Special Mathematics Lecture

Introduction to Data Assimilation

Nagoya University, Spring 2022

Lecturer: Serge Richard

Goals of these Lectures notes:

Provide the necessary background information for understanding the main ideas and concepts of data assimilation
These notes correspond to 15 lectures lasting 90 minutes each.

These lecture notes have been typed and edited by
Vic Austen
Qiwen Sun
Serge Richard

Website for this course:
http://www.math.nagoya-u.ac.jp/~richard/SMLspring2022.html

Comments or corrections are welcome:
richard@math.nagoya-u.ac.jp

Contents

About these lecture notes:
These notes and the corresponding course have been mainly inspired by the book [5], with additional material borrowed from [1] and [4]. Other references about data assimilation consulted for this course are [2] and [7]. The probability part is based [3].

Chapter 1

The Basics

In this chapter, we provide some basic ideas about the framework in which data assimilation can be used.

1.1 Background Information

Definition 1.1.1 (From Wikipedia)**.** *Data assimilation is a mathematical discipline that seeks to optimally combine theory with observations.*

The main applications of data assimilation (DA) include:

1) The forecast of the evolution of a dynamical system,

2) The determination of the current state of a partially observed system,

3) The determination of an initial condition for the evolution of a system,

4) The determination of the value of some unknown parameters.

As mentioned in the above definition, data assimilation is a *mathematical discipline*, but was mainly developed by meteorologists and earth scientists. Now, its usefulness has been recognized and applied to several research fields. Let us also emphasize that data assimilation always requires two ingredients: a model and some data. A typical flowchart for any data assimilation process is given in Figure 1.1. Alternatively, a typical

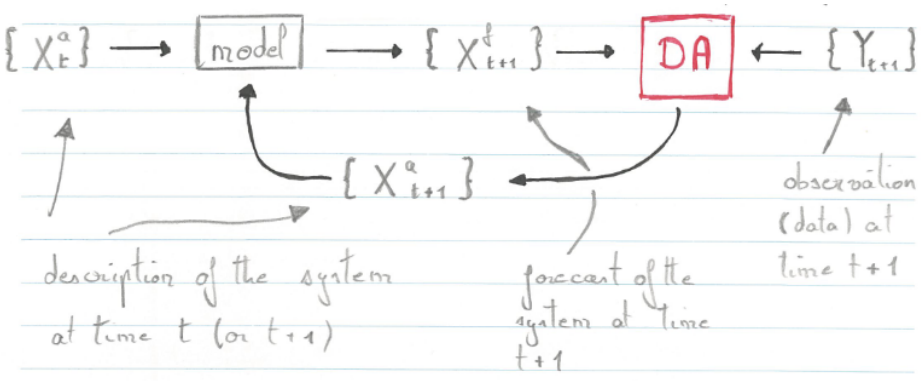

Figure 1.1: A typical flowchart of a data assimilation process.

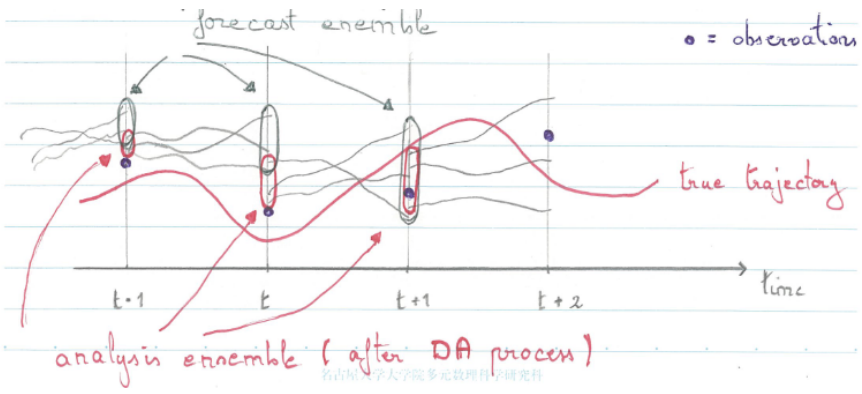

Figure 1.2: A typical representation of a data assimilation process with three simulations.

representation of a data assimilation process with three simulations is represented in Figure 1.2. The above definition also contains the idea of *optimality*, but this is subject to many constraints (as developed later): the model error (since a model is always a simplification), the precision error (no system is perfectly known or described), the partial observations (in space and time), the computational errors, the simplifications due to high dimensionality, the mathematical restriction (no exact solution known), *etc*. Note also that data assimilation consists of many techniques and approaches, a diagram showing these developments is given on page 12 of [1].

The goal of this course is to develop the mathematical tools for fully understanding Figures 1.1 and 1.2, and to understand some of the data assimilation methods. In addition, we shall describe some of the most familiar data assimilation algorithms. Note that we shall always work in the discrete time setting. This assumption simplifies part of the mathematics, and is sufficient for most of the applications. In addition, even for continuous time systems, a discretization in time is usually applied during the computation process.

Let us finally stress that we will not implement these tools in any programming language (*e.g.* Python). This is certainly unfortunate, but it would require too much additional time. On the other hand, students are encouraged to work on the applications and on the implementations by themselves, and to look at many examples available on internet.

1.2 Synthetic Data and Tests

We start by mentioning synthetic data since they are often used for testing any data assimilation algorithm. We sketch how they are generated and how they can be used in some elementary data assimilation techniques. This section is slightly cumbersome since we have not introduced the mathematical tools yet: they will be introduced in Chapter 2.

Consider a (discrete time) physical system whose evolution is described by a function $f : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^N$, and a function $x : \mathbb{Z} \to \mathbb{R}^N$ satisfying the equation

$$x(n+1) = f(x(n), n). \tag{1.2.1}$$

Here $x(n)$ represents the state of the physical system at time n, and the space $\mathbb{R}^N$ is called the *state space*. For example, it can be the position space, or the position $\times$ velocity space, or a much more complicated space, as we shall see in some examples later. Observe that the evolution of the system is provided by the function f

in (1.2.1), which gives the state of the system at time $n+1$ as a function of the state of the system at time n. Note that the function f can also depend explicitly on the time n, as emphasized by the second variable of f.

Let us now emphasize that the equation (1.2.1) is already a model, an *idealization* of the reality, also called *the truth*. Thus, we assume that the true system $x^t : \mathbb{Z} \to \mathbb{R}^N$ has an evolution given by:

$$x^t(n+1) = f(x^t(n), n) + g(n) \tag{1.2.2}$$

where $g : \mathbb{Z} \to \mathbb{R}^N$ is a function which is never known. This means that x^t can also never be known exactly. Since g is not known, this approach can not be used for generating synthetic data. We shall thus introduce an equivalent trick.

For the creation of synthetic data, we usually assume that the truth $x^t : \mathbb{Z} \to \mathbb{R}^N$ satisfies the equation (1.2.1), but that the model described by $x : \mathbb{Z} \to \mathbb{R}^N$ satisfies the following equation:

$$x(n+1) = f(x(n), n) + \xi(n), \tag{1.2.3}$$

where $\xi(n)$ is called *the model error* at time n, and this model error is usually not known explicitly. In this setting, the evolution of the truth follows the perfect system described by (1.2.1), while the evolution of the model is described by (1.2.3) and involves an error related to the simplification of our model compared to the truth.

Now even if the state of the true system at time n is supposed to be fully described by $x^t(n)$, it is usually not possible to observe it in its entirety. The observations are possible only on part of the system, or through a function defined on the system. For that reason, we define *the observation* $y(n)$ of the physical system at time n by

$$y(n) := H(x^t(n)) + \varepsilon(n), \tag{1.2.4}$$

where $H : \mathbb{R}^N \to \mathbb{R}^M$ corresponds to what is measured on the system, and $\varepsilon(n)$ is the *observation error*, namely the error in the measurement process. The function H is called *the observation operator*, or *the observation matrix* in the simplest situation where H is a $M \times N$ matrix. Note that usually, $M \leq N$, but this is not required. In this setting, the set of observations $\{y(n)\}_{n\in\mathbb{Z}}$ corresponds to the *synthetic data*.

Based on the notations introduced above, we can now describe more precisely one of the goals of data assimilation. As a time reference, we assume that the present time is denoted by n_*. Let us assume that $\{y(n)\}_{n\leq n_*}$ are known. These observations can be either real observations, or synthetic observations as constructed above. Then, one of the aims of data assimilation is to determine the observation forecasts $\{y^f(n)\}$ for $n > n_*$, or the more challenging aim, to determine the system forecasts $\{x^f(n)\}$ for $n > n_*$. Uncertainties have to be attached to these forecasts.

For testing the forecasts $\{y^f(n)\}_{n>n_*}$ one can usually only compare them with the set $\{y(n)\}_{n>n_*}$ when these observations are available (which means in the future, if n_* is the current time). However, in the present setting with synthetic data, we can compare them with $\{H(x^t(n))\}_{n>n_*}$, or compare $\left\{x^f(n)\right\}_{n>n_*}$ with $\{x^t(n)\}_{n>n_*}$, since these quantities exist. One simple way to measure the error is with the *Root Mean Square Error*, abbreviated RMSE.

The RMSE for a window of length p can be found by evaluating any one of the three expressions:

$$\sqrt{\sum_{i=1}^{p} \frac{1}{p} \left|y^f(n_*+i) - y(n_*+i)\right|^2} \tag{1.2.5}$$

or

$$\sqrt{\sum_{i=1}^{p} \frac{1}{p} \left| y^f(n_* + i) - H(x^t(n_* + i)) \right|^2} \tag{1.2.6}$$

or

$$\sqrt{\sum_{i=1}^{p} \frac{1}{p} \left| x^f(n_* + i) - x^t(n_* + i) \right|^2} \tag{1.2.7}$$

Let us emphasize that (1.2.5) is always possible, if we wait until the time $n_* + p$, while (1.2.6) and (1.2.7) are usually unavailable. They are available precisely in the framework of synthetic data, and this is the interest of this framework. In fact, by comparing (1.2.6) with the expression

$$\sqrt{\sum_{i=1}^{p} \frac{1}{p} \left| y(n_* + i) - H(x^t(n_* + i)) \right|^2} \tag{1.2.8}$$

we can conclude that the forecast process has been efficient if (1.2.6) is smaller than (1.2.8), while the prediction scheme is not efficient if (1.2.6) is bigger than (1.2.8).

As a final remark, observe that having fixed n_* to the present time is arbitrary, n_* can also be fixed at any particular time point in the past. This does not change any outcome, except that $y(n)$ for some $n > n_*$ might already be available for comparisons.

1.3 Data Driven Forecasting

In this section, we mention the simplest forecasting methods, and give a flavour of future developments. Clearly, everything is very simple, since no mathematical tools have been introduced yet.

Let us start with *polynomial interpolation*. We consider a family $\{t_i\}_{i=0}^{\ell} \subset \mathbb{R}$, $t_i \neq t_j$ for $i \neq j$, and set $c_i : \mathbb{R} \to \mathbb{R}$ with

$$c_i(t) := \frac{\prod_{j \neq i}(t - t_j)}{\prod_{j \neq i}(t_i - t_j)} = \frac{(t - t_0)(t - t_1) \dots (t - t_{i-1})(t - t_{i+1}) \dots (t - t_\ell)}{(t_i - t_0)(t_i - t_1) \dots (t_i - t_{i-1})(t_i - t_{i+1}) \dots (t_i - t_\ell)}.$$

Clearly, $c_i(t_i) = 1$ but $c_i(t_j) = 0$ if $i \neq j$.

If $\{y(t_0), \dots, y(t_\ell)\} \subset \mathbb{R}^M$ are observations at time $t_0, \dots, t_\ell$, then one can set $q : \mathbb{R} \to \mathbb{R}^M$ as:

$$q(t) := \sum_{i=0}^{\ell} c_i(t)\, y(t_i) \tag{1.3.1}$$

Thus, q is a function of the variable t with values in $\mathbb{R}^M$. Observe that the graph of this function goes through the points $(t_i, y(t_i))$, since $q(t_i) = y(t_i)$, for any $i \in \{0, 1, \dots, \ell\}$. Then, for any any $t \notin \{t_0, \dots, t_\ell\}$ we can set the observation forecast $y^f(t) := q(t)$. Clearly, this approach can be applied to the discrete time example of Section 1.2, but it performs poorly in general.

Exercise 1.3.1 (Linear interpolation)**.** *Write the expression for $y^f(t)$ if the two observations $y(t_0)$ at time t_0 and $y(t_1)$ at time t_1 are provided. Prove that the corresponding polynomial represents a line in $\mathbb{R}^M$.*

The next forecast scheme consists in *a linear combination with training set*. The idea is still to use a linear combination of the past $\ell + 1$ observations, as in (1.3.1), but without imposing the explicit coefficients $c_i(t)$.

