt{r+1} \mathbf{L}_k C_{k-1} (\Omega_k)_i^T$$

- Propagation

$$(\mathbf{x}_{k+1}^f)_i = M_{k+1} (\mathbf{x}_k^a)_i$$

- Calcul des nouvelles directions

$$\mathbf{L}_{k+1} = \left[\begin{array}{ccc} (\mathbf{x}_{k+1}^f)_1 & \cdots & (\mathbf{x}_{k+1}^f)_{r+1} \end{array} \right] \mathbf{T}$$

- Calcul de la matrice de covariance d'erreur réduite

$$\mathbf{U}_{k+1}^f = (r+1) \mathbf{T}^T \mathbf{T} + (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1} \mathbf{L}_{k+1}^T \mathbf{Q}_k \mathbf{L}_{k+1} (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1}$$

Troisième partie

Quelques problèmes et solutions

Chapitre 7

Quelques problèmes et solutions

- 7.1 Linéarisation, boucle interne et méthode incrémentale
- 7.2 Sous-estimation des erreurs et inflation
- 7.3 Sous-échantillonnage et localisation

Quatrième partie

Modélisation inverse

Dans cette partie est introduite la problématique de la modélisation inverse en géophysique. Les idées principales sont présentées sur un exemple simplifié d'assimilation de radiances satellitaires dans l'infra-rouge. Le mauvais conditionnement des opérateurs d'inversion est souligné et on évoque des remèdes à ce problème, dont la régularisation. La notion de quantification de l'information extraite des données est exposée sommairement. La présentation suit étroitement celle de Clive Rodgers [16].

Chapitre 8

La modélisation inverse en géophysique : exemple de l'assimilation de radiances et concepts associés

Dans ce chapitre, quelques idées clés associées aux problèmes inverses en géophysique sont présentées. Le problème sera très progressivement introduit dans le cadre de l'assimilation de données satellitaires (de l'inversion de radiances en particulier). Le problème physique sera d'abord posé : quelques notions du transport radiatif seront rappelées (résumées de [18]). Une fois le problème posé, quelques concepts propres à la modélisation inverse seront introduits. La notion de contenu en information sera également développée : conditionnement, degrés de liberté et mesure du contenu en information. Dans le domaine applicatif de la physique atmosphérique, les ouvrages [16, 5] constituent des références importantes. D'autres monographies, plus mathématiques, comme [9, 25] peuvent également être consultées. Les problèmes inverses sont également étudiés en physique de la Terre solide [20, 21]. Le canevas suivi est fortement inspiré des premiers chapitres de l'excellente monographie de Clive Rodgers [16].

8.1 Le sondage de l'atmosphère : la physique

Une application de la mesure par satellite est le sondage de l'atmosphère, pour en explorer le profil en température, en humidité, ou bien en déterminer la distribution des gaz traces (comme l'ozone). Or les instruments d'un satellite mesurent en général des radiances dans différentes longueurs d'ondes, mais n'effectuent pas de mesures in situ. Ils sont en orbite à des altitudes bien supérieures au sommet de l'atmosphère (environ 100 km), de 400 à 800 km pour les orbites basses (satellites défilants) jusqu'à 36000 km pour les plate-formes géostationnaires. Il est donc nécessaire de remonter aux quantités inaccessibles directement (colonnes, profils). Cela constitue un problème inverse.

Nous allons considérer la simplification d'un problème typique de sondage atmosphérique : la reconstruction d'un profil de température à partir de mesures de radiances effectuées par un sondeur NADIR (opérant dans l'infrarouge). Nous n'introduisons que les notions

nécessaires à une exposition circonstanciée du problème inverse.

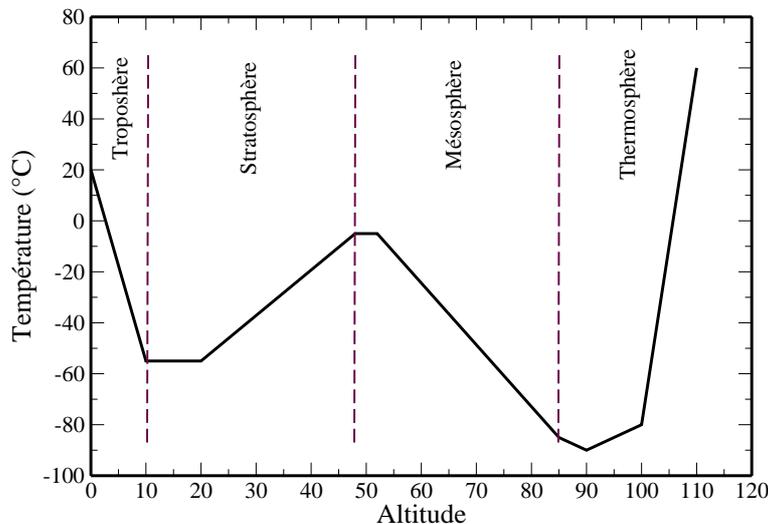
8.1.1 Profil vertical de température dans l'atmosphère

En partant du sol, la première couche atmosphérique conséquente est la troposphère. Son épaisseur varie de 8 kms au dessus des pôles à 18 kms au dessus de l'équateur. Les minima de température sont de 220 K au dessus des pôles et 190 K au dessus de l'équateur dans cette couche. Le gradient moyen de température est de -6.5 K par km.

Au sommet de la troposphère se trouve la tropopause qui mène à la stratosphère. Cette dernière s'étend jusqu'à 50 kms d'altitude. La température y est constante, puis croissante, jusqu'à atteindre des températures typiques de 270 K. Cette augmentation est causée par la photolyse de l'ozone et de l'oxygène moléculaire par le rayonnement ultraviolet. Le pic de température est atteint à la stratopause.

Vient ensuite la mésosphère jusqu'à 86 kms. La température y est décroissante jusqu'à 170 K. Au delà, dans la thermosphère et l'ionosphère, les espèces sont très diluées (typiquement 10^{13} mol/cm³, contre 10^{19} mol/cm³ au niveau de la mer) voire ionisées. Il n'y a plus d'équilibre thermodynamique et la notion de température est abstraite.

Les espèces majoritaires (N₂ et O₂) présentent un rapport de mélange uniforme jusqu'à une altitude de 90 kms (homéosphère). C'est pourquoi on peut considérer que du point de vue chimique l'atmosphère se termine à une altitude d'environ 100 kms.



8.1.2 Processus radiatifs dans l'atmosphère et définition des radiances

Les processus radiatifs sont cruciaux pour la dynamique de l'atmosphère. Le rayonnement solaire est de loin le premier forçage de l'atmosphère. Le profil de température est déterminé par le flux reçu mais aussi par l'ensemble des processus physiques faisant intervenir, directement ou indirectement, ce rayonnement. Les espèces chimiques interagissent avec le rayonnement au niveau moléculaire par trois voies principales : l'absorption, la diffusion et l'émission.

Toute théorie de transfert radiatif doit donc rendre compte de ses processus physico-chimiques.

Le rayonnement reçu par un détecteur (en particulier les instruments d'un satellite) est appelé **radiance**. Comme la notion de rayonnement est directionnelle, la définition de la notion de radiances n'est pas immédiate. En voici donc un résumé succinct.

On rappelle qu'un angle solide $d\Omega$ correspond à une surface de la sphère de rayon unité découpée par un cône de sommet le centre de la sphère. Un angle solide est exprimé en stéradians. Lorsque la surface intersectée correspond à la sphère complète, alors $\Omega = 4\pi$. Pour un angle solide infinitésimal, on a :

$$d\Omega = \sin\theta d\phi d\theta, \quad (8.1)$$

où θ est l'angle zénithal et ϕ est l'angle azimutal.

Considérons l'énergie reçue par l'élément d'aire dA pendant l'intervalle de temps dt , dans la partie du spectre $[\lambda, \lambda + d\lambda]$ en provenance de l'angle solide $d\Omega$:

$$dE_\lambda = I_\lambda \cdot dt \cdot d\lambda \cdot d\Omega \cdot \cos\theta \cdot dA. \quad (8.2)$$

Ceci définit la radiance I_λ .

8.1.3 Quantification du rayonnement et rayonnement du corps noir

On rappelle que les niveaux électroniques d'un atome sont quantifiés. Les niveaux de l'atome d'hydrogène sont de la forme $E = -E_0/n^2$ avec $E_0 = 13,6$ eV et n l'indice entier du niveau électronique. Pour les molécules composées de plusieurs atomes se rajoutent de nouveaux niveaux énergétiques. Ils correspondent à des degrés de liberté moléculaires, comme la rotation.