For simplicity, we shall assume that $M = 1$ and that $t_{i+1} - t_i = \text{cst}$ for all i, meaning that the observations are taken at regular time intervals. We then define the *residual* r_j as the difference between observation and forecast, namely for $j \in \{1, \dots, J\}$

$$r_j := y\left(t_{j+\ell+1}\right) - y^f(t_{j+\ell+1}) = y(t_{j+\ell+1}) - \sum_{i=0}^{\ell} x_i y(t_{j+\ell-i}). \tag{1.3.2}$$

Clearly, the sum term is a linear combination of the $\ell + 1$ previous observations, but the coefficients x_i are not fixed yet.

Suppose that we have $J + \ell + 1$ data available, then we can write in a matrix form:

$$r := \begin{pmatrix} r_1 \\ \vdots \\ r_J \end{pmatrix} = \begin{pmatrix} y(t_{1+\ell+1}) \\ \vdots \\ y(t_{J+\ell+1}) \end{pmatrix} - \mathcal{A} \begin{pmatrix} x_0 \\ \vdots \\ x_\ell \end{pmatrix} \in \mathbb{R}^J \quad \text{with} \quad \mathcal{A} = \begin{pmatrix} y(t_{\ell+1}) & y(t_\ell) & \dots & y(t_1) \\ y(t_{\ell+2}) & y(t_{\ell+1}) & \dots & y(t_2) \\ \vdots & \vdots & & \vdots \\ y(t_{\ell+J}) & y(t_{\ell+J-1}) & \dots & y(t_J) \end{pmatrix} \in M_{J\times\ell+1}(\mathbb{R})$$

Let us emphasize that the above construction is based on the assumption of *the stationarity of the time series*, in the sense that the coefficients $x_0, \dots, x_\ell$ can be used at any time! The regularity of the time between the observations is necessary for this assumption.

We write the matrix equation above as $r = y - \mathcal{A}x$, where x is still unknown. If $J = \ell + 1$ and if $\mathcal{A}$ is invertible, we can set $x := \mathcal{A}^{-1}y$, and then $r = \mathbf{0} \in \mathbb{R}^J$, but often $J \gg \ell + 1$, and then there is in general no $x \in \mathbb{R}^{\ell+1}$ such that $r = \mathbf{0}$.

The *method of least squares* consists of finding x such that $\|r\|^2 := \sum_{j=1}^{J} r_j^2$ is minimum. Note that such an x might not be unique, but at least there exists one. Thus, we denote these value(s) of x as x_*, and write

$$x_* := \arg\min \|r\|^2 = \arg\min_{x\in\mathbb{R}^{\ell+1}} \|\mathcal{A}x - y\|^2. \tag{1.3.3}$$

Once one x_* are found (see exercise below), we can use (1.3.2) and obtain:

$$y^f\left(t_{n_*+1}\right) = \sum_{i=0}^{\ell} (x_*)_i y\left(t_{n_*-i}\right).$$

In the sequel, we shall often consider square matrices with a special property: A matrix $\mathcal{A} \in M_{N\times N}(\mathbb{R})$ is *positive definite* if $\mathcal{A}$ is symmetric and $x^T\mathcal{A}x > 0$ for all $x \in \mathbb{R}^N \setminus \{\mathbf{0}\}$. We write simply $\mathcal{A} > 0$ if $\mathcal{A}$ is positive definite. Note that since $\mathcal{A}$ is symmetric, its eigenvalues are well-defined, and this condition means that its N eigenvalues are strictly positive.

Exercise 1.3.2. *Consider $F : \mathbb{R}^{\ell+1} \to \mathbb{R}$, with $F(x) = \|\mathcal{A}x - y\|^2 = \langle \mathcal{A}x - y, \mathcal{A}x - y\rangle$, where the last notation denotes the scalar product in $\mathbb{R}^J$. Show that $\nabla F(x) = 2\mathcal{A}^T(\mathcal{A}x - y)$, and that the Hessian Matrix of F is $\mathcal{H}_F = 2\mathcal{A}^T\mathcal{A} \in M_{\ell+1\times\ell+1}(\mathbb{R})$. Show that A^TA is positive definite if* $\operatorname{rank}\mathcal{A} = \ell + 1$. *If so, show that the solution of*

$$\mathcal{A}^T\mathcal{A}x = \mathcal{A}^T y$$

is the unique solution to the least squares equation (1.3.3). This implies that x_* in the equation is unique.

Let us finally remark that the two approaches introduced above do not rely on any model for the dynamical system generating the observations, they are part of the so-called *autoregressive models* based on data only. In other words, there are instances of purely *data driven models*. More sophisticated models of this type exist, but they use random variables, introduced in Section 2.1, as for example the ARIMA model.

1.4 Model Driven Forecasting

We come back to the construction of Section 1.2, and assume that the evolution of the model is described by

$$x(n+1) = f(x(n))$$

with $f : \mathbb{R}^N \to \mathbb{R}^N$. This is a simplification of (1.2.1) once the system is assumed to be autonomous.

Recall also that true evolution x^t is supposed to satisfy $x^t(n+1) = f(x^t(n)) + g(n)$ but g is not known, see (1.2.2). Suppose also that an initial condition $x(n_*)$ is given. Then, for generating the forecast at $n_* + k$ one has

$$x(n_* + k) = f(x(n_* + k - 1)) = \cdots = f\Big(f(\dots f(x(n_*))\dots)\Big) = f^{\circ k}(x(n_*)), \tag{1.4.1}$$

where $f^{\circ k}$ means "f composed k times". On the other hand, for the truth one has

$$x^t(n_*+k) = f\left(x^t(n_* + k - 1)\right) + g(n_*+k-1) = f\Big(f(x^t(n_*+k-2)) + g(n_*+k-2)\Big) + g(n_*+k-1) = \dots \tag{1.4.2}$$

Observe that a new term $g(x_* + \ell)$ enters at each step. Due to these additional terms, and in particular if f is non-linear, the difference between $x(n_* + k)$ and $x^t(n_* + k)$ will quickly grow, and the same for the difference between

$$y(n_* + k) = H(x^t(n_* + k)) + \varepsilon(n_* + k) \quad \text{and} \quad y^f(n_* + k) = H(x(n_* + k)), \tag{1.4.3}$$

where H denotes the observation operator, and ε the observation error, as introduced in (1.2.4). In addition, this construction depends on the initial value $x(n_*)$, which is not precisely known in general. If we define *the residual* at time $n_* + k$ by

$$r(n_* + k) := y(n_* + k) - y^f(n_* + k)$$

then one observes with (1.4.1), (1.4.2), and (1.4.3), that this residual is highly dependent on $x(n_*)$, with a very complicated dependence on this initial condition.

Let us now present how past observations can improve our knowledge of $x(n_*)$, and therefore improve the forecast $x(n_* + k)$. The approach is again based on the method of least squares. For simplicity we assume that $M = 1$ and that the observation operator $H : \mathbb{R}^N \to \mathbb{R}$ is linear, *i.e.* $H(x) = Hx \in \mathbb{R}$ for some $H \in M_{1\times N}(\mathbb{R})$.

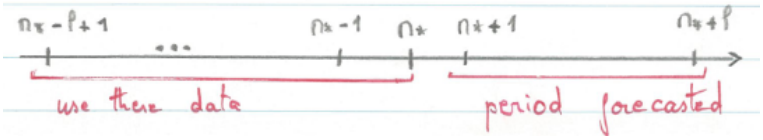

Figure 1.3: A window with ℓ past observations, and a window with ℓ forecast values.

We assume that ℓ past observations are available at regular time intervals, and denote them by $\{y(n_*-\ell+i)\}_{i=1}^{\ell} \subset \mathbb{R}$. Set $L : \mathbb{R}^N \to \mathbb{R}$ defined by:

$$\begin{aligned} L(x) &:= \frac{1}{2}\sum_{i=1}^{\ell} \left(r(n_* - \ell + i)\right)^2 = \frac{1}{2}\sum_{i=1}^{\ell} \Big(y^f(n_* - \ell + 1) - y(n_* - \ell + 1)\Big)^2 \\ &= \frac{1}{2}\sum_{i=1}^{\ell} \Big(Hx(n_* - \ell + 1) - y(n_* - \ell + i)\Big)^2 = \frac{1}{2}\sum_{i=1}^{\ell} \Big(Hf^{\circ i}(x) - y(n_* - \ell + i)\Big)^2 . \end{aligned}$$

In the last step, we set : $x = x(n_* - \ell)$. A representation is given in Figure 1.3. Hence, we shall look for an initial condition $x \equiv x(n_* - \ell)$ which *minimizes* L, and then use $x(n_*)$ as the initial condition for forecasting $x(n_* + 1), \dots, x(n_* + \ell)$. This operation is repeated every ℓ intervals of time, often called *a window of length* ℓ.

Let us emphasize that the minimization problem is much more involved than the one seen in Section 1.3, since the x-dependence is much more complicated. A local extremum x_* of L satisfies $\nabla L(x_*) = \mathbf{0}$, but even if we find one it is necessary to check that it is indeed a local minimum and not a local maximum. We set $H = (h_1, \dots, h_N)$ and observe firstly that

$$\begin{aligned}\nabla\left[\frac{1}{2}\left(Hf^{\circ i}(x) - y(n_* - \ell + i)\right)^2\right] &= \left(Hf^{\circ i}(x) - y(n_* - \ell + i)\right)\nabla Hf^{\circ i}(x)\\ &= -r(n_* - \ell + i)\nabla\sum_{j=1}^{N} h_j\left(f^{\circ i}(x)\right)_j\\ &= -r(n_* - \ell + i)\left((J_{f^{\circ i}})(x)\right)^T H^T.\end{aligned}$$

For the last equality, we recall that for a continuously differentiable function $k : \mathbb{R}^N \to \mathbb{R}^N$, the Jacobian matrix J_k is given by

$$J_k(x) = \begin{pmatrix} \frac{\partial k_1}{\partial x_1} & \cdots & \frac{\partial k_1}{\partial x_N} \\ \vdots & & \vdots \\ \frac{\partial k_N}{\partial x_1} & \cdots & \frac{\partial k_N}{\partial x_N} \end{pmatrix}(x).$$

Exercise 1.4.1. *Check and justify the above computations.*

Thus, by collecting the previous equalities we infer that

$$\nabla L(x) = -\sum_{i=1}^{\ell} r(n_* - ell + 1)\left((J_{f^{\circ i}})(x)\right)^T H^T$$

Let us emphasize that the expression $\nabla L(x)$ is very complicated but the orange part does not depend on any data, it depends only on the model and can be computed once for all. The cyan part depends on the data and has to be updated every *variational data assimilation cycle* of length ℓ.

There exist algorithms for finding x_* such that $\nabla L(x_*) \approx 0$, such as the *gradient descent method*. Once an optimal x_* is chosen, we set $x^a(n_*) := f^{\circ\ell}(x_*)$, where a stands to *analysis*, and define the forecasts $y^f(n_*+1) := f(x^a(n_*))$, up to $y^f(n_* + \ell) := f^{\circ\ell}(x^a(n_*))$. Usually, these forecasts are quite good, meaning that they are quite close to the truth. When $x^f(n_* + \ell) := f^{\circ\ell}(x^a(n_*))$ is forecast, we add the superscript f for forecast. Later, when the observations $y(n_* + 1), \dots, y(n_* + \ell)$ are available, and when the method described above is applied on the window $n_* + 1, \dots, n_* + \ell$, one gets an improved initial condition for $x^f(n_* + \ell)$ for the next cycle, and denotes it by $x^a(n_* + \ell)$. Figure 1.4 corresponds to this scheme with $\ell = 5$.

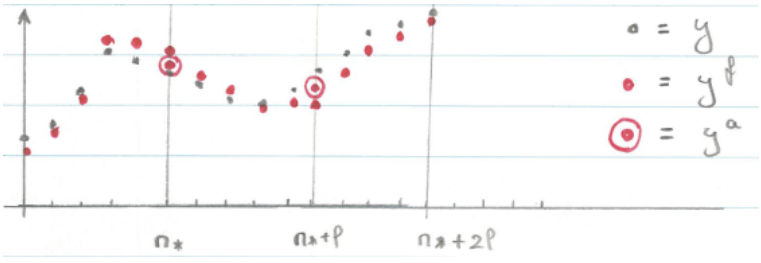

Figure 1.4: An illustration of the model driven forecasting scheme.