Selon le principe de quantification du rayonnement, le rayonnement ne peut interagir avec un atome que par l'intermédiaire de ces niveaux. Ce qui a conduit Planck à écrire

$$\Delta E = \frac{hc}{\lambda} = h\nu, \quad (8.3)$$

où $h = 6.63 \times 10^{-34}$ Js est la constante de Planck, $c = 3 \times 10^8$ m s⁻¹ est la vitesse de la lumière, et ν est la fréquence du rayonnement associée à la longueur d'onde λ . Toutefois les raies dans le spectre de ces espèces chimiques ne sont pas nécessairement fines car elles peuvent être élargies par différents processus dynamiques (effet Doppler).

Une conséquence des principes de la physique quantique et de la physique statistique est le rayonnement de corps noir. À l'équilibre thermodynamique, le maximum d'énergie qu'un corps peut émettre par unité de surface et de temps (et par unité de longueur d'onde) est l'énergie émise par ce que l'on appelle le *corps noir* :

$$B(\lambda, T) = \frac{8\pi hc}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda kT}} - 1}, \quad (8.4)$$

où $k = 1.38 \times 10^{-23}$ J K⁻¹ est la constante de Boltzmann. En première approximation, le Soleil, comme la Terre, peuvent être considérés comme des corps noir, et leur spectre approché par la distribution de Planck.

À température fixée T , $B(\lambda, T)$ a un maximum en $\lambda_{\max} = 2.90 \times 10^6/T$, exprimée en nanomètres si T est en Kelvin. C'est la loi de déplacement de Wien : un corps chaud

rayonne plutôt dans les courtes longueurs d'ondes. Le flux total d'énergie radiative émis par le corps noir est

$$B(T) = \int_0^{\infty} d\lambda B(\lambda, T) = \sigma T^4 \quad (8.5)$$

C'est la loi de Stefan-Boltzmann.

Deux exemples essentiels de distribution de Planck sont donnés par les rayonnements de corps noir du Soleil et de la Terre. Ceux-ci sont représentés sur la figure 8.1.

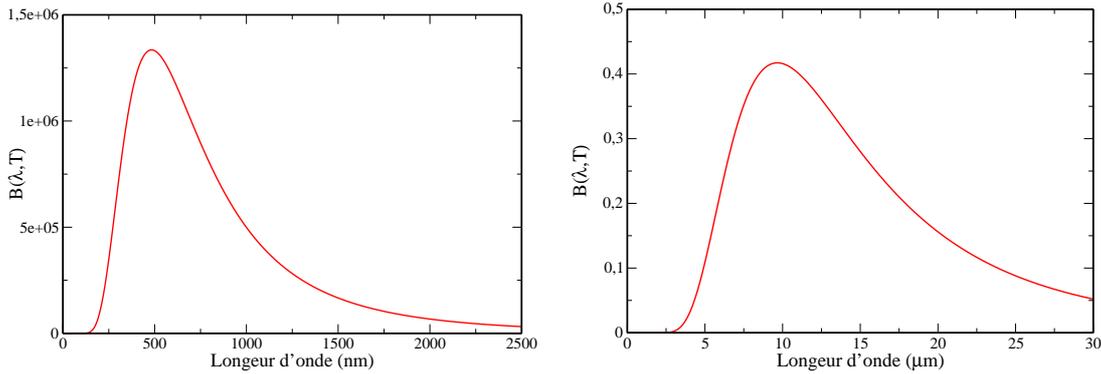


FIGURE 8.1 – Radiation de corps noir du Soleil ($T \simeq 6000K$) à gauche et radiation de corps noir de la Terre ($T \simeq 300K$) à droite.

À partir de la distribution de Planck est définie l'émissivité d'un corps. Considérons le rayonnement effectif d'un corps

$$E(\lambda, T) = \epsilon_{\lambda}(T)B(\lambda, T). \quad (8.6)$$

L'émissivité correspond donc au rapport entre le rayonnement effectif d'un corps et le rayonnement théorique de corps noir. Pour le sol nu $\epsilon \sim 0.94$, alors que pour la mer $\epsilon \sim 1$.

8.1.4 Loi de Beer-Lambert et loi de Kirchoff

Le rayonnement est donc altéré lors de la traversée d'un milieu. En particulier il est affaibli par les processus d'absorption. La loi de Beer-Lambert définit l'atténuation du rayonnement en fonction de l'altitude :

$$\frac{dI_{\lambda}(z)}{dz} = -a_{\lambda}(z)I_{\lambda}(z) \quad (8.7)$$

où a_{λ} est le coefficient d'absorption à la longueur d'onde λ . Lorsqu'on souhaite davantage se concentrer sur le trajet optique, plutôt que sur l'altitude, il est préférable de définir l'épaisseur optique

$$ds = a_{\lambda}(z)dz, \quad (8.8)$$

de sorte que (changement de variable $z \rightarrow s$)

$$\frac{dI_{\lambda}(s)}{ds} = -I_{\lambda}(s). \quad (8.9)$$

La loi de Kirchoff est une approximation qui consiste à dire qu'à l'équilibre thermodynamique local, l'émissivité à la longueur d'onde λ vaut le coefficient d'absorption à cette même longueur d'onde :

$$\epsilon_\lambda = a_\lambda . \quad (8.10)$$

8.1.5 La diffusion

La diffusion est un processus plus complexe que les précédents. La nature de cette diffusion dépend de façon déterminante du rapport entre la longueur du rayonnement (λ) et la taille des molécules / particules (d) qui diffusent ce rayonnement.

- si la taille des particules est totalement négligeable par rapport à la longueur d'onde, on parle de diffusion isotrope. La distribution angulaire est sphérique autour du diffuseur.
- si $d \ll \lambda$, ce qui est le cas pour un gaz de molécules, on parle de diffusion de Rayleigh. L'intensité du rayonnement diffusé est

$$I(\theta, R) = I_0 \frac{\pi^4 d^6}{8r^2 \lambda^4} \left(\frac{m^2 - 1}{m^2 + 1} \right)^2 (1 + \cos^2 \theta) . \quad (8.11)$$

I_0 est l'intensité du rayonnement incident. m est l'indice de réfraction qui dépend de la nature du gaz. θ et R déterminent la direction et la distance d'observation. La distribution angulaire est axisymétrique, en forme de cacahuète.

- si $d \gg \lambda$, typiquement pour des gouttelettes d'eau, des cristaux de glaces, les lois de l'optique géométrique s'applique. La diffusion ne dépend plus alors de la longueur d'onde.
- si $d \sim \lambda$, typiquement pour les aérosols atmosphériques, aucune approximation simple n'est possible, et on recourt à la théorie de Mie. L'intensité du rayonnement diffusé est alors de la forme

$$I(\theta, R) = I_0 \frac{\lambda^2}{8\pi R^2} (i_1 + i_2) , \quad (8.12)$$

où i_1 et i_2 sont les paramètres d'intensité de Mie qui sont des fonctions complexes de d/λ , θ et m . Il n'y a pas de symétrie particulière similaire à celle observée pour la diffusion de Rayleigh (excepté l'axisymétrie).

Une conséquence de cela est une faible diffusion du rayonnement infrarouge par les gaz de l'atmosphère, puisque la diffusion de Rayleigh est proportionnelle à $1/\lambda^4$.

8.1.6 Équation du transfert radiatif pour le rayonnement infrarouge

Pour le rayonnement infrarouge, la diffusion peut donc être négligée, ce qui simplifie considérablement l'équation du transfert radiatif. On a donc, dans cette partie du spectre :

$$\frac{dI_\lambda(z)}{dz} = -a_\lambda(z)I_\lambda(z) + \epsilon_\lambda(T)B(\lambda, T) . \quad (8.13)$$

Invokant la loi de Kirchoff, et en utilisant l'épaisseur optique $s(z, \lambda) = \int_0^z a_\lambda(z')dz'$, on obtient

$$\frac{dI_\lambda(s)}{ds} = -I_\lambda(s) + B(\lambda, T(s)) . \quad (8.14)$$

La solution de cette équation est immédiate :

$$I(\lambda, s) = I(\lambda, 0)e^{-s} + \int_0^s B(\lambda, T(s')) e^{s'-s} ds' \quad (8.15)$$

Le premier terme correspond à un terme d'amortissement pur du rayonnement. Le second terme correspond à la contribution (amortie) du rayonnement propre du milieu traversé.

8.1.7 Définition de la transmittance

La transmittance est une coordonnée transformée de l'épaisseur optique qui tient compte de l'amortissement. La transmittance entre l'altitude z et le sommet de l'atmosphère (donc essentiellement $z = \infty$ du point de vue optique), est solution de

$$\frac{d\tau(z)}{dz} = a_\lambda(z)\tau(z), \quad (8.16)$$

avec la condition $\tau(\infty) = 1$. Il en résulte :

$$\tau(z) = \tau(\infty) \exp\left(-\int_z^\infty a(\lambda, z) dz\right) = \tau(\infty) \exp(s(z) - s(\infty)). \quad (8.17)$$

Si l'on suppose $\tau(0) = 0$, c'est-à-dire si le rayonnement depuis le niveau de la mer est totalement absorbé, alors la radiance au sommet de l'atmosphère (et donc du point de vue de l'instrument satellitaire) vaut

$$L(\lambda) = \int_0^\infty dz B(\lambda, T(z)) \frac{d\tau(\lambda, z)}{dz} = I(\lambda, \infty). \quad (8.18)$$

8.1.8 Fonctions de poids

Un instrument opère dans une petite gamme de longueur d'onde, centrée par exemple sur une longueur d'onde moyenne $\bar{\lambda}$. Un sondeur infrarouge embarqué sur un satellite mesure le rayonnement infrarouge terrestre (puisque c'est dans ce domaine qu'émet essentiellement la Terre), après interaction avec la totalité de l'atmosphère. Le rayonnement de corps noir du milieu traversé ne varie que faiblement sur la gamme de longueurs d'onde sondée. Cela n'est pas le cas de l'épaisseur optique (et donc de la transmittance), puisqu'elle dépend de la physico-chimie des espèces traversées (donc de nombreuses raies plus ou moins élargies). Supposons que l'instrument embarqué mesure la radiance L_i à la longueur d'onde λ_i . On définit alors la fonction

$$K_i(z) = \frac{d\tau(\lambda_i, z)}{dz} \quad (8.19)$$

appelée **fonction de poids**. Compte tenu des approximations utilisées (notamment $\tau(0) = 0$), on a $\int_0^\infty K_i(z) dz = 1$, ce qui justifie la terminologie de *poids*. La fonction de poids attribue une pondération aux différentes couches de l'atmosphère. Suivant l'exemple filé de C. Rodgers [16], la figure 8.2 correspond à des fonctions de poids synthétiques caractéristiques d'un sondage au NADIR. La figure 8.3 donne l'exemple réel des fonctions de poids de l'instrument AIRS.

La mesure de l'instrument à la longueur d'onde λ_i s'écrit donc approximativement

$$L_i = \int_0^\infty dz B(\bar{\lambda}, T(z)) K_i(z). \quad (8.20)$$

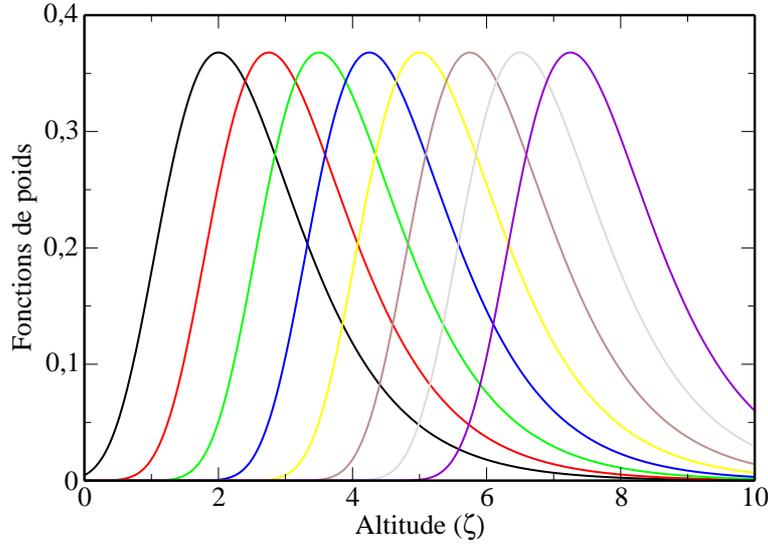


FIGURE 8.2 – Fonctions de poids typiques. $\zeta = -\ln(p/p_0)$ est une mesure de l'altitude. p est la pression et p_0 est la pression au sol.

8.2 Un problème inverse typique

8.2.1 L'équation du problème inverse

Le sondeur infrarouge permet de réaliser p mesures dans la gamme. Le problème inverse est donc le suivant : retrouver le profil de température $T(z)$, connaissant les résultats des mesures Eq.(8.20) pour $i = 1, \dots, p$.

Il est évident qu'on ne peut inverser un profil aussi fin que l'on souhaite. Il est donc nécessaire de discrétiser le profil de température en n niveaux. Une approche plus générale et plus abstraite consiste à utiliser des fonctions de base sur lesquelles décomposer les fonctions inconnues. Par exemple, on suppose que

$$B(\bar{\lambda}, T(z)) = \sum_{k=1}^n \omega_k W_k(z), \quad (8.21)$$

où $W_k(z)$ sont les fonctions de bases. Les ω_k sont les poids inconnus à reconstruire. Alors on a

$$L_i = \sum_{k=1}^n \omega_k \int_0^\infty dz W_k(z) K_i(z). \quad (8.22)$$

ou encore

$$L_i = \sum_{k=1}^n [\mathbf{C}]_{ik} \omega_k, \quad (8.23)$$

où \mathbf{C} est une matrice $p \times n$.

Les fonctions de base retenues pour l'exemple introduit auparavant, sont des fonctions polynomiales de $\zeta = -\ln(p/p_0)$ de la forme ζ^j avec $j = 0, \dots, m-1$.

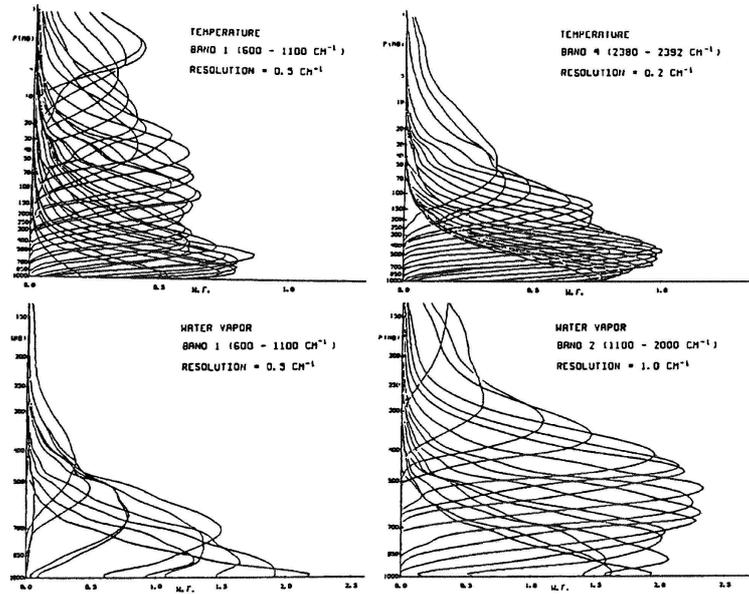


FIGURE 8.3 – Fonctions de poids de l’instrument infra-rouge AIRS de la plate-forme AQUA (d’après John Eyre)

8.2.2 Conditionnement du problème - Régularisation

In fine, le problème consiste donc à résoudre le système linéaire Eq.(8.23). Celui-ci peut être mal posé si $p < n$: le système linéaire est sous-déterminé. Si $p \geq n$, alors le problème inverse est correctement posé. Le système linéaire est déterminé, voire sur-déterminé.

Cependant, même s’il on dispose d’autant de mesures (ou plus) que d’inconnues, le problème peut être mal conditionné pour des raisons physiques, parce que les mesures ne sont pas assez discriminantes pour retrouver le profil désiré. Cela se traduit mathématiquement par une matrice \mathbf{C} mal conditionnée¹. Ce problème se pose parce que le modèle physique et les mesures sont entachées d’erreurs. En l’absence de ces erreurs, le problème du conditionnement n’est que purement numérique et n’intervient qu’à des niveaux de précision physiquement irréalistes (dans un cadre géophysique).