Remark 1.4.2. *1) If $M > 1$, the same approach holds, but one instead has to sum over the components of*

$$\left(y^f(n_* - \ell + i) - y(n_* - \ell + i)\right).$$

2) If H is non-linear, the same approach is possible, the same (cyan and orange) separation holds, but the orange part is slightly more complicated.

3) What is missing are confidence intervals, or information about how much we can trust a forecast. This notion is related to probability, which are going to be discussed in Section 2.1.

Chapter 2

Mathematical Background

This chapter contains the mathematical background necessary for the understanding of data assimilation techniques.

2.1 Probability and Random Variables

The aim of this section is to describe and quantify any non-predictable experiment. We give a framework suitable for many applications.

Definition 2.1.1 (Measurable space)**.** *A* measurable space $(\Lambda, \mathcal{E})$ *is a set* Λ *together with a collection of subsets* $\mathcal{E}$ *closed under complement, countable unions and countable intersections: if* $A \in \mathcal{E}$, $A^c := \Lambda \backslash A \in \mathcal{E}$, *if* $\{A_j\}_{j\in\mathbb{N}} \subset \mathcal{E}$, *then* $\cup_j A_j \in \mathcal{E}$ *and* $\cap_j A_j \in \mathcal{E}$. *One also says that* $\mathcal{E}$ *is a* σ-algebra.

If Λ also belongs to $\mathcal{E}$, then one condition can be dropped, as one can show in this exercise:

Exercise 2.1.2. *Prove this statement: if* $\mathcal{E}$ *is a collection of subsets of* Λ *with* $\Lambda \in \mathcal{E}$ *and which is closed under complement and countable unions, then it is closed under countable intersections.*

An example of a measurable space is the usual space $\mathbb{R}^N$ together with the family of sets generated by intervals by considering countable unions, intersections, and complements. In this case, one speaks about the Borel σ-algebra σ_{B}. Thus, $(\mathbb{R}^N, \sigma_{\mathrm{B}})$ is the most common measurable space, and one usually denotes it simply by $\mathbb{R}^N$. An other example of a measurable space is provided by $\Lambda = \{\lambda_1, \ldots, \lambda_N\}$ a finite set and $\mathcal{E}$ the *power set* of Λ consisting of all subsets of Λ. Two standard examples are

$$\Lambda = \{\text{heads, tails}\} \quad \text{or} \quad \Lambda = \left\{ \boxdot, \boxdot, \boxdot, \boxdot, \boxdot, \boxdot \right\}.$$

This second example can also be extended to an infinite set, like for example $\Lambda = \mathbb{N}$ or $\Lambda = \mathbb{Z}$, also endowed with their respective power set.

Definition 2.1.3 (Probability space)**.** *A* probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *consists of a measurable space* $(\Omega, \mathcal{F})$ *with* $\Omega \in \mathcal{F}$, *and a function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *satisfying* $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$ *and*

$$\mathbb{P}\left(\cup_{j\in\mathbb{N}} A_j\right) = \sum_{j\in\mathbb{N}} \mathbb{P}(A_j)$$

whenever $A_j \cap A_k = \emptyset \ \forall j \neq k$. *We call* Ω the sample space, $\mathcal{F}$ the event space, $\omega \in \Omega$ an elementary event *and* $A \in \mathcal{F}$ an event, *and finally* $\mathbb{P}$ the probability measure.

Usually, Ω is very complicated or unknown. Functions defined on Ω are more important.

Definition 2.1.4 (Random variable)**.** *Consider a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and a measurable space* $(\Lambda, \mathcal{E})$*. A* random variable $X : \Omega \to \Lambda$ *is a function satisfying for any* $A \in \mathcal{E}$

$$\{\omega \in \Omega \mid X(\omega) \in A\} \equiv X^{-1}(A) \in \mathcal{F}. \tag{2.1.1}$$

In the special case $\Lambda = \mathbb{R}^N$, or more precisely if we consider $(\Lambda, \mathcal{E}) = (\mathbb{R}^N, \sigma_{\mathrm{B}})$, then (2.1.1) is satisfied if $X = (X_1, \ldots, X_N)$ verifies

$$\{\omega \in \Omega \mid X_j(\omega) \leq x_j \ \forall j = 1, \ldots, N\} \in \mathcal{F}$$

for any $(x_1, \ldots, x_N) \in \mathbb{R}^N$. Note that the special case $N = 1$ corresponds to the standard univariate random variables, while $N > 1$ corresponds to the multivariate random variables. Based on the notion of random variable, we can define a new measure:

Definition 2.1.5 (Induced probability measure)**.** *When* $X : \Omega \to \Lambda$ *is a random variable from a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *to a measurable space* $(\Lambda, \mathcal{E})$*, the map* $\mu_X : \mathcal{E} \to [0, 1]$ *defined by*

$$\mu_X(A) = \mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) \in A\}\big) = \mathbb{P}\left(X^{-1}(A)\right) \equiv \mathbb{P}(X \in A)$$

is called the induced probability measure*.* μ_X *is also called* the law *of* X*, and we write* $X \sim \mu_X$ *for this correspondence.*

Usually, μ_X is much simpler than X, since it is defined on $(\Lambda, \mathcal{E})$, as for example on $(\mathbb{R}^N, \sigma_{\mathrm{B}})$, and not on $(\Omega, \mathcal{F})$. Let us mention that there exist two principal types of random variables (but others also exist). Quite often, only these 2 types are mentioned.

Definition 2.1.6 (Absolutely continuous random variable)**.** *The random variable* $X : \Omega \to \mathbb{R}^N$ *is* absolutely continuous *if the induced probability measure is absolutely continuous with respect to the Lebesgue measure, namely if there exists a (measurable) function* $\Pi_X : \mathbb{R}^N \to [0, \infty)$ *satisfying for any* $A \in \sigma_{\mathrm{B}}$

$$\mu_X(A) = \int_A \Pi_X(x)\,\mathrm{d}x.$$

The function Π_X *is called* the probability density function*, or simply the* pdf.

Definition 2.1.7 (Discrete valued random variable)**.** *The random variable* $X : \Omega \to \Lambda$ *is* discrete valued *if* $X(\Omega) = \{X(\omega) \mid \omega \in \Omega\}$ *is finite or countable. In this case, we define the function* $p_X : X(\Omega) \to [0, 1]$ *by*

$$p_X(x) := \mathbb{P}\left(X^{-1}(\{x\})\right)$$

for any $x \in X(\Omega)$*. The function* p_X *is called the* the probability mass function*, or simply* pmf.

In these two situations, we still write $X \sim \Pi_X$ or $X \sim p_X$. It is clear that the following properties hold: $\int_{\mathbb{R}^N} \Pi_X(x)\,\mathrm{d}x = 1$ and $\sum_{x \in X(\Omega)} p_X(x) = 1$. Observe also that for any absolutely continuous random variable X, one has $\mu_X(x) = 0$ for any $x \in \mathbb{R}^N$ while $\Pi_X(x) \in [0, \infty)$ for (almost every) $x \in \mathbb{R}^N$.

Remark 2.1.8. *Any function* $\Pi : \mathbb{R}^N \to [0, \infty)$ *satisfying* $\int \Pi(x)\,\mathrm{d}x = 1$*, or any function* p *from a finite set or a countable set* Λ *to* $[0, 1]$ *satisfying* $\sum_x p(x) = 1$*, defines the pdf or the pmf of a random variable. However, in such a situation we don't have the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*, we just have the law. If necessary, a probability space can be constructed, but it is somewhat artificial.*

Let us still introduce a new measurable space $(\Xi, \mathcal{G})$. In applications, this space will always be either $\mathbb{R}$, or $\mathbb{R}^N$, or $M_{n\times m}(\mathbb{R})$, the set of $n \times m$ matrices with entries in $\mathbb{R}$. Since $M_{n\times m}(\mathbb{R})$ can be identified with $\mathbb{R}^{nm}$, it is also a measurable space. We then consider functions $f : \Lambda \to \Xi$. These functions are *measurable* if $f^{-1}(B) \in \mathcal{E}$ for all $B \in \mathcal{G}$. Note that this notion of measurability was already the one imposed on the random variable $X : \Omega \to \Lambda$.

Definition 2.1.9 (Expectation)**.** *Let $X : \Omega \to \Lambda$ be a random variable. For any measurable function $f : \Lambda \to \Xi$, the* expectation *of $f(X)$ is defined by (the Lebesgue type integral)*

$$\mathbb{E}(f(X)) = \int_\Lambda f(x)\, \mu_X(\mathrm{d}x). \tag{2.1.2}$$

Note that when writing such an expression, we assume that it exists even with f replaced by $|f|$ (absolute convergence of the integral). When Λ is $\mathbb{R}$, or $\mathbb{R}^N$, or $M_{n\times m}(\mathbb{R})$, we simply write $\mathbb{E}(X)$ for $\mathbb{E}(\mathrm{id}(X))$, where id *denotes the identity function satisfying* $\mathrm{id}(x) = x$*,*

For any $A \in \mathcal{E}$, we define the characteristic function $\mathbf{1}_A$ by $\mathbf{1}_A(x) = 1$ if $x \in A$ and $\mathbf{1}_A(x) = 0$ if $x \notin A$. Thus, $\mathbf{1}_A : \Lambda \to \mathbb{R}$ is a measurable function, and one observe that the following equalities hold:

$$\mathbb{P}(X \in A) \equiv \mathbb{P}\Big(\{\omega \in \Omega \mid X(\omega) \in A\}\Big) = \mu_X(A) = \int_A \mu_X(\mathrm{d}x) = \int_\Lambda \mathbf{1}_A(x)\, \mu_X(\mathrm{d}x) = \mathbb{E}(\mathbf{1}_A(X)).$$

Exercise 2.1.10. *For $\sigma > 0$ and $\bar{x} \in \mathbb{R}$ set $\Pi : \mathbb{R} \to [0, \infty)$ by*

$$\Pi(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}\,(x - \bar{x})^2\right).$$

Check that $\int \Pi(x)\, \mathrm{d}x = 1$. We write $X = N(\bar{x}, \sigma^2)$ for the corresponding univariate random variable, called Gaussian random variable*. Check that $\mathbb{E}(X) \equiv \mathbb{E}(\mathrm{id}(X)) = \bar{x}$. and* $\mathrm{Var}(X) := \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = \mathbb{E}\left((X - \bar{x})^2\right) = \sigma^2$.

More generally, for $\bar{x} \in \mathbb{R}^N$ and $P \in M_{N\times N}(\mathbb{R})$ with $P > 0$, set $\Pi : \mathbb{R}^N \to [0, \infty)$ with

$$\Pi(x) := \frac{1}{(2\pi)^{N/2}|P|^{1/2}} \exp\left(-\frac{1}{2}(x - \bar{x})^T P^{-1}(x - \bar{x})\right),$$

*with $|P| := \det(P)$. Check that $\int \Pi(x)\, \mathrm{d}x = 1$. We write $X = N(\bar{x}, P)$ for the corresponding multivariate random variable, called N-*dim Gaussian random variable*. Check that $\mathbb{E}(X) = \bar{x}$, and that $P = \mathbb{E}\left((X - \bar{x})(X - \bar{x})^T\right)$. Here, P is called the* covariance matrix*.*

Exercise 2.1.11. *If $X : \Omega \to \mathbb{R}^N$ is absolutely continuous with pdf Π_X and if $\phi : \mathbb{R}^N \to \mathbb{R}^N$ is bijective and C^∞, show that $Y := \phi(X) : \Omega \to \mathbb{R}^N$ is a new absolutely continuous random variable, with pdf Π_Y given by $\Pi_Y(y) = \Pi_X\left(\phi^{-1}(y)\right)|J_{\phi^{-1}}(y)|$. Here, $|J_{\phi^{-1}}(y)|$ denotes the determinant of the Jacobian matrix of ϕ^{-1}.*

Consider now two measurable spaces $(\Lambda^1, \mathcal{E}^1)$ and $(\Lambda^2, \mathcal{E}^2)$, and two random variables $X^1 : \Omega \to \Lambda^1$ and $X^2 : \Omega \to \Lambda^2$ defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The induced probability measures are denoted by μ_{X^1} and μ_{X^2}. Set $Z = (X^1, X^2) : \Omega \to \Lambda^1 \times \Lambda^2$ with $\Lambda^1 \times \Lambda^2 = \{(x^1, x^2) \mid x^1 \in \Lambda^1, x^2 \in \Lambda^2\}$. The set $\Lambda^1 \times \Lambda^2$ is endowed with the σ-algebra generated by boxes $A^1 \times A^2 = \{(x^1, x^2) \mid x^1 \in A^1 \text{ and } x^2 \in A^2\}$ for any $A^1 \in \mathcal{E}^1$ and $A^2 \in \mathcal{E}^2$. This σ-algebra is denoted by $\mathcal{E}^1 \times \mathcal{E}^2$. The induced probability measure μ_Z is called the *joint measure*. By definition, for any set $A \in \mathcal{E}^1 \times \mathcal{E}^2$, one has

$$\mathbb{P}\left(\Big\{\omega \in \Omega \mid (X^1(\omega), X^2(\omega)) \in A\Big\}\right) = \mu_Z(A) = \int_A \mu_Z(\mathrm{d}x^1 \times \mathrm{d}x^2) = \int_{\Lambda^1\times\Lambda^2} \mathbf{1}_A(x^1, x^2)\mu_Z(\mathrm{d}x^1 \times \mathrm{d}x^2).$$

The following equalities then hold:

$$\mu_{X^1}(A^1) = \mu_Z(A^1 \times \Lambda^2) \quad \text{and} \quad \mu_{X^2}(A^2) = \mu_Z(\Lambda^1 \times A^2)$$

for any $A^1 \in \mathcal{E}^1$ and $A^2 \in \mathcal{E}^2$. The probability measures μ_{X^1} and μ_{X^2} are called the *marginal measures* of μ_Z.