Pour simplifier supposons que $p = n$ et que le problème est suffisamment bien conditionné. \mathbf{C} est alors carrée et inversible (y compris du point de vue numérique). Alors les poids solutions sont

$$\omega_k^* = \sum_{i=1}^p [\mathbf{C}^{-1}]_{ki} L_i \tag{8.24}$$

donc

$$B^*(\bar{\lambda}, T(z)) = \sum_{i=1}^p \left\{ \sum_{k=1}^n [\mathbf{C}^{-1}]_{ki} W_k(z) \right\} \cdot L_i = \sum_{i=1}^p G_i(z) L_i. \tag{8.25}$$

$G_i(z)$ est la *fonction de contribution* de la mesure i à la reconstruction. Supposons

1. Pour une matrice symétrique définie positive, une mesure de son conditionnement est le rapport entre la plus grande et la plus petite valeur propre. Plus ce rapport est élevé, moins bon est le conditionnement.

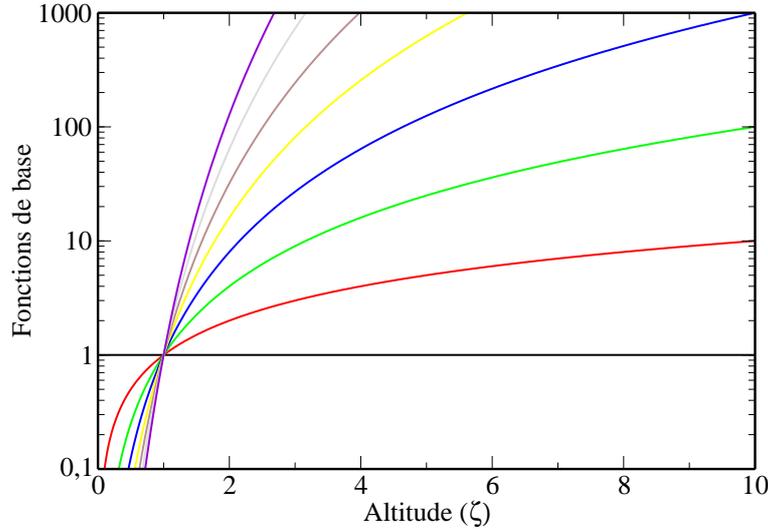


FIGURE 8.4 – Fonctions de base sur lesquelles est décomposé le profil en température

qu'une erreur soit commise sur $L_i : L_i \rightarrow L_i + \varepsilon_i$, alors l'erreur commise sur B est $G_i(z)\varepsilon_i$. Ceci permet de qualifier le mauvais conditionnement par rapport à une mesure et une altitude. On a représenté sur la figure 8.5, les fonctions de contributions correspondantes à l'exemple synthétique introduit avec la figure 8.2.

Si on suppose que les mesures de radiances sont indépendantes, mais possèdent une variance similaire, disons σ^2 , alors la variance de la solution reconstruite est $\sum_{i=1}^p G_i^2(z)\sigma^2$. Par conséquent, $\sum_{i=1}^p G_i^2(z)$ peut être interprété comme un **facteur d'amplification** de l'erreur σ^2 . Poursuivant l'exemple synthétique introduit plus haut, la figure 8.6 représente le facteur d'amplification d'erreur.

La solution du problème synthétique est représentée sur la figure 8.7. Y est également représenté le profil supposé réel de la température en fonction de l'altitude.

8.3 Solution standard et conséquences

Le cas $m = n$ est assez singulier en pratique. En effet lorsqu'on considère un système réel (où sévit de l'erreur de toute nature), la transition $p = n$ entre un système sous-déterminé et un système sur-déterminé n'est pas nécessairement prononcée.

La généralisation du problème inverse Eq.(8.22) est donnée par :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \epsilon, \quad (8.26)$$

où \mathbf{y} est le vecteur d'observation, \mathbf{H} est la matrice d'observation de dimension $p \times n$, qui contient le modèle physique, \mathbf{x} est le vecteur d'état et ϵ est l'erreur d'observation / modèle.

Une solution variationnelle du problème est obtenue en minimisant la fonction coût sur le vecteur d'état \mathbf{x}

$$\mathcal{L} = \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) \quad (8.27)$$

\mathbf{R} est la matrice de covariance d'erreur, obtenue de la connaissance statistique a priori de

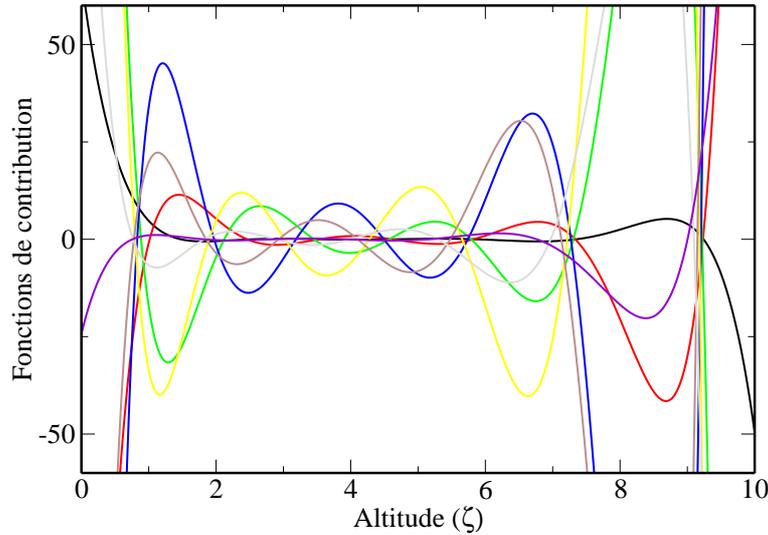


FIGURE 8.5 – Fonctions de contribution typiques

ϵ . Si le problème original est sous-déterminé, il y aura de nombreux minima (un espace affine de solutions). Et pourtant, on souhaite une réponse unique au problème inverse !

Un artifice mathématique permet de contourner ce problème. Il s'agit du procédé de *régularisation*. La justification de son utilisation en assimilation de données tient à son interprétation physique comme nous allons le voir. La fonctionnelle régularisée est

$$\mathcal{L} = \frac{1}{2} (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}) + \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \mathbf{B}^{-1} (\mathbf{x} - \boldsymbol{\mu}) , \quad (8.28)$$

où le second terme constitue la régularisation (dite de *Tikhonov*). Physiquement, ce terme s'interprète comme une ébauche, c'est-à-dire de l'information a priori détenue avant la mesure. \mathbf{B} est la matrice de covariance de l'erreur attachée à cette ébauche.

Si l'information a priori est suffisamment discriminante, on peut espérer qu'une unique solution sera minimum de cette fonctionnelle. Il suffit ici que \mathbf{B} soit inversible, c'est-à-dire que \mathbf{B} existe.

La solution de ce problème variationnel (confer chapitre 1) s'obtient ainsi :

$$\mathbf{P}^{-1} = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1} , \quad (8.29)$$

$$\mathbf{x}^* = \boldsymbol{\mu} + (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\boldsymbol{\mu}) . \quad (8.30)$$

Comme on l'a vu, \mathbf{P} s'interprète comme la matrice de covariance d'erreur d'analyse. Dans ce contexte, il s'agit de l'erreur statistique supposée que l'on commet en utilisant le résultat de l'inversion. On peut récrire la solution au moyen de la formule de Sherman-Morrison-Woodbury (voir chapitre 1) :

$$\mathbf{x}^* = \boldsymbol{\mu} + \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1} (\mathbf{y} - \mathbf{H}\boldsymbol{\mu}) . \quad (8.31)$$

En l'absence d'ébauche $\mathbf{B}^{-1} = \mathbf{0}$ et si $\mathbf{R} = \mathbf{I}$, alors $\mathbf{x}^* = \mathbf{H}^T (\mathbf{H}\mathbf{H}^T)^{-1} \mathbf{y}$. $\mathbf{H}^T (\mathbf{H}\mathbf{H}^T)^{-1}$ est appelée l'*inverse généralisé* (ou de Moore-Penrose), puisqu'il généralise l'inverse d'une matrice carrée non-singulière.