Let us now assume that $\Lambda^j = \mathbb{R}^{N_j}$ and and that X^j are absolutely continuous random variables. Then $Z : \Omega \to \mathbb{R}^{N_1+N_2}$ is absolutely continuous, with pdf denoted by Π_{X^1,X^2} and its marginal pdfs given by

$$\Pi_{X^1}(x^1) = \int_{\mathbb{R}^{N_2}} \Pi_{X^1,X^2}(x^1, x^2)\,\mathrm{d}x^2 \quad \text{and} \quad \Pi_{X^2}(x^2) = \int_{\mathbb{R}^{N_1}} \Pi_{X^1,X^2}(x^1, x^2)\,\mathrm{d}x^1.$$

For $X^j : \Omega \to \mathbb{R}^{N_j}$ we set $\mathbb{E}(X^j) := \mathbb{E}\left(\mathrm{id}(X^j)\right) \in \mathbb{R}^{N_j}$ and the *cross-covariance matrix*

$$\mathrm{Cov}(X^1, X^2) = \mathbb{E}\left(\left(X^1 - \mathbb{E}(X^1)\right)\left(X^2 - \mathbb{E}(X^2)\right)^T\right) \in M_{N_1\times N_2}(\mathbb{R}).$$

In particular, for $X : \Omega \to \mathbb{R}^N$ the *covariance matrix* is given by $\mathrm{Cov}(X) = \mathrm{Cov}(X, X) \in M_{N\times N}(\mathbb{R})$. In the special case $N_1 = N_2 = 1$ (a univariate random variable), the *correlation* is defined by

$$\mathrm{Corr}\left(X^1, X^2\right) = \frac{\mathbb{E}\left(\left(X^1 - \mathbb{E}(X^1)\right)\left(X^2 - \mathbb{E}(X^2)\right)\right)}{\sqrt{\mathbb{E}\left(\left(X^1 - \mathbb{E}(X^1)\right)^2\right) \cdot \mathbb{E}\left(\left(X^2 - \mathbb{E}(X^2)\right)^2\right)}} \in [-1, 1].$$

Definition 2.1.12 (Independence)**.** *The random variables $X^1 : \Omega \to \Lambda^1$ and $X^2 : \Omega \to \Lambda^1$ are* independent *if for any $A^1 \in \mathcal{E}^1$, $A^2 \in \mathcal{E}^2$ one has*

$$\mu_Z(A^1 \times A^2) = \mu_{X^1}(A^1)\,\mu_{X^2}(A^2).$$

In the special case of absolutely continuous random variables $X^1 : \Omega \to \mathbb{R}^{N_1}$ and $X^2 : \Omega \to \mathbb{R}^{N_2}$, the independence of X^1 and X^2 is equivalent to the condition $\Pi_{X^1,X^2} = \Pi_{X^1}\Pi_{X^2}$.

So far, X^1 and X^2 were defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and so does $Z = (X^1, X^2)$. Let us now briefly mention a converse situation, which is related to the famous optimal transportation problem.

Definition 2.1.13 (Transference plan)**.** *Given two probability measures μ_1, μ_2 on Λ, a* coupling *for them consists in a pair of random variables $Z := (X^1, X^2)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $X^j : \Omega \to \Lambda$ has induced probability measure μ_j. The joint measure μ_Z is called the* transference plan*, and the set of all transference plans is denoted by $\Pi(\mu_1, \mu_2)$.*

Note that a coupling always exists by considering X^1, X^2 independent, with the resulting joint measure given by the product of the two initial measures. Another special situation is considered in:

Definition 2.1.14 (Deterministic coupling)**.** *Let μ_1, μ_2 be two probability measures on $\mathbb{R}^N$. If there exists $\phi : \mathbb{R}^N \to \mathbb{R}^N$ bijective, differentiable and with inverse differentiable such that*

$$\int_{\mathbb{R}^N} f(x^2)\mu_2(\mathrm{d}x^2) = \int_{\mathbb{R}^N} f\big(\phi(x^1)\big)\mu_1(\mathrm{d}x^1)$$

for all "suitable" $f : \mathbb{R}^N \to \mathbb{R}$, then there exists a coupling $Z = (X, \phi(X))$ with transference plan

$$\mu_Z(\mathrm{d}x^1 \times \mathrm{d}x^2) = \delta(x^2 - \phi(x^1))\mu_1(\mathrm{d}x^1)\,\mathrm{d}x^2,$$

where δ denotes the Dirac delta function. Such couplings are called deterministic couplings*.*

Usually, one looks for a coupling having a special property:

Definition 2.1.15 (Monge-Kantorovitch problem)**.** *A transference plan μ_Z^* is a* solution of the Monge-Kantorovitch problem *for $\mu_1, \mu_2 : \mathbb{R}^N \to [0, \infty)$ with cost function $c(x^1, x^2) = \left\|x^1 - x^2\right\|^2$ if μ_Z^* satisfies*

$$\mu_Z^* = \underset{\mu_Z \in \Pi(\mu_1, \mu_2)}{\arg\inf} \ \mathbb{E}\left(\left\|X^1 - X^2\right\|^2\right).$$

Exercise 2.1.16. *If $N = 1$, show that this problem is equivalent to the maximization of* $\mathrm{Cov}(X^1, X^2)$.

For discrete-valued random variables, the notion of conditional pmf is quite clear, as for example when two random variables are defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and one of them takes a given value, what is the probability distribution of the other one? For absolutely continuous random variables, we shall use the following definition (but heuristic derivations exist).

Definition 2.1.17 (Conditional probability)**.** *For $j \in \{1, 2\}$ let $X^j : \Omega \to \mathbb{R}^{N_j}$ be absolutely continuous random variables, we set*

$$\Pi_{X^2|X^1}(x^2 \mid x^1) = \frac{\Pi_{X^2,X^1}(x^2, x^1)}{\Pi_{X^1}(x^1)} \tag{2.1.3}$$

whenever the denominator is not 0*, and call $\Pi_{X^2|X^1}(\cdot \mid x^1)$* the conditional probability distribution function *for X^2 under the assumption that X^1 takes the value x^1.*

Informally, we shall also say that $\Pi_{X^2|X^1}(\cdot \mid x^1)$ is the pdf of X^2 knowing x^1. For the numerator of (2.1.3) observe that $\Pi_{X^1,X^2}(x^1, x^2) = \Pi_{X^2,X^1}(x^2, x^1)$.

Definition 2.1.18 (Disintegration formula)**.** *Let $X^0, X^1, \ldots, X^n$ be a family of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. The following formula holds and is called the* disintegration formula*:*

$$\begin{aligned}\Pi_{X^n,\ldots,X^0}(x^n, \ldots, x^0) =& \Pi_{X^n|X^{n-1},\ldots,X^0}(x^n \mid x^{n-1}, \ldots, x^0) \cdot \Pi_{X^{n-1}|X^{n-2},\ldots,X^0}(x^{n-1} \mid x^{n-2}, \ldots, x^0) \ldots \\ & \ldots \Pi_{X^1|X^0}(x^1 \mid x^0) \cdot \Pi_{X^0}(x^0).\end{aligned}$$

Lemma 2.1.19 (Marginal and conditional pdf)**.** *In the above framework one has*

$$\Pi_{X^1}(x^1) = \mathbb{E}\left(\Pi_{X^1|X^0}(x^1 \mid X^0)\right).$$

Exercise 2.1.20. *Prove the above lemma.*

2.2 Discrete Time Markov Process

Recall that a discrete time dynamical system is defined by $x : \mathbb{Z} \to \mathbb{R}^N$ and by $f : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^N$ satisfying (1.2.1).

Definition 2.2.1 (Autonomous system)**.** *A system is* autonomous *if the function f does not depend explicitly on n, namely $f : \mathbb{R}^N \to \mathbb{R}^N$.*

Given $x(n_*)$ we have seen that $x(n_* + k)$ can be obtained by applying (1.2.1) k times. Instead of considering points, we can also consider sets $A(n_*) \subset \mathbb{R}^N$ and assume that $x(n_*)$ is inside $A(n_*)$. Then $A(n_* + 1) = f(A(n_*), n_*) = \{f(x, n_*) \mid x \in A(n_*)\}$, and similarly for $A(n_* + k)$. We also would like to associate some probabilities to these sets, and use the framework of Section 2.1.

Let $A \in \sigma_{\mathrm{B}}$ and set $\mathbb{P}^n(A)$ to be the probability that $x(n)$ is in A. Then, do we have $\mathbb{P}^{n+1}(f(A,n)) = \mathbb{P}^n(A)$? or what is the relation between $\mathbb{P}^{n+1}(B)$ and $\mathbb{P}^n(A)$ if B is another Borel set of $\mathbb{R}^N$? Thus, we are interested in defining a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable $X^n : \Omega \to \mathbb{R}^N$ with pdf $\Pi^n : \mathbb{R}^N \to [0,\infty)$ such that for any $A \in \sigma_{\mathrm{B}}$

$$\mathbb{P}^n(A) = \mathbb{P}\left(X^n \in A\right) = \mathbb{E}(\mathbf{1}_A\left(X^n\right)) = \int_A \Pi^n(x)\,\mathrm{d}x \equiv \int_A \Pi^n\left(x^n\right)\,\mathrm{d}x^n.$$

In this setting, Π^n represents the probability density that the system is at the point $x \in \mathbb{R}^N$ at time n.

We are also interested in knowing the relation between X^n and X^{n+1}, or equivalently between Π^n and Π^{n+1}, in order to see the relation between $\mathbb{P}^n(A)$ and $\mathbb{P}^{n+1}(A)$, if any? For example, if X^0 is given and allows us to define for any Borel set of $\mathbb{R}^N$ the probability that $x(0)$ belongs to this set, can one deduce X^1, which will define for any Borel set of $\mathbb{R}^N$ the probability that $x(1)$ is in this set? This question can (and will) be answered iteratively.

Definition 2.2.2 (Stochastic process)**.** *Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and set $\mathbb{N} := \{0, 1, 2, \dots\}$. A (discrete time)* stochastic process *is a family $(X^n)_{n\in\mathbb{N}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in a fixed measurable space $(\Lambda, \mathcal{E})$.*

This definition is too general for our purposes. A more convenient definition is:

Definition 2.2.3 (Absolutely continuous discrete time Markov process)**.** *An* absolutely continuous discrete time Markov process *with values in $\mathbb{R}^N$ consists of a stochastic process with each $X^n : \Omega \to \mathbb{R}^N$ absolutely continuous and satisfying for any $n \geq 1$ and $x^0, \dots, x^n \in \mathbb{R}^N$*

$$\Pi_{X^n|X^{n-1},\dots,X^0}(x^n \mid x^{n-1}, x^{n-2}, \dots, x^0) = \Pi_{X^n|X^{n-1}}(x^n \mid x^{n-1}). \tag{2.2.1}$$

In other words, in a discrete time Markov process, if X^{n-1} is known, knowing $X^{n-2}, \dots, X^0$ does not bring any additional information.

Remark 2.2.4. *1) Discrete time Markov process can be defined without the assumption of absolute continuity, but then the general definition is more involved since one can not use the notion of densities. However, discrete valued Markov process can easily be defined and are mentioned below. In the sequel, we shall only consider absolutely continuous or discrete valued Markov processes, and simply call them* Markov process.