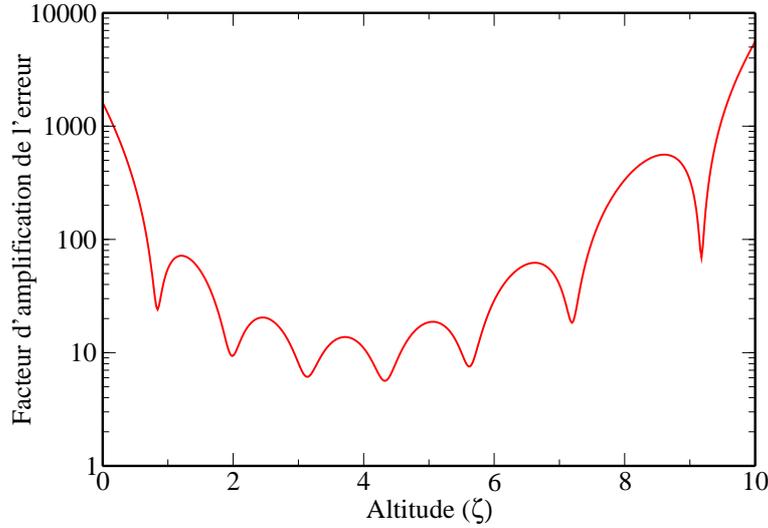


FIGURE 8.6 – Facteur d'amplification d'erreur

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

$$\mathbf{x}^* - \mu = \mathbf{K}(\mathbf{y} - \mathbf{H}\mu) = \mathbf{K}\mathbf{H}(\mathbf{x}_t - \mu) + \mathbf{K}\epsilon, \quad (8.32)$$

où \mathbf{x}_t est la vraie valeur du vecteur d'état. La valeur inférée est ainsi reliée à la valeur réelle au travers de l'opérateur $\mathbf{A} = \mathbf{K}\mathbf{H}$, appelé matrice d'*averaging kernel* :

$$\frac{\partial x_k^*}{\partial x_l^t} = [\mathbf{A}]_{kl}. \quad (8.33)$$

La matrice associée $\mathbf{H}\mathbf{K}$, de même trace, exprime le même type de sensibilité mais dans l'espace des observations. Elle est appelée matrice d'*influence*, mais aussi matrice *DRM* (ou matrice de résolution) pour *Data Resolution Matrix*.

On souhaite maintenant analyser le contenu en information de cette inversion. Pour cela, on normalise les matrices de covariance d'erreur. On définit ainsi $\hat{\mathbf{H}} = \mathbf{R}^{-1/2} \mathbf{H} \mathbf{B}^{1/2}$, $\hat{\mathbf{y}} = \mathbf{R}^{-1/2} \mathbf{y}$, $\hat{\mathbf{x}} = \mathbf{B}^{-1/2} \mathbf{x}$, $\hat{\epsilon} = \mathbf{R}^{-1/2} \epsilon$ et $\hat{\mu} = \mathbf{B}^{-1/2} \mu$. Alors, on obtient

$$\hat{\mathbf{y}} = \hat{\mathbf{H}} \hat{\mathbf{x}} + \hat{\epsilon}. \quad (8.34)$$

La fonctionnelle se réécrit

$$\mathcal{L} = \frac{1}{2} (\hat{\mathbf{y}} - \hat{\mathbf{H}} \hat{\mathbf{x}})^T (\hat{\mathbf{y}} - \hat{\mathbf{H}} \hat{\mathbf{x}}) + \frac{1}{2} (\hat{\mathbf{x}} - \hat{\mu})^T (\hat{\mathbf{x}} - \hat{\mu}), \quad (8.35)$$

Ainsi, les matrices de covariances d'erreur d'ébauche et d'observation sont normalisées à la matrice unité.

On utilise alors la *décomposition en valeurs singulières* de la matrice $\hat{\mathbf{H}}$. $\hat{\mathbf{H}}$ peut s'écrire $\mathbf{U} \mathbf{\Lambda} \mathbf{V}^T$, avec $\mathbf{\Lambda}$ une matrice diagonale de taille $p \times p$ qui contient les valeurs singulières de $\hat{\mathbf{H}}$, \mathbf{U} une matrice orthogonale de taille $p \times p$ et \mathbf{V} une matrice de taille $n \times p$, telle que $\mathbf{V}^T \mathbf{V} = \mathbf{I}$. On effectue alors une rotation (pseudo-rotation) dans l'espace des observations (des états) qui préserve les propriétés statistiques des variables :

$$\mathbf{y}' = \mathbf{U}^T \hat{\mathbf{y}} \quad \mathbf{x}' = \mathbf{V}^T \hat{\mathbf{x}} \quad \epsilon' = \mathbf{U}^T \hat{\epsilon}, \quad (8.36)$$

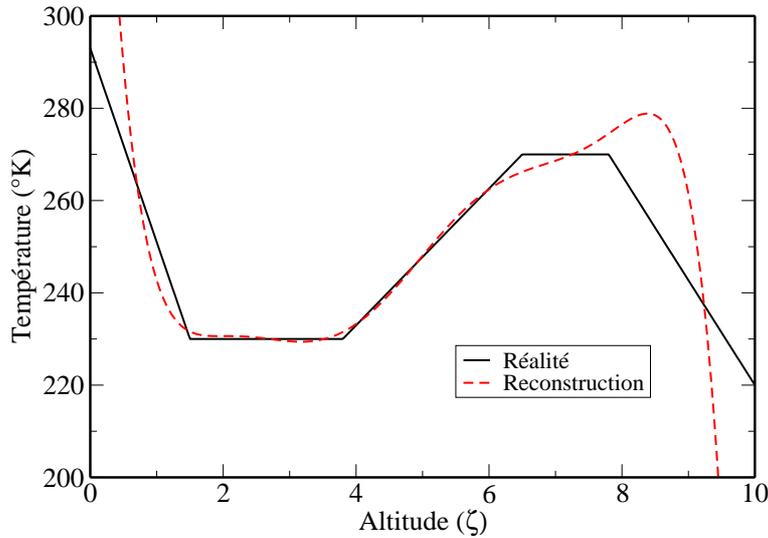


FIGURE 8.7 – Reconstruction du profil de température et comparaison au profil réel

ce qui mène à

$$\mathbf{y}' = \mathbf{\Lambda} \mathbf{x}' + \epsilon' . \quad (8.37)$$

L'équation d'observation a été réduite formellement à une équation qui décrit des variables indépendantes de statistiques normalisées. Alors, a priori, les fluctuations des mesures sont décrites par

$$\mathbf{E}(\mathbf{y}' \mathbf{y}'^T) = \mathbf{E}[(\mathbf{\Lambda} \mathbf{x}' + \epsilon')(\mathbf{\Lambda} \mathbf{x}' + \epsilon')^T] = \mathbf{\Lambda}^2 + \mathbf{I} . \quad (8.38)$$

Elles sont la somme d'une contribution dues aux erreurs (observation et modèle) normalisée à l'unité ici, et des variations physiquement significatives et qui sont décrites par les valeurs singulières contenues dans la diagonale de $\mathbf{\Lambda}$. Si une valeur singulière est nettement inférieure à un, on peut donc penser qu'elle sera noyée dans le bruit. Au contraire, si elle est plus grande que l'unité, elle dominera le bruit et contribuera à amener de l'information pour l'inversion. L'inversion sera de qualité si le nombre de ces valeurs significatives, c'est-à-dire le rang effectif du problème, atteint le nombre de variables inconnues.

8.4 Quantifier l'information

Cette analyse semi-qualitative peut être généralisée par une approche plus quantitative permettant de déterminer objectivement la richesse d'un ensemble d'information.

8.4.1 Nombre de degrés de libertés

Considérons la variable aléatoire $\chi^2 = 2\mathcal{L}$, c'est-à-dire

$$\chi^2 = (\mathbf{y} - \mathbf{H} \mathbf{x}^a)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H} \mathbf{x}^a) + (\mathbf{x}^a - \boldsymbol{\mu})^T \mathbf{B}^{-1} (\mathbf{x}^a - \boldsymbol{\mu}) , \quad (8.39)$$

où \mathbf{x}^a est vue comme une variable aléatoire, dépendant des fluctuations de l'ébauche et de l'observation. La démarche et les notations sont similaires à l'analyse BLUE du chapitre

1. On peut alors montrer que l'espérance de χ^2 est égal à p , le nombre de degrés de liberté indépendants. Ces p degrés de libertés se répartissent entre ceux attribuables au vecteur d'état, i.e.

$$d_e = \mathbb{E} \left[(\mathbf{x}^a - \boldsymbol{\mu})^T \mathbf{B}^{-1} (\mathbf{x}^a - \boldsymbol{\mu}) \right], \quad (8.40)$$

et ceux attribuables au bruit, i.e.

$$d_b = \mathbb{E} [\boldsymbol{\epsilon}^T \mathbf{R}^{-1} \boldsymbol{\epsilon}] = \mathbb{E} \left[(\mathbf{y} - \mathbf{H}\mathbf{x}^a)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^a) \right]. \quad (8.41)$$

d_e est donc le nombre de degrés de liberté du signal. Suivant l'approche du chapitre 1, le calcul de l'espérance de l'écart pondéré à l'ébauche conduit à

$$\begin{aligned} d_e &= \mathbb{E} \left[(\mathbf{x}^a - \boldsymbol{\mu})^T \mathbf{B}^{-1} (\mathbf{x}^a - \boldsymbol{\mu}) \right] \\ &= \mathbb{E} \left[\text{Tr} \left\{ \mathbf{B}^{-1} (\mathbf{x}^a - \boldsymbol{\mu}) (\mathbf{x}^a - \boldsymbol{\mu})^T \right\} \right] \\ &= \mathbb{E} \left[\text{Tr} \left\{ \mathbf{B}^{-1} \mathbf{K} (\mathbf{y} - \mathbf{H}\boldsymbol{\mu}) (\mathbf{y} - \mathbf{H}\boldsymbol{\mu})^T \mathbf{K}^T \right\} \right] \\ &= \text{Tr} \left\{ \mathbf{B}^{-1} \mathbf{K} \mathbb{E} \left[(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b) (\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)^T \right] \mathbf{K}^T \right\} \\ &= \text{Tr} \left\{ \mathbf{B}^{-1} \mathbf{K} (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) \mathbf{K}^T \right\} \\ &= \text{Tr} (\mathbf{B}^{-1} \mathbf{B} \mathbf{H}^T \mathbf{K}^T) \\ &= \text{Tr} (\mathbf{K}\mathbf{H}) = \text{Tr} (\mathbf{A}), \end{aligned} \quad (8.42)$$

où l'on a utilisé $\mathbf{K} = \mathbf{B}\mathbf{H}^T (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}$. Ceci renforce l'importance de l'opérateur \mathbf{A} , l'*averaging kernel*. Pour les fluctuations dues au bruit, on a :

$$\begin{aligned} d_b &= \mathbb{E} \left[(\mathbf{y} - \mathbf{H}\mathbf{x}^a)^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^a) \right] \\ &= \mathbb{E} \left[\text{Tr} \left\{ \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x}^a) (\mathbf{y} - \mathbf{H}\mathbf{x}^a)^T \right\} \right] \\ &= \mathbb{E} \left[\text{Tr} \left\{ \mathbf{R}^{-1} (\mathbf{e}^o - \mathbf{H}\mathbf{e}^a) (\mathbf{e}^o - \mathbf{H}\mathbf{e}^a)^T \right\} \right]. \end{aligned} \quad (8.43)$$

Mais l'erreur d'analyse est $\mathbf{e}^a = \mathbf{e}^b + \mathbf{K} (\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)$, de sorte que

$$\mathbf{e}^o - \mathbf{H}\mathbf{e}^a = (\mathbf{I} - \mathbf{H}\mathbf{K}) (\mathbf{e}^o - \mathbf{H}\mathbf{e}^b). \quad (8.44)$$

Alors

$$\begin{aligned} d_b &= \text{Tr} \left\{ \mathbf{R}^{-1} (\mathbf{I} - \mathbf{H}\mathbf{K}) \mathbb{E} \left[(\mathbf{e}^o - \mathbf{H}\mathbf{e}^b) (\mathbf{e}^o - \mathbf{H}\mathbf{e}^b)^T \right] (\mathbf{I} - \mathbf{H}\mathbf{K})^T \right\} \\ &= \text{Tr} \left\{ \mathbf{R}^{-1} (\mathbf{I} - \mathbf{H}\mathbf{K}) (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T) (\mathbf{I} - \mathbf{H}\mathbf{K})^T \right\} \\ &= \text{Tr} \left\{ \mathbf{R}^{-1} (\mathbf{R} + \mathbf{H}\mathbf{B}\mathbf{H}^T - \mathbf{H}\mathbf{B}\mathbf{H}^T) (\mathbf{I} - \mathbf{H}\mathbf{K})^T \right\} \\ &= \text{Tr} (\mathbf{I} - \mathbf{H}\mathbf{K}) = p - \text{Tr} (\mathbf{A}). \end{aligned} \quad (8.45)$$

On vérifie bien que $d_e + d_b = p$.

En termes de valeurs singulières λ_i de $\widehat{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}^{1/2}$, on a

$$d_e = \sum_{i=1}^m \frac{\lambda_i^2}{1 + \lambda_i^2} \quad d_b = \sum_{i=1}^m \frac{1}{1 + \lambda_i^2}, \quad (8.46)$$

car $\text{Tr}(\mathbf{A}) = \text{Tr} \left[\widehat{\mathbf{H}}^T \left(\mathbf{I} + \widehat{\mathbf{H}}\widehat{\mathbf{H}}^T \right)^{-1} \widehat{\mathbf{H}} \right]$. Ces formules confirment que les valeurs singulières de $\widehat{\mathbf{H}} = \mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}^{1/2}$ significativement supérieures à l'unité comptent comme autant de degrés de liberté qui apportent de l'information.

8.4.2 Résolution

On a vu que le nombre de degrés de liberté du signal est égal à la trace de la matrice d'*averaging kernel*. Ce nombre peut donc être décomposé comme la somme des éléments diagonaux de \mathbf{A} , chacun de ces éléments concernant un niveau vertical. L'élément de la diagonale affecté à un niveau peut être interprété comme le nombre de degrés de liberté de signal nécessaires à la reconstruction de ce niveau. L'inverse de ce nombre donne le nombre de niveaux qu'il est possible de reconstruire avec un degré de liberté du signal (en général bien inférieur à 1). Autrement dit, il donne une indication locale sur la résolution de l'inversion.

Une manière de quantifier cette notion de résolution en tenant compte de la matrice d'*averaging kernel* est de mesurer l'étalement de cette fonction (combien d'éléments $[\mathbf{A}]_{ij}$ sont-ils significativement non nuls pour j fixé ?). Comme elle peut être de signe quelconque, on s'intéresse plutôt au carré des entrées de la matrice $([\mathbf{A}]_{ij})^2$. Une notion de taille de résolution est fournie, par exemple, par la moyenne locale :

$$r_i = \frac{\sum_j |i - j| \times ([\mathbf{A}]_{ij})^2}{\sum_j ([\mathbf{A}]_{ij})^2}. \quad (8.47)$$

Plus l'*averaging kernel* est piqué (plus la matrice est concentrée sur la diagonale), plus faible est r_i . Si la matrice \mathbf{A} est purement diagonale, alors r_i est même nul.

8.4.3 Information au sens de Shannon - Entropie

Suivant Boltzmann, puis Gibbs, Shannon a redéfini le concept d'entropie dans le contexte de la théorie de l'information. L'opposé de l'entropie s'identifie alors à une quantité d'information. On peut attribuer une entropie, donc un contenu en information à une densité de probabilité. En particulier, on sait calculer la différence de contenu en information entre deux densités, disons $p(x)$ et $q(x)$. Cette différence s'exprime par la formule de Kullback-Leibler :

$$K[p, q] = \int dx p(x) \ln \frac{p(x)}{q(x)}. \quad (8.48)$$

Cette formule peut être appliquée à l'inversion que l'on a opérée précédemment, à l'aide de l'analyse variationnelle. On peut considérer que la distribution initiale correspond à celle fournie par l'ébauche, et que la distribution finale est donnée par la description statistique

de l'analyse. On cherche donc à mesurer l'écart entre deux distributions, évalué par $\Delta\mathcal{I} = K[p(\mathbf{x}|\mathbf{y}), p(\mathbf{x})]$. $p(\mathbf{x})$ est la distribution a priori du vecteur d'état :

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\mathbf{B}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{B}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right) \quad (8.49)$$

et $p(\mathbf{x}|\mathbf{y})$ est la distribution a posteriori du vecteur d'état :

$$p(\mathbf{x}|\mathbf{y}) = \frac{1}{(2\pi)^{n/2}|\mathbf{P}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{x}^a)^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}^a)\right) \quad (8.50)$$