2) Equation (2.2.1) can also be stated by the disintegration formula:

$$\Pi_{X^n,\dots,X^0}(x^n, x^{n-1}, \dots, x^0) = \Pi_{X^n|X^{n-1}}(x^n \mid x^{n-1})\Pi_{X^{n-1}|X^{n-2}}(x^{n-1} \mid x^{n-2}) \dots \Pi_{X^1|X^0}(x^1 \mid x^0)\Pi_{X^0}(x^0).$$

3) It also follows from (2.2.1) that the Chapman-Kolmogorov equation *holds:*

$$\Pi_{X^n}(x^n) = \int_{\mathbb{R}^N} \Pi_{X^n|X^{n-1}}(x^n \mid x^{n-1})\Pi_{X^{n-1}}(x^{n-1})\,\mathrm{d}x^{n-1}. \tag{2.2.2}$$

This equation is often written as $\Pi_{X^n} = \mathcal{P}\Pi_{X^{n-1}}$, or simply $\Pi^n = \mathcal{P}\Pi^{n-1}$. $\mathcal{P}$ is the transition operator, *assumed to be autonomous. If it is not autonomous, it is denoted by $\mathcal{P}^n$.*

Consider Λ a finite set $\{\lambda_i\}_{i=1}^N$ or a countable set $\{\lambda_i\}_{i\in\mathbb{N}}$, and let $\{X^n\}_{n\in\mathbb{N}}$ with $X^n : \Omega \to \Lambda$. This stochastic process is a *discrete valued discrete time Markov process* for any $x^0, \dots, x^n \in \Lambda$ the following condition is satisfied:

$$\mathbb{P}(X^n = x^n \mid X^{n-1} = x^{n-1}, \dots, X^0 = x^0) = \mathbb{P}(X^n = x^n \mid X^{n-1} = x^{n-1}).$$

In this setting and if assume that the system is autonomous, then one sets

$$p_{jk} := \mathbb{P}(X^n = \lambda_j \mid X^{n-1} = \lambda_k)$$

and this quantity is independent of n. It follows that $p_{jk} \geq 0$ and that $\sum_j p_{jk} = 1$. In this setting, equation (2.2.2) becomes

$$\mathbb{P}(X^n = \lambda_j) = \sum_k p_{jk}\, \mathbb{P}(X^{n-1} = \lambda_k)\,.$$

The matrix (p_{jk}) is called the *transition matrix* in the finite case, or the *transition operator* in the countable case. For shortness, we still write $\Pi^n = \mathcal{P}\Pi^{n-1}$ with $\Pi^n_j = \mathbb{P}(X^n = \lambda_j)$.

Definition 2.2.5 (Steady state)**.** *Let $\mathcal{P}$ be a transition operator for a discrete time Markov process. A solution of the equation $\Pi^n = \mathcal{P}\Pi^n$ is called a* steady state Π^*. *The corresponding measure $\Pi^* \mathrm{d}x$ in the absolutely continuous case, or Π^* in the discrete valued case, is called an* invariant measure.

If Π^0 is a steady state, all $\Pi^n := (\mathcal{P})^n\Pi^0$ have the distribution Π^0, and $\mathbb{E}(f(X^n)) = \mathbb{E}(f(X^0))$ for any $f : \Lambda \to \Xi$.

Definition 2.2.6 (Weakly stationary)**.** *A discrete time Markov process is* weakly stationary *if for all n_1, n_2, and $n \in \mathbb{N}$, $\mathbb{E}(X^{n_1}) = \mathbb{E}(X^{n_2})$ and $\mathrm{Cov}(X^{n_1+n}, X^{n_2+n}) = \mathrm{Cov}(X^{n_1}, X^{n_2})$.*

For a weakly stationary Markov process, we define the *autocovariance function*

$$C(n, m) := \mathrm{Cov}(X^n, X^m)$$

and observe that $C(n, m) = C(0, m - n)$ for any $m \geq n$. This function provides information about the loss of memory of the Markov process. Let us check that our initial example defines a Markov process:

Consider $f : \mathbb{R}^N \to \mathbb{R}^N$, let $(\xi^n)_{n\in\mathbb{N}}$ be a sequence of *i.i.d* (independent and identically distributed) absolutely continuous random variables of law Π_ξ, and let X^0 be an absolutely continuous random variable. Assume that

$$X^{n+1} = f(X^n) + \xi^n,$$

and observe that the deterministic part of this equation is autonomous. Then one has

$$\Pi_{X^n|X^{n-1},\dots,X^0}(x^n \mid x^{n-1}, \dots, x^0) = \Pi_{X^n|X^{n-1}}(f(x^{n-1}) + \xi^{n-1} \mid x^{n-1})\,.$$

This relation corresponds to the *Markov property*. It also follows from the assumption that for fixed x^n and for any $A \in \sigma_{\mathrm{B}}$

$$\begin{aligned} X^{n+1} = f(x^n) + \xi^n &\Longrightarrow \Pi_{X^{n+1}|X^n}(x^{n+1} \mid x^n) = \Pi_\xi(x^{n+1} - f(x^n)) \\ &\Longrightarrow \mathbb{P}(X^{n+1} \in A \mid x^n) = \int_A \Pi_\xi(x^{n+1} - f(x^n))\,\mathrm{d}x^{n+1} \end{aligned} \tag{2.2.3}$$

Exercise 2.2.7. *Check the veracity of the relations* (2.2.3).

We also have that

$$\begin{aligned} \mathbb{P}(X^{n+1} \in A) &= \int_A \Pi_{X^{n+1}}(x^{n+1})\,\mathrm{d}x^{n+1} \\ &= \int_A \Big[\int_{\mathbb{R}^N} \Pi_{X^{n+1}}(x^{n+1} \mid x^n)\,\Pi_{X^n}(x^n)\,\mathrm{d}x^n \Big] \mathrm{d}x^{n+1} \end{aligned} \tag{2.2.4}$$

$$= \int_A \Big[\int_{\mathbb{R}^N} \Pi_\xi(x^{n+1} - f(x^n)) \, \Pi_{X^n}(x^n) \, \mathrm{d}x^n \Big] \mathrm{d}x^{n+1} \tag{2.2.5}$$

By comparing (2.2.4) and (2.2.5) one then infers that

$$\Pi_{X^{n+1}}(x^{n+1}) = \int_{\mathbb{R}^N} \Pi_\xi(x^{n+1} - f(x^n)) \, \Pi_{X^n}(x^n) \, \mathrm{d}x^n$$

Alternatively, one writes this equation as

$$\Pi_{X^{n+1}} = \mathcal{P}\Pi_{X^n} \quad \text{with} \quad [\mathcal{P}\Pi](x) := \int_{\mathbb{R}^N} \Pi_\xi(x - f(x')) \, \Pi(x') \, \mathrm{d}x' \,. \tag{2.2.6}$$

Observe that this is the form of the Chapman-Kolmogorov equation mentioned in the item 3) of Remark 2.2.4.

2.3 Bayesian Inference

The aim of this section is to quantify everything (prior and posterior uncertainty) by using Bayes' theorem.

Recall that for any probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the conditional probability is defined for $A, B \in \mathcal{F}$ by

$$P(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Then Bayes' relation reads

$$\mathbb{P}(A \mid B) = \mathbb{P}(B \mid A) \frac{\mathbb{P}(A)}{\mathbb{P}(B)} \,.$$

Similarly, if X^1, X^2 are absolutely continuous random variables on $(\Omega, \mathcal{F}, \mathbb{P})$, one has

$$\Pi_{X^1|X^2}(x^1 \mid x^2) = \Pi_{X^2|X^1}(x^2 \mid x^1) \frac{\Pi_{X^1}(x^1)}{\Pi_{X^2}(x^2)} \,. \tag{2.3.1}$$

As an example of application, let us come back to the *inference model*:

$$Y = H(X) + \varepsilon$$

Here, $Y \in \mathbb{R}^M$ is the observation, $H : \mathbb{R}^N \to \mathbb{R}^M$ the continuous observation operator, X an absolutely continuous random variable with values in $\mathbb{R}^N$ and law Π_X, and ε the observation error, an absolutely continuous random variable with values in $\mathbb{R}^M$ and law Π_ε. In this case, Y is also an absolutely continuous random variable with

$$\Pi_Y(y) = \int_{\mathbb{R}^N} \Pi_\varepsilon(y - h(x)) \, \Pi_X(x) \, \mathrm{d}x. \tag{2.3.2}$$

The proof for this equality can be mimicked from the argument given at the end of the previous section. Computing Y, given X, is called the *forward problem*, but computing X, given Y, is the *inverse inference problem*. In this setting, one has:

$$\Pi_{X|Y}(x \mid y) = \frac{\Pi_{Y|X}(y \mid x) \, \Pi_X(x)}{\Pi_Y(y)} \,. \tag{2.3.3}$$

Here, Π_X is the distribution of X before any observation, called the *prior pdf*. $\Pi_{X|Y}(\cdot \mid y)$ is the distribution of X after the observation, named the *posterior pdf*. Lastly, $\Pi_{Y|X}(\cdot \mid x)$ is the likelihood of the observation given a particular value of X, called the *likelihood function*. Equation (2.3.3) is often written as

$$\Pi_{X|Y}(x \mid y) \propto \Pi_{Y|X}(y \mid x) \, \Pi_X(x) \,, \tag{2.3.4}$$

since we know that the density must integrate to 1. When $y = y_{\text{obs}}$, which means after that the observations took place, we write Π_{X^a} for $\Pi_{X|Y}(\cdot \mid y_{\text{obs}})$ and rewrite (2.3.4) as

$$\Pi_{X^a}(x) \propto \Pi_{Y|X}(y_{\text{obs}} \mid x)\Pi_{X^f}(x)\,, \tag{2.3.5}$$

where Π_{X^f} is the distribution of X before the observation. In this setting, Π_{X^a} is called the the pdf of *the analysis random variable*, while Π_{X^f} is called the pdf of *the forecast random variable*.

Let us now work on the simplest but seminal example.

Example 2.3.1. *Consider $Y = H(X) + \varepsilon$ with the following assumptions:*

$$\begin{aligned}
&H \in M_{M\times N}(\mathbb{R}),\\
&X \sim N(\bar{x}, P) \quad \textit{with} \quad \bar{x} \in \mathbb{R}^N \quad \textit{and} \quad P \in M_{N\times N}(\mathbb{R}), \quad P > 0,\\
&\varepsilon \sim N(\boldsymbol{0}, R) \quad \textit{with} \quad \boldsymbol{0} \in \mathbb{R}^M \quad \textit{and} \quad R \in M_{M\times M}(\mathbb{R}), \quad R > 0.
\end{aligned}$$

We are to solve both the forward problem and the inverse inference problem. Recall that

$$\begin{aligned}
\Pi_X(x) &= \frac{1}{(2\pi)^{N/2}|P|^{1/2}} \exp\left(-\frac{1}{2}(x-\bar{x})^T P^{-1}(x-\bar{x})\right)\\
\Pi_\varepsilon(z) &= \frac{1}{(2\pi)^{M/2}|R|^{1/2}} \exp\left(-\frac{1}{2}z^T P^{-1} z\right).
\end{aligned}$$

For the forward problem, by using (2.3.2) one has

$$\begin{aligned}
\Pi_Y(y) &= \int_{\mathbb{R}^N} \frac{(2\pi)^{-(N+M)/2}}{|P|^{1/2}|R|^{1/2}} \exp\left(-\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx) - \frac{1}{2}(x-\bar{x})^T P^{-1}(x-\bar{x})\right) \mathrm{d}x\\
&\propto \exp\left(-\frac{1}{2}y^T R^{-1} y + \frac{1}{2}d^T C^{-1} d\right),
\end{aligned}$$

where

$$C := P^{-1} + H^T R^{-1} H \ > 0 \quad \textit{and} \quad d := H^T R^{-1} y + P^{-1}\bar{x} \ \in \mathbb{R}^N.$$

For this computation, the completing-the-square formula

$$x^T C x - 2d^T x = (x - C^{-1}d)^T C (x - C^{-1}d) - d^T C^{-1} d$$

has been used.