On a alors

$$\begin{aligned} \Delta\mathcal{I} &= K[p(\mathbf{x}|\mathbf{y}), p(\mathbf{x})] \\ &= \int_{\mathbf{x}} d\mathbf{x} p(\mathbf{x}|\mathbf{y}) \left\{ -\frac{1}{2}(\mathbf{x} - \mathbf{x}^a)^T \mathbf{P}^{-1}(\mathbf{x} - \mathbf{x}^a) - \frac{1}{2} \ln |\mathbf{P}| \right. \\ &\quad \left. + \frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{B}^{-1}(\mathbf{x} - \boldsymbol{\mu}) + \frac{1}{2} \ln |\mathbf{B}| \right\} \\ &= \frac{1}{2} \ln |\mathbf{B}\mathbf{P}^{-1}| + \frac{1}{2}(\mathbf{x}^a - \boldsymbol{\mu})^T \mathbf{B}^{-1}(\mathbf{x}^a - \boldsymbol{\mu}) + \frac{1}{2} \text{Tr}(\mathbf{P}\mathbf{B}^{-1} - \mathbf{I}) . \end{aligned} \quad (8.51)$$

Cette dernière ligne s'obtient à l'aide d'intégrations gaussiennes élémentaires en utilisant :

$$\begin{aligned} \int_{-\infty}^{\infty} d^n \mathbf{x} p(\mathbf{x}|\mathbf{y}) &= 1, \quad \int_{-\infty}^{\infty} d^n \mathbf{x} p(\mathbf{x}|\mathbf{y})(\mathbf{x} - \mathbf{x}^a) = 0 \\ \int_{-\infty}^{\infty} d^n \mathbf{x} p(\mathbf{x}|\mathbf{y}) \{(\mathbf{x} - \mathbf{x}^a)_i (\mathbf{x} - \mathbf{x}^a)_j\} &= [\mathbf{P}]_{ij} . \end{aligned} \quad (8.52)$$

Or il est très facile de vérifier que $\mathbf{P}\mathbf{B}^{-1} = \mathbf{I} - \mathbf{K}\mathbf{H} = \mathbf{I} - \mathbf{A}$. Donc

$$\frac{1}{2} \text{Tr} \ln (\mathbf{B}\mathbf{P}^{-1}) = -\frac{1}{2} \text{Tr} \ln (\mathbf{I} - \mathbf{A}) \quad \text{et} \quad \frac{1}{2} \text{Tr} (\mathbf{P}\mathbf{B}^{-1} - \mathbf{I}) = -\frac{1}{2} \text{Tr} (\mathbf{A}) . \quad (8.53)$$

Reste le troisième terme qui mesure l'information tirée de l'écart entre l'estimateur a priori et l'estimateur a posteriori. Or le gain $\Delta\mathcal{I}$ doit être compris comme le gain moyen sur un ensemble significatif d'expériences et d'analyses. On a alors

$$\begin{aligned} \text{E}[\Delta\mathcal{I}] &= -\frac{1}{2} \text{Tr} \ln (\mathbf{I} - \mathbf{A}) - \frac{1}{2} \text{Tr}(\mathbf{A}) + \text{E} \left[\frac{1}{2}(\mathbf{x}^a - \boldsymbol{\mu})^T \mathbf{B}^{-1}(\mathbf{x}^a - \boldsymbol{\mu}) \right] \\ &= -\frac{1}{2} \text{Tr} \ln (\mathbf{I} - \mathbf{A}) - \frac{1}{2} \text{Tr}(\mathbf{A}) + \frac{1}{2} \text{Tr}(\mathbf{A}) \\ &= -\frac{1}{2} \text{Tr} \ln (\mathbf{I} - \mathbf{A}) , \end{aligned} \quad (8.54)$$

où l'on a tenu compte du résultat Eq.(8.42). En termes de valeurs singulières λ_i de $\hat{\mathbf{H}} = \mathbf{R}^{-1/2} \mathbf{H}\mathbf{B}^{1/2}$, on a

$$\text{E}[\Delta\mathcal{I}] = \frac{1}{2} \sum_{i=1}^m \ln (1 + \lambda_i^2) . \quad (8.55)$$

8.5 Autres approches

8.5.1 Méthode de Backus-Gilbert

On a vu que l'on pouvait obtenir une idée de la résolution atteinte lors d'une reconstruction à l'aide de l'opérateur \mathbf{A} , la *matrice de résolution*, ou *averaging kernel*. Plutôt que de suivre une démarche bayésienne strictement probabiliste, on essaie de construire ici un estimateur de la solution de $\mathbf{y} = \mathbf{H}\mathbf{x}$, qui optimise la résolution. On souhaite donc que l'opérateur \mathbf{A} , qui est une matrice $n \times n$, soit aussi *proche* de la matrice identité \mathbf{I} , avec $[\mathbf{I}]_{kl} = \delta_{k,l}$, que possible.

Dans le cas d'une analyse BLUE, ou dans le cas (équivalent sous l'hypothèse de linéarité) d'une analyse de type moindres carrés, l'opérateur \mathbf{A} est déterminé par les conditions d'optimalité et d'absence de biais. On suppose que ce n'est pas le cas ici, et on cherche à calculer \mathbf{A} pour en déduire un autre estimateur.

Pour cela on suppose que l'estimateur est une combinaison linéaire des mesures

$$\mathbf{x}^a = \mathbf{\Omega}\mathbf{y}, \quad (8.56)$$

où $\mathbf{\Omega}$ est une matrice $n \times p$ à déterminer. Or on a $\mathbf{y} = \mathbf{H}\mathbf{x}^t$, ce qui entraîne $\mathbf{x}^a = \mathbf{\Omega}\mathbf{H}\mathbf{x}^t$. Avec $\mathbf{A} = \mathbf{\Omega}\mathbf{H}$ l'opérateur de résolution. On cherche alors la solution du problème de minimisation

$$\min_{\mathbf{\Omega}} \text{Tr} \left\{ (\mathbf{A} - \mathbf{I}) (\mathbf{A} - \mathbf{I})^T \right\}. \quad (8.57)$$

On a

$$\begin{aligned} \text{Tr} (\mathbf{A} - \mathbf{I}) (\mathbf{A} - \mathbf{I})^T &= \text{Tr} (\mathbf{\Omega}\mathbf{H} - \mathbf{I}) (\mathbf{H}^T\mathbf{\Omega}^T - \mathbf{I}) \\ &= \text{Tr} (\mathbf{\Omega}\mathbf{G}\mathbf{\Omega}^T - 2\mathbf{\Omega}\mathbf{H} + \mathbf{I}), \end{aligned} \quad (8.58)$$

où $\mathbf{G} = (\mathbf{H}\mathbf{H}^T)$ est la matrice de Gram (parfois appelée *matrice d'information*). Dans les cas sous-déterminés réalistes cette matrice est inversible ($p \ll n$). La variation de cet écart vaut

$$\delta \left\{ \text{Tr} (\mathbf{A} - \mathbf{I}) (\mathbf{A} - \mathbf{I})^T \right\} = 2\text{Tr} \left\{ \delta\mathbf{\Omega} (\mathbf{G}\mathbf{\Omega}^T - \mathbf{H}) \right\}. \quad (8.59)$$

Il en découle qu'une matrice $\mathbf{\Omega}$ assurant l'optimalité est de la forme

$$\mathbf{\Omega}^* = \mathbf{H}^T\mathbf{G}^{-1} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}. \quad (8.60)$$

Ce qui conduit à l'estimateur $\mathbf{x}^* = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\mathbf{y}$. On retrouve là l'inverse généralisé de Moore-Penrose. Autrement dit la méthode de Backus-Gilbert conduit essentiellement au même résultat que la méthode variationnelle (ou BLUE), pourvu que $\mathbf{B}^{-1} = 0$ et que $\mathbf{R} = \mathbf{I}$.

Dans la littérature géophysique où la technique de Backus-Gilbert est utilisée, l'écart à minimiser est légèrement différent et conduit à des résultats formellement distincts. Toutefois la nature des résultats est inchangée.

8.5.2 Méthode du maximum d'entropie

Le principe du maximum d'entropie est une méthode essentiellement Bayésienne. Comme la méthode de Backus-Gilbert, cette méthode est bien adaptée à des problèmes sous-déterminés, ou très mal conditionnés.

On cherche toujours à résoudre $\mathbf{y} = \mathbf{H}\mathbf{x}$. Ici, on souhaite minimiser l'information gagnée, autre que celle apportée par les données \mathbf{y} . Toute autre information introduit en effet un biais d'origine non-contrôlée dans l'inversion. Or on sait quantifier la quantité d'information, grâce au point de vue de Shannon, en utilisant la fonction de Kullback-Leibler. La densité de probabilité de départ est celle de l'ébauche, soit $p(\mathbf{x})$. On cherche à obtenir la densité a posteriori $p(\mathbf{x}|\mathbf{y})$.