For the inverse inference problem, one infers from (2.3.4) that

$$\begin{aligned}
\Pi_{X|Y}(x \mid y) &\propto \Pi_{Y|x}(y \mid x)\Pi_X(x)\\
&\propto \exp\left(-\frac{1}{2}(y-Hx)^T R^{-1}(y-Hx) - \frac{1}{2}(x-\bar{x})^T P^{-1}(x-\bar{x})\right)\\
&\propto \exp\left(-\frac{1}{2}(x-C^{-1}d)^T C (x-C^{-1}d)\right).
\end{aligned}$$

This means that $\Pi_{X^a} \equiv \Pi_{X|Y}(\cdot \mid y_{\text{obs}}) \sim N(\bar{x}^a, P^a)$ with

$$P^a := C^{-1} = (P^{-1} + H^T R^{-1} H)^{-1} \tag{2.3.6}$$

$$\bar{x}^a := C^{-1} d = \bar{x} - P^a H^T R^{-1}(H\bar{x} - y_{\text{obs}}) \tag{2.3.7}$$

Exercise 2.3.2. *Prove equations* (2.3.6) *and* (2.3.7).

Note that the previous expressions for P^{a} is quite inefficient since it involves the inverse of three matrices. This can be improved with the following lemma:

Lemma 2.3.3 (Woodbury formula)**.** *If $E \in M_{n\times n}(\mathbb{R})$, $G \in M_{m\times m}(\mathbb{R})$, both invertible, and if $F \in M_{n\times m}(\mathbb{R})$, $H \in M_{m\times n}(\mathbb{R})$, then the following equality holds:*

$$(E + FGH)^{-1} = E^{-1} - E^{-1}F(G^{-1} + HE^{-1}F)^{-1}HE^{-1}\,. \tag{2.3.8}$$

Exercise 2.3.4. *Prove Woodbury formula.*

By using this formula one gets the following expressions:

$$\begin{aligned} P^{\mathrm{a}} &= P - PH^T(HPH^T + R)^{-1}HP \\ \bar{x}^{\mathrm{a}} &= \bar{x} - PH^T(HPH^T + R)^{-1}(H\bar{x} - y_{\mathrm{obs}}). \end{aligned}$$

These formulae are called the *Kalman update formula.*

In the previous expressions, the posterior probability distribution $\Pi_{X|Y}$ of X knowing Y could be obtained. This is usually not the case, and one often gets only a partial information on this distribution, as for example a *point estimate*, namely a single value instead of a distribution. Let us introduce two general definitions:

Definition 2.3.5 (Cost function)**.** *A* cost function*, or a* loss function *is a measurable function from $\Lambda \times \Lambda$ to $\mathbb{R}$, where $(\Lambda, \mathcal{E})$ is the target space of some random variable.*

Definition 2.3.6 (Bayesian estimator)**.** *Given a posterior pdf $\Pi_{X|Y}(\cdot \mid y)$, a* Bayesian estimator *for X with loss function ℓ is defined by*

$$x^* := \underset{x'\in\Lambda}{\arg\min} \int_\Lambda \ell(x', x)\,\Pi_{X|Y}(x \mid y)\,\mathrm{d}x\,. \tag{2.3.9}$$

Note that in the previous definition, we have assumed that $\Lambda = \mathbb{R}^N$, which allows us to write the integral, but it could be more general. We continue with $\mathbb{R}^N$ since this situation appears often.

Examples 2.3.7. *1) If ℓ is the* quadratic loss function *defined by $\ell(x', x) = (x' - x)^2$, then,*

$$x^* = \int_{\mathbb{R}^N} x\,\Pi_{X|Y}(x \mid y)\,\mathrm{d}x.$$

This estimator is called the posterior mean estimate*.*

2) If ℓ is the absolute loss function *defined by $\ell(x', x) = |x' - x|$ and if $N = 1$, then x^* satisfies*

$$\int_{-\infty}^{x^*} \Pi_{X|Y}(x \mid y)\,\mathrm{d}x = \int_{x^*}^{\infty} \Pi_{X|Y}(x \mid y)\,\mathrm{d}x.$$

This estimator is called the posterior median estimate*.*

3) If the loss function is given by

$$\ell_\varepsilon(x', x) = \begin{cases} 1 & \text{if } \left\|x' - x\right\| > \varepsilon \\ 0 & \text{otherwise,} \end{cases}$$

then in the limit $\varepsilon \to 0_+$, it leads to the maximum a posterior (MAP) estimate*, where $\Pi_{X|Y}(\cdot \mid y)$ takes its global maximum.*

The Bayesian estimators still rely on the knowledge of the posterior distributions. Since these distributions are rarely known explicitly, we often use sampling techniques. Let us introduce the notion of *importance sampling*. For this, we come back to a measurable space $(\Lambda, \mathcal{E})$, and recall that a *measure* on it is a function $\mu : \Lambda \to [0, \infty]$ satisfying $\mu(\emptyset) = 0$ and

$$\mu\left(\cup_{j\in\mathbb{N}} A_j\right) = \sum_{j\in\mathbb{N}} \mu(A_j)$$

whenever $A_j \in \mathcal{E}$ and $A_j \cap A_k = \emptyset \ \forall j \neq k$. If ν is another measure on $(\Lambda, \mathcal{E})$, then ν is *absolutely continuous* with respect to a measure μ if $\nu(A) = 0$ whenever $A \in \mathcal{E}$ verifies $\mu(A) = 0$. In this case and if we assume an additional small technical assumption (μ is σ-finite, namely Λ can be decomposed into a countable union of measurable sets of finite μ-measure), then there exists a function $\rho : \Lambda \to [0, \infty)$ called the *Radon-Nikodym derivative* such that

$$\nu(A) = \int_A \rho(x)\mu(\mathrm{d}x)\,.$$

Observe that the notion just introduced is a generalization of the notion already discussed in Definition 2.1.6.

Now, if $\Lambda = \mathbb{R}^N$, if ν is absolutely continuous with respect a σ-finite measure μ, and if μ is iself absolutely continuous with respect to the usual Lebesgue measure on $\mathbb{R}^N$, then $\mu(\mathrm{d}x) = \Pi_\mu(x)\,\mathrm{d}x$ (or equivalently $\mu(A) = \int_A \Pi_\mu(x)\,\mathrm{d}x$) in which case $\nu(A) = \int_A \rho(x)\Pi_\mu(x)\mathrm{d}x$ which means that ν is absolutely continuous with respect to the Lebesgue measure. In this case we write $\nu(\mathrm{d}x) = \rho(x)\,\Pi_\mu(x)\,\mathrm{d}x$ for an equality between measures, or equivalently $\Pi_\nu(x) = \rho(x)\,\Pi_\mu(x)$ for an equality between pdfs.

In the discrete case, the notion of absolutely continuity of one measure with respect to another one is easy to understand. More precisely, if Λ is finite or countable, and if μ, ν are measures on Λ, then ν is absolutely continuous with respect to μ if $\nu(x) \neq 0$ for any $x \in \Lambda$ satisfying $\mu(x) \neq 0$.

Definition 2.3.8 (Importance sampling)**.** *Let X, X' be random variables with values in Λ and such that μ_X is absolutely continuous with respect to $\mu_{X'}$. Assume that $\mu_{X'}$ is a σ-finite measure, and let ω denote the Radon-Nikodym derivative. If $\{x_j\}_{j=1}^J \subset \Lambda$ are J independent samples from the random variable X, and if $\{x'_j\}_{j=1}^{J'} \subset \Lambda$ are J' independent samples from the random variable C', then for any measurable function $g : \Lambda \to \Xi$ one has*

$$\mathbb{E}(g(X)) \cong \frac{1}{J}\sum_{j=1}^{J} g(x_j) \cong \frac{1}{\sum_{k=1}^{J'} \omega(x'_k)} \sum_{j=1}^{J'} \omega(x'_j)\, g(x'_j). \tag{2.3.10}$$

If μ_X and $\mu_{X'}$ have a pdf or a pmf (in which case $\Lambda = \mathbb{R}^N$) denoted respectively by Π_X and $\Pi_{X'}$, then $\omega(x'_j) = \frac{\Pi_X(x'_j)}{\Pi_{X'}(x'_j)}$.

This approach is useful if we need the distribution of X while only the distribution of X' is provided. In the framework of the inference problem satisfying (2.3.5), one infers that for $g : \mathbb{R}^N \to \Xi$,

$$\mathbb{E}(g(X^{\mathrm{a}})) \cong \sum_{j=1}^{J} \frac{\Pi_{Y|X}(y_{\mathrm{obs}} \mid x_j^f)}{\sum_{k=1}^{J} \Pi_{Y|X}(y_{\mathrm{obs}} \mid x_k^f)} g(x_j^f)$$

where $\{x_j^f\}_{j=1}^J$ are J independent samples from the random variable X^f.

Chapter 3

Basic Algorithms of Data Assimilation

This chapter contains a few basic algorithms of data assimilation, such as the Kalman filter.

We firstly introduce the general framework. Consider a function $f : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^N$, A family of i.i.d. random variables $(\xi^n)_{n\in\mathbb{Z}}$ following a common distribution ξ, and let $(X^n)_{n\in\mathbb{Z}}$ be an absolutely continuous discrete time Markov process with values in $\mathbb{R}^N$. The evolution is described by the *discrete time dynamical system* defined by

$$X^{n+1} = f(X^n, n) + \xi^n. \tag{3.0.1}$$

The observations in this framework is described by the random variables

$$Y^{t_j} = H(X^{t_j}, t_j) + \varepsilon^{t_j},$$

where $\{t_1, \dots, t_k\} \subset \mathbb{Z}$ are the time point of the observations, $H : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^M$ is the *observation operator* which can be time dependent, and $\{\varepsilon^{t_j}\}$ are the *observation noise*, namely a family of i.i.d. random variables with values in $\mathbb{R}^M$.

Remark 3.0.1. *As already mentioned in Section 1.2, the model error $(\xi^n)_{n\in\mathbb{Z}}$ is replacing the unknown function g which is related to the true evolution, see (1.2.2).*

Set $y_{\text{obs}}^{t_1:t_k} := (y^{t_1}, \dots, y^{t_k}) \in M_{M\times k}$ for a matrix of observations. We then want to find the pdf $\Pi_{X^n|Y^{t_1:t_k}}(\cdot, n \mid y_{\text{obs}}^{t_1:t_k})$.

Definition 3.0.2 (Names of processes)**.** *If $n > t_k$, the process is called a* prediction process*. If $n \approx t_k$, the process is called a* filtering process*, while for $n < t_k$ one speaks about a* smoothing process*.*

This distinction is important since the tools developed for the 3 processes are slightly different.

3.1 The Filtering Process

Suppose that the observation error ε follows a $N(\mathbf{0}, R)$, and assume for simplicity that the observations take place at $t_n = n$ for $n \geq 1$. By the Chapman-Kolmogorov equation (2.2.2),

$$\Pi_{X^{1,f}}(x^1) := \Pi_{X^1}(x^1) = \int_{\mathbb{R}^N} \Pi_{X^1|X^0}(x^1 \mid x^0)\Pi_{X^0}(x^0)\,\mathrm{d}x^0\,. \tag{3.1.1}$$

By Bayes' theorem,

$$\Pi_{X^{1,a}}(x^1) := \Pi_{X^1|Y^1}(x^1 \mid y^1) = \frac{\Pi_{Y^1|X^1}(y^1 \mid x^1)\Pi_{X^{1,f}}(x^1)}{\int \Pi_{Y^1|X^1}(y^1 \mid x^1)\Pi_{X^{1,f}}(x^1)\,\mathrm{d}x^1}\,, \tag{3.1.2}$$

and with the assumption on ε, the first factor is given by:

$$\Pi_{Y^1|X^1}(y^1 \mid x^1) = \frac{1}{(2\pi)^{M/2}|R|^{1/2}} \exp\left(-\frac{1}{2}(y^1 - H(x^1,1))^T R^{-1}(y^1 - H(x^1,1))\right). \tag{3.1.3}$$

By Chapman-Kolmogorov, we can predict the pdf of X^2 (and this prediction includes the information obtained by the first observation):

$$\Pi_{X^{2,f}}(x^2) := \Pi_{X^2|Y^1}(x^2 \mid y^1) = \int_{\mathbb{R}^N} \Pi_{X^2|X^1}(x^2 \mid x^1)\Pi_{X^{1,a}}(x^1)\,\mathrm{d}x^1\,. \tag{3.1.4}$$

In (3.1.1) and (3.1.4), the expressions $\Pi_{X^1|X^0}(x^1 \mid x^0)$ and $\Pi_{X^2|X^1}(x^2 \mid x^1)$ are referred to as the *Markov kernel*, and are related to the equation (3.0.1). In this sense, it is natural to call these expressions a forecast, and to attach the superscript f to them. Once a new observation y^2 is available, one gets by Bayes' formula and as in (3.1.2):

$$\Pi_{X^{2,a}}(x^2) := \Pi_{X^2|Y^{1:2}}(x^2 \mid y^{1:2}) = \frac{\Pi_{Y^2|X^2}(y^2 \mid x^2)\Pi_{X^{2,f}}(x^2)}{\int \Pi_{Y^2|X^2}(y^2 \mid x^2)\Pi_{X^{2,f}}(x^2)\,\mathrm{d}x^2}\,, \tag{3.1.5}$$

where the first factor takes the form

$$\Pi_{Y^2|X^2}(y^2 \mid x^2) = \frac{1}{(2\pi)^{M/2}|R|^{1/2}} \exp\left(-\frac{1}{2}(y^2 - H(x^2,2))^T R^{-1}(y^2 - H(x^2,2))\right). \tag{3.1.6}$$

The expression (3.1.5) is called the analysis for X^2 and uses all available information. For arbitrary $t_n = n$, one has:

$$\Pi_{X^{n,f}}(x^n) := \Pi_{X^n|Y^{1:n-1}}(x^n \mid y^{1:n-1}) = \int_{\mathbb{R}^N} \Pi_{X^n|X^{n-1}}(x^n \mid x^{n-1})\Pi_{X^{n-1,a}}(x^{n-1})\,\mathrm{d}x^{n-1}\,, \tag{3.1.7}$$

$$\Pi_{X^{n,a}}(x^n) := \Pi_{X^n|Y^{1:n}}(x^n \mid y^{1:n}) = \frac{\Pi_{Y^n|X^n}(y^n \mid x^n)\Pi_{X^{n,f}}(x^n)}{\int_{\mathbb{R}^n} \Pi_{Y^n|X^n}(y^n \mid x^n)\Pi_{X^{n,f}}(x^n)\,\mathrm{d}x^n}\,. \tag{3.1.8}$$

This approach is called the *sequential data assimilation*.

Remark 3.1.1. *1) In the previous construction, we have assumed that observations are available at every $t_n = n$. If not, if we suppose that observations are available only every ℓ steps, then we evolve the system ℓ times with the Chapman-Kolmogorov equation, namely compute the forecast ℓ times, before performing the analysis step. This corresponds to applying the transition operator ℓ times instead of one time between each assimilation step.*

2) In the previous construction, we have assumed that the observation operator could be explicitly time dependent, but the same possibility holds for the model error ξ^n and for the observation errors ε^n. In this case, the expression for the Markov kernel would be different at each time step, and the expressions (3.1.3) and (3.1.6) would have additional dependences on n.

3.2 Kalman Filter (KF)

The *Kalman filter* is a special instance of the previous construction with the following additional assumptions:

$$\varepsilon^n \sim N(\mathbf{0}, R),$$
$$\xi^n \sim N(\mathbf{0}, Q),$$

$$f(x) = Dx + b, \quad D \in M_{N\times N}(\mathbb{R}),\ b \in \mathbb{R}^N,$$
$$H(x) = Hx, \quad H \in M_{M\times N}(\mathbb{R}),$$

where $Q \in M_{N\times N}(\mathbb{R})$ with $Q > 0$, and $R \in M_{M\times M}(\mathbb{R})$ with $R > 0$. Thus, the Kalman filter corresponds to a "linear" and autonomous evolution

$$X^{n+1} = DX^n + b + \xi^n \tag{3.2.1}$$

together with a "linear" and autonomous observation

$$Y^n = HX^n + \varepsilon^n .$$

Here, ε^n and ξ^n are Gaussian distributions of mean $\mathbf{0}$.

Remark 3.2.1. *1) Recall that if $X^n \sim N(\bar{x}, P^n)$, then as a consequence of Equation (3.2.1),*

$$X^{n+1} \sim N\left(D\bar{x}^n + b, DP^nD^T + Q\right).$$

2) Recall from Example 2.3.1 and from Equation (2.3.8) that if $X^f \sim N(\bar{x}^f, P^f)$, then $X^{\mathrm{a}} \sim N(\bar{x}^{\mathrm{a}}, P^{\mathrm{a}})$ with

$$P^{\mathrm{a}} := P^f - KHP^f$$
$$\bar{x}^{\mathrm{a}} := \bar{x}^f - K(H\bar{x}^f - y_{\mathrm{obs}})$$
$$K := P^fH^T(HP^fH^T + R)^{-1}$$

The matrix K is known as the Kalman gain matrix.

Thus, if we summarize the above results, and if $X^0 \equiv X^{0,\mathrm{a}} \sim N(\bar{x}^0, P^0)$, one gets:

Algorithm 3.2.2 (Kalman filter)**.** *The* Kalman filter algorithm *(KF) is given by:*

$$X^{n,f} \sim N\left(D\bar{x}^{n-1,\mathrm{a}} + b, DP^{n-1,\mathrm{a}}D^T + Q\right) =: N(\bar{x}^{n,f}, P^{n,f}),$$
$$X^{n,a} \sim N\left(\bar{x}^{n,f} - K^n(H\bar{x}^{n,f} - y^n), P^{n,f} - K^nHP^{n,f}\right) =: N(\bar{x}^{n,\mathrm{a}}, P^{n,\mathrm{a}}),$$
$$K^n := P^{n,f}H^T(HP^{n,f}H^T + R)^{-1}.$$

Remark 3.2.3. *In the previous setting, D (the evolution matrix), b (the evolution vector), Q (the model error covariance matrix), R (the observation error covariance matrix), and H (the observation matrix) could be time dependent.*

3.3 Variational Data Assimilation

This section is an improvement of what was sketched in Section 1.4 with the method of least squares. This variational approach is a smoothing technique which leads to the "best" initial condition, illustrated in Figure 3.1. We consider the autonomous model with no error: $x^{n+1} = f(x^n)$, and the observation operator $H : \mathbb{R}^N \to \mathbb{R}^M$. Let $x^{t_0,\mathrm{b}} \in \mathbb{R}^N$ be a prior (background) knowledge of the system at time t_0. If the assimilation window consists of n observations, then we define the cost function $L : \mathbb{R}^N \to \mathbb{R}$ by

$$L(x) = \frac{1}{2}(x - x^{t_0,\mathrm{b}})^T B^{-1}(x - x^{t_0,\mathrm{b}}) + \frac{1}{2}\sum_{i=1}^{n}\left(H(f^{\circ i}(x)) - y^{t_i}\right)^T R^{-1}\left(H(f^{\circ i}(x)) - y^{t_i}\right).$$

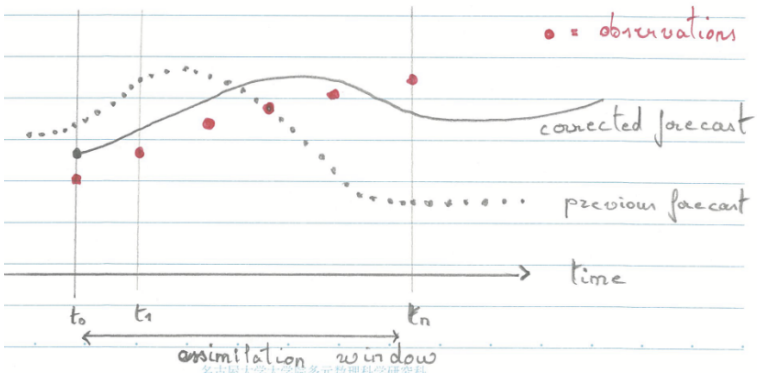

Figure 3.1: An illustration of the smoothing process, and the resulting corrected forecast.

Here, $\{y^{t_1}, \dots, y^{t_n}\} \subset \mathbb{R}^M$ are the observations, $R \in M_{M\times M}(\mathbb{R})$ is the observation error covariance matrix, and $B \in M_{N\times N}(\mathbb{R})$ is the prior (background) covariance matrix. R and B are chosen a priori, and are related to our confidence in the observations and in our prior knowledge of the system. The choice $R = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $B^{-1} = 0$ leads to the least squares method.

By a minimization process, we determine x^{t_0} which provides the minimum value for L. Then we get some forecast by computing $f^{\circ i}(x^{t_0})$ for any i. In particular, we can use $f^{\circ t_n}(x^{t_0})$ as the prior knowledge for the next step (the next assimilation window).

Remark 3.3.1. *If we set $x^{\mathrm{b}} := x^{t_0,\mathrm{b}}$ and $x = x^{\mathrm{b}} + h$ with h "small", then we can linearize some expansions:*

$$H(f^{\circ i}(x^{\mathrm{b}} + h)) \simeq H(f^{\circ i}(x^{\mathrm{b}})) + J_H(f^{\circ i}(x^{\mathrm{b}}))\, J_{f^{\circ i}}(x^{\mathrm{b}})\, h =: H(f^{\circ i}(x^{\mathrm{b}})) + \mathcal{H}^i L^i h$$

with $J_{f^{\circ i}}(x^{\mathrm{b}}) = \prod_{k=0}^{i-1} J_f(f^{\circ k}(x^k))$. Then,

$$L(h) = \frac{1}{2} h^T B^{-1} h + \frac{1}{2} \sum_{i=1}^{n} (\mathcal{H}^i L^i h - \delta^i)^T R^{-1} (\mathcal{H}^i L^i h - \delta^i),$$

where $\delta^i = y^{t_i} - H(f^{\circ i}(x^{\mathrm{b}}))$.

There exist several methods for finding the minimum of a function. The gradient descent method is a popular one, but the Gauss-Newton algorithm or the nonlinear conjugate gradient method are other (and more powerful) methods.

It is possible to add some randomness in the model by considering $X^{n+1} = f(X^n) + \xi_n$, but the framework becomes more complicated: see weak constraint 4DVar on pages 198–199 of [5].

3.4 Particle Filter

This is a method based on an ensemble of prediction combined with the importance sampling approach, see Definition 2.3.8. Numerous predictions are made, and the best ones are selected with the likelihood function, as introduced in equation (2.3.3).

We keep the general framework with ξ^n, ε^n i.i.d. random variables and the system of equations

$$X^{n+1} = f(X^n, n) + \xi^n$$

$$Y^n = H(X^n, n) + \varepsilon^n$$

without imposing any specific form to f, H, ξ^n, ε^n. We also assume that an observation is available at every $n \in \mathbb{N}$, but this is not a restriction, as emphasized in Remark 3.1.1 (otherwise we evolve the system ℓ steps before performing the assimilation).

We now further discuss two types of particle filters.

3.4.1 Sequential Importance Sampling (SIS)

Consider a random variable $X^0$and a sample $\{x_j^0\}_{j=1}^J$ from this random variable. Set the weights $\omega_j^0 := \frac{1}{J}$. The forecast system is realized at time 1 by $\{x_j^1, \omega_j^0\}_{j=1}^J$ with $x_j^1 := f(x_j^0, 0) + \xi_j^0$. For any measurable $g : \mathbb{R}^N \to \Xi$, the expectation value of the forecast is given by

$$\mathbb{E}(g(X^{1,f})) := \sum_{j=1}^{J} \omega_j^0 \, g(x_j^1).$$

Once the observation y^1 is available, we set

$$\omega_j^1; = \frac{\omega_j^0 \, \Pi_{Y^1|X^1}(y^1 \mid x_j^1)}{\sum_{j=1}^J \omega_j^0 \, \Pi_{Y^1|X^1}(y^1 \mid x_j^1)},$$

with $\Pi_{Y^1|X^1}(y^1 \mid x_j^1)$ the likelihood of the observation y^1 given x_j^1.

Example 3.4.1. *If we assume that $Y^n = H(X^n) + \varepsilon^n$ with $\varepsilon^n \sim N(\boldsymbol{0}, R)$, then*

$$\Pi_{Y^1|X^1}(y^1 \mid x_j^1) = \frac{1}{(2\pi)^M |R|^{1/2}} \exp\left(-\frac{1}{2}(y^1 - H(x_j^1))^T R^{-1}(y^1 - H(x_j^1))\right).$$

Thus, the weight ω_j^1 is small if $H(x_j^1)$ is far from the observation y^1.

The analysis (assimilated) system consists of $\{x_j^1, \omega_j^1\}_{j=1}^J$, and one sets $\mathbb{E}(g(X^{1,\mathrm{a}})) := \sum_{j=1}^J \omega_j^1 \, g(x_j^1)$.