Plutôt que de minimiser cette quantité d'information (ou maximiser l'entropie si l'on préfère) sous la contrainte que $\mathbf{y} = \mathbf{H}\mathbf{x}$ soit respecté, on assouplit cette contrainte en demandant qu'elle soit vérifiée en moyenne : $\mathbf{y} = \mathbb{E}_{p(\mathbf{x}|\mathbf{y})}(\mathbf{H}\mathbf{x})$. En introduisant un vecteur de multiplicateurs de Lagrange β pour imposer la contrainte, la fonction coût à optimiser est alors

$$\mathcal{L} = \sum_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) \ln \frac{p(\mathbf{x}|\mathbf{y})}{p(\mathbf{x})} + \beta^T \sum_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) (\mathbf{y} - \mathbf{H}\mathbf{x}) . \quad (8.61)$$

La solution est plus complexe que dans le cas variationnel où la fonction coût est quadratique. Cette fonctionnelle (une fonction définie sur un ensemble de fonctions), est à optimiser sur $p(\mathbf{x}|\mathbf{y})$ et sur β . Sous la contrainte additionnelle que $\sum_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) = 1$, on obtient en optimisant sur $p(\mathbf{x}|\mathbf{y})$:

$$p(\mathbf{x}|\mathbf{y}) = \frac{p(\mathbf{x}) \exp(\beta^T \mathbf{H}\mathbf{x})}{\sum_{\mathbf{x}} p(\mathbf{x}) \exp(\beta^T \mathbf{H}\mathbf{x})} . \quad (8.62)$$

Cette densité sera complètement déterminée lorsque le vecteur β sera identifié. Injectant cette expression de $p(\mathbf{x}|\mathbf{y})$ dans la fonctionnelle \mathcal{L} , on est conduit à minimiser la fonction de coût

$$\Psi = \ln Z(\beta) - \beta^T \mathbf{y} \quad \text{avec} \quad Z(\beta) = \sum_{\mathbf{x}} p(\mathbf{x}) \exp(\beta^T \mathbf{H}\mathbf{x}) . \quad (8.63)$$

$Z(\beta)$ est appelé *fonction de partition* par analogie avec la physique statistique.

Une fois connue $p(\mathbf{x}|\mathbf{y})$, un estimateur est alors nécessaire pour proposer une solution objective au problème d'inversion. L'estimateur moyen

$$\bar{\mathbf{x}} = \sum_{\mathbf{x}} \mathbf{x} p(\mathbf{x}|\mathbf{y}) , \quad (8.64)$$

possède la propriété désirée de satisfaire $\mathbf{y} = \mathbf{H}\bar{\mathbf{x}}$.

Dans le cas où la densité de probabilité a priori $p(\mathbf{x})$ est gaussienne, alors la fonction coût n'est autre que la fonction coût quadratique, mais exprimée dans l'espace des observations, c'est-à-dire en suivant le formalisme PSAS.

En effet supposons $p(\mathbf{x})$ de la forme de Eq.(8.49), alors la fonction de partition se calcule simplement :

$$\begin{aligned} Z(\beta) &= \int_{-\infty}^{\infty} d^n \mathbf{x} p(\mathbf{x}) \exp(\beta^T \mathbf{H}\mathbf{x}) \\ &= \int_{-\infty}^{\infty} d^n \mathbf{x} \frac{1}{(2\pi)^{n/2} |\mathbf{B}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu)^T \mathbf{B}^{-1}(\mathbf{x} - \mu) + \beta^T \mathbf{H}(\mathbf{x} - \mu) + \beta^T \mathbf{H}\mu\right) \\ &= \frac{(2\pi)^{n/2}}{|\mathbf{B}^{-1}|^{1/2}} \frac{1}{(2\pi)^{n/2} |\mathbf{B}|^{1/2}} \exp\left(\frac{1}{2} \beta^T \mathbf{H} \mathbf{B} \mathbf{H}^T \beta + \beta^T \mathbf{H}\mu\right) . \end{aligned} \quad (8.65)$$

À l'aide de cette intégration gaussienne, on a donc

$$Z(\beta) = \exp\left(\frac{1}{2}\beta^T \mathbf{H} \mathbf{B} \mathbf{H}^T \beta + \beta^T \mathbf{H} \mu\right). \quad (8.66)$$

Il en résulte la fonction coût effective :

$$\Psi = \frac{1}{2}\beta^T \mathbf{H} \mathbf{B} \mathbf{H}^T \beta - \beta^T (\mathbf{y} - \mathbf{H} \mu). \quad (8.67)$$

Cette fonction coût est à comparer à la fonction coût Eq.(1.44) du chapitre 1.

Bibliographie

- [1] P. BOUTTIER, F. ET COURTIER. Data assimilation, concepts and methods. *Training course notes of ECMWF, European Centre for Medium-range Weather Forecasts, Reading, UK.*, **53**, 1997.
- [2] G. BURGERS, P.J. VAN LEEUWEN, et G. EVENSEN. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126**, 1719–1724, 1998.
- [3] S. COHN. An Introduction to Estimation Theory. *Journal of the Meteorological Society of Japan*, pages 257–288, 1997.
- [4] R. DALEY. *Atmospheric Data Analysis*. Cambridge University Press, 1993.
- [5] I. G. ENTING. *Inverse Problems in Atmospheric Constituent Transport*. Cambridge, Atmospheric and Space Science Series, 2002.
- [6] G. EVENSEN. The Ensemble Kalman Filter : Theoretical Formulation and Practical Implementation. *Ocean Dynamics*, **53**, 343–367, 2003.
- [7] N. J. GORDON, D. J. SALMOND, et A. F. M. SMITH. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE-F*, **140**, 107–113, 1993.
- [8] A. GRIEWANK. *Evaluating derivatives : Principles and Techniques of Automatic Differentiation*. SIAM, 2000.
- [9] C. W. GROETSCH. *Inverse Problems in the Mathematical Sciences*. Vieweg and Verlag, 1993.
- [10] K. IDE, P. COURTIER, M. GHIL, et A. LORENC. Unified notation for data assimilation : Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189, 1999.
- [11] E. KALNAY. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 2003.
- [12] F. LE DIMET, H.-E. NGODOCK, B. LUONG, et J. VERRON. Sensitivity analysis in variational data assimilation. *J. Met. Soc. Japan*, **75**, 245–255, 1997.
- [13] E.N. LORENZ et K.E. EMMANUEL. Optimal sites for supplementary weather observations : simulation with a small model. *J. Atmos. Sci.*, **55**, 399–414, 1998.
- [14] D. T PHAM. Stochastic Methods for Sequential Data Assimilation in Strongly Non-linear Systems. *Monthly Weather Review*, pages 1194–1207, 2001.
- [15] D. T. PHAM, J. VERRON, et M. C. ROUBAUD. A Singular Evolutive Extended Kalman Filter for Data Assimilation in Oceanography. *J. Marine Systems*, pages 323–340, 1998.

-
- [16] C. D. RODGERS. *Inverse Methods for atmospheric sounding*. World Scientific, Series on Atmospheric, Oceanic and Planetary Physics, 2000.
- [17] A. J. SEGERS. *Data assimilation in atmospheric chemistry models using Kalman filtering*. Thèse de doctorat, Delft University, Netherlands, 2002.
- [18] B. SPORTISSE. *Pollution atmosphérique : Des processus à la modélisation*, chapter 2. Springer, 2007.
- [19] O. TALAGRAND. Assimilation of Observations, an Introduction. *Journal of the Meteorological Society of Japan*, **75**, 191–209, 1997.
- [20] A. TARANTOLA. *Inverse Problem Theory*. Elsevier, 1987.
- [21] A. TARANTOLA. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.
- [22] A. TARANTOLA et B. VALETTE. Generalized nonlinear inverse problems solved using the least square criterion. *Reviews of Geophysics and Space Physics*, **20**, 219–232, 1982.
- [23] R. TODLING. Estimation theory and Foundations of Atmospheric Data Assimilation. *Data Assimilation Office Note, Goddard Laboratory for Atmospheres*, 1999.
- [24] P. J. VAN LEEUWEN. A variance-minimizing filter for large-scale applications. *Mon. Wea. Rev.*, **131**, 2071–2084, 2003.
- [25] C. R. VOGEL. *Computational Methods for Inverse Problems*. SIAM, Frontiers in Applied Mathematics, 2002.
- [26] A. E. WAN et R. VAN DER MERWE. *The Unscented Kalman Filter*, chapter 7. Kalman Filtering and Neural Networks, Wiley Publishing, 2001.