Iteratively, given the assimilated system $\{x_j^n, \omega_j^n\}$ at time n, the forecast system at time $n+1$ is given by $\{x_j^{n+1}, \omega_j^n\}$ with $x_j^{n+1} = f(x_j^n, n) + \xi_j^n$, and the analysis system at time $n+1$ is given by $\{x_j^{n+1}, \omega_j^{n+1}\}_{j=1}^{n+1}$ with

$$\omega_j^{n+1} = \frac{\omega_j^n \, \Pi_{Y^{n+1}|X^{n+1}}(y^{n+1} \mid x_j^{n+1})}{\sum_{j=1}^J \omega_j^n \, \Pi_{Y^{n+1}|X^{n+1}}(y^{n+1} \mid x_j^{n+1})}.$$

The expectation values are respectively given by

$$\mathbb{E}(g(X^{n+1,f})) = \sum_{j=1}^{J} \omega_j^n \, g(x_j^{n+1}) \quad \text{and} \quad \mathbb{E}(g(X^{n+1,\mathrm{a}})) = \sum_{j=1}^{J} \omega_j^{n+1} \, g(x_j^{n+1}).$$

Unfortunately, this approach does not work well generally, because the weight will quickly concentrate on very few particles, and the system will become very unstable. Therefore, we need to concentrate more on realistic systems.

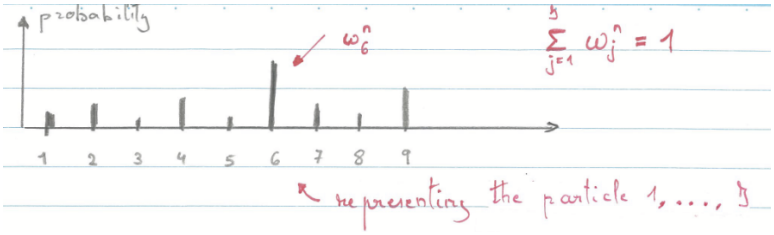

Figure 3.2: An illustration of the SIR process

3.4.2 Sequential Importance Resampling (SIR)

This method is also known as *bootstrap filter* or *condensation algorithm*. The approach is the same as the SIS, but a resampling is triggered whenever necessary.

Set $J^n_{\text{eff}} := \frac{1}{\sum_{j=1}^{J}(\omega_j^n)^2}$ to be the *effective sampling size*, and observe that $J^n_{\text{eff}} = J$ if $\omega_j^n = \frac{1}{J} \ \ \forall j$, while $J^n_{\text{eff}} \simeq 1$ if $\omega_j^n \simeq 1$ for one j, and $\omega_k^n \simeq 0$ for all $k \neq j$. Thus, J^n_{eff} is a way to measure the uniformity of the set of weights.

The SIR consists of the SIS together with a resampling when $J^n_{\text{eff}} < \alpha J$ for a fixed $\alpha < 1$. Once this threshold is met, a new sampling following the distribution $\{\omega_j^n\}_{j=1}^J$ is chosen. More precisely, we use a *multinomial distribution*: *Multinomial distribution* To each $j \in \{1, \dots, J\}$ we associate the probability ω_j^n (with $\sum_{j=1}^J \omega_j^n = 1$) and draw J independent numbers in $\{1, \dots, J\}$ according to these probabilities. An illustration is given in Figure 3.2. Clearly, some j's will be chosen more than once, while some j's might not be chosen at all. It turns out that $\mathbb{E}(\text{choosing particle } j \text{ at the resampling}) = J\omega_j^n$. We denote $\{z_k^n\}_{k=1}^J$ the corresponding set of new particles (for example, $z_1^n = x_2^n$, $z_2^n = x_6^n$, $z_3^n = x_6^n$, ...) according to the J trials. It then follows that

$$\mathbb{E}(g(X^{n,\mathrm{a}})) = \sum_{j=1}^{J} \omega_j^n \, g(x_j^n) \cong \sum_{j=1}^{J} \frac{1}{J} \, g(z_j^n).$$

We can then continue the SIS process with $\{z_j^n, \omega_j^n := \frac{1}{J}\}_{j=1}^J$ until the next threshold is met. Thus, the resampling allows us to concentrate on the most meaningful particles. When the resampling takes place, some particles will be identical, but because of the randomness, their trajectories will quickly move apart.

Remark 3.4.2. *This approach can be used for estimating some parameters, even time-dependent parameters, as we shall see in the next section.*

3.5 Application of Particle Filter to COVID-19

This section is about an application of the SIR approach introduction in the previous section. Its goal is to outline a method to estimate the effective reproduction number in real time for COVID-19, by using an agent-based model and a particle filter scheme. In this setting, one particle is one simulation of the epidemic in Tokyo, and we shall deal with $J = 100000$.

The model that we will be using is illustrated in Figure 3.3. Here, P_S, P_a, P_q, P_t, P_d, and P_r are probabilities. To each compartment (box), one associates a distribution for the number of days spent in the compartment (based on the medical information). In this setting, n represents the time in days. Then the random variable X^n is defined by

$$X^n := \left(E^n, I_a^n, I_s^n, T^n, D^n, R^n, R_t(n), P_d(n), T_h(n)\right)$$

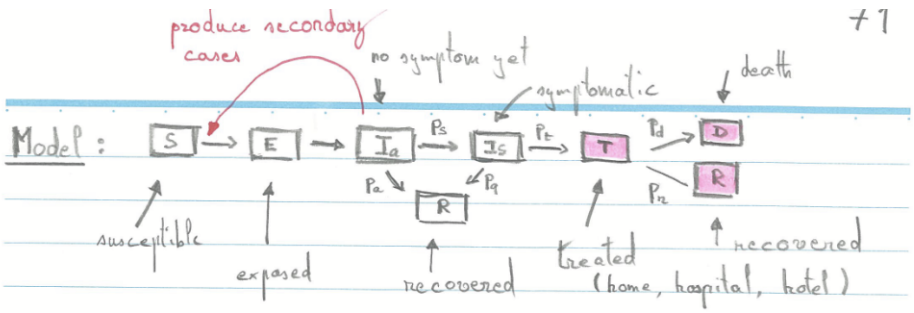

Figure 3.3: The model, with various boxes representing various states in the proliferation of COVID-19.

with $X^n \in \mathbb{N}^6 \times [0,\infty) \times [0,1] \times [0,\infty) \subset \mathbb{R}^9$. There are three unknown parameters, namely the effective reproduction number R_t, the probability of dying P_d, and the average time T_h spent in compartment T. We also define the random variable of observations $Y^n := (T^n, D^n, R^n) \in \mathbb{N}^3 \subset \mathbb{R}^3$ since they are the only reliable observations (highlighted pink in the figure).

In this setting, the observation operator $H : \mathbb{R}^9 \to \mathbb{R}^3$ is simply the projection on variables 4, 5, and 6, so $H \in M_{3\times 9}(\mathbb{R})$. Then, we get the system:

$$\begin{cases} X^{n+1} = f(X^n), \\ Y^n = HX^n + \varepsilon^n, \end{cases}$$

with $\varepsilon^n \sim N(\mathbf{0}, R^n)$, and f defined as a function of several variables and with values in $\mathbb{R}^9$. Note that the precise content of f is depends on the model chosen, but its nine components are sketched below. Note that we do not indicate the dependence on the fixed probabilities P_S, P_a, P_q, P_t, P_d, and P_r, but only on the components of the random variable X^n:

$$\begin{aligned} E^{n+1} &= f_1(I_a^n, R_t(n), n) \\ I_a^{n+1} &= f_2(E^n, n) \\ I_s^{n+1} &= f_3(I_a^n, n) \\ T^{n+1} &= f_4(I_s^n, T_h(n), n) \\ D^{n+1} &= f_5(T^n, T_h(n), P_d(n), n) \\ R^{n+1} &= f_6(T^n, T_h(n), 1 - P_d(n), n) \\ R_t(n+1) &= f_7(R_t(n), n) \\ P_d(n+1) &= f_8(P_d(n), n) \\ T_h(n+1) &= f_9(T_n(n), n) \end{aligned}$$

Note that each function contains some randomness, implicitly represented by the explicit dependence on the time variable n. By choosing a diagonal covariance matrix

$$R^n = \begin{pmatrix} \sigma_T(n)^2 & 0 & 0 \\ 0 & \sigma_D^2 & 0 \\ 0 & 0 & \sigma_R^2 \end{pmatrix} \in M_{3\times 3}(\mathbb{R}),$$

one gets the likelihood function

$$\Pi_{Y^n|X^n}(y^n \mid x^n) \propto \exp\left(-\frac{(T^n - T^n_{\text{obs}})^2}{2\sigma_T(n)^2} - \frac{(D^n - D^n_{\text{obs}})^2}{2\sigma_D^2} - \frac{(R^n - R^n_{\text{obs}})^2}{2\sigma_R^2} \right)$$

with $y^n = (T^n_{\text{obs}}, D^n_{\text{obs}}, R^n_{\text{obs}})$ the observations provided by the ministry of health.

By evolving simultaneously J systems like this, and by computing the weights ω^n_j for each of them, a SIR process can be implemented. Then, if X^n_k represents the k component of X^n for $k \in \{1, \dots, 9\}$ one infers its analysis value by

$$\mathbb{E}(X^{n,\text{a}}_k) = \sum_{j=1}^{J} \omega^n_j (x^n_j)_k.$$

For $k = 7$, one gets the effective reproduction number R_t. For more details, see [8].

Chapter 4

Variants and Extensions of the Kalman Filter

In this chapter, we extend the algorithms discussed in Chapter 3.

4.1 Extended Kalman Filter (EKF)

We consider again the system

$$\begin{cases} X^{n+1} = f(X^n, n) + \xi^n \\ Y^n = H(X^n, n) + \varepsilon^n \end{cases}, \tag{4.1.1}$$

where $f : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^N$ and $H : \mathbb{R}^N \times \mathbb{Z} \to \mathbb{R}^M$ are C^1-functions. We also suppose that all ξ^n and ε^n are independent and satisfy

$$\mathbb{E}(\xi^n) = \mathbf{0}$$
$$\mathbb{E}(\varepsilon^n) = \mathbf{0}.$$

The covariance matrices $Q^n \in M_{N\times N}(\mathbb{R})$ and $R^n \in M_{M\times M}(\mathbb{R})$ as then defined by

$$Q^n := \mathbb{E}(\xi^n {\xi^n}^T)$$
$$R^n := \mathbb{E}(\varepsilon^n {\varepsilon^n}^T)$$

Let us now assume that the analyzed random variable $X^{n,a} := X^n \mid Y^{1:n}$ is known by its mean value $\bar{x}^{n,a} = \mathbb{E}(X^{n,a})$ and by its covariance matrices $P^{n,a} = \mathbb{E}\Big((X^{n,a} - \bar{x}^{n,a})(X^{n,a} - \bar{x}^{n,a})^T\Big)$. We then set

$$\bar{x}^{n+1,f} := \mathbb{E}(f(X^{n,a}, n) + \xi^n) = \mathbb{E}(f(X^{n,a}, n)),$$

but this quantity is usually not computable since the full distribution of $X^{n,a}$ is not known. Thus we shall do an approximation, namely we assume that

$$f(x^{n,a}, n) \cong f(\bar{x}^{n,a}, n) + [J_f(\bar{x}^{n,a}, n)](x^{n,a} - \bar{x}^{n,a}) \tag{4.1.2}$$

where $J_f(\cdot, n)$ is the Jacobian matrix of $f(\cdot, n)$, obtained for any fixed n. The approximation is suitable if the system f is not too chaotic, and if the distribution of $X^{n,a}$ is rather peaked around its mean value. By (4.1.2), we infer that

$$\bar{x}^{n+1,f} = \mathbb{E}(f(X^{n,a}, n)) \cong f(\bar{x}^{n,a}, n) \tag{4.1.3}$$

and also deduce that

$$\begin{aligned} P^{n+1,f} &= \mathbb{E}\Big((f(X^{n,a}, n) + \xi^n - f(\bar{x}^{n,a}, n))(f(X^{n,a}, n) + \xi^n - f(\bar{x}^{n,a}, n))^T\Big) \\ &\cong \mathbb{E}\Big(\Big([J_f(\bar{x}^{n,a}, n)](X^{n,a} - \bar{x}^{n,a})\Big)\Big([J_f(\bar{x}^{n,a}, n)](X^{n,a} - \bar{x}^{n,a})\Big)^T\Big) + \mathbb{E}(\xi^n \xi^{nT}) \\ &= [J_f(\bar{x}^{n,a}, n)]P^{n,a}[J_f(\bar{x}^{n,a}, n)]^T + Q^n. \end{aligned} \tag{4.1.4}$$

The second equivalence is by using the approximation (4.1.2). With (4.1.3) and (4.1.4), we have obtained the mean value and the covariance matrix of the forecast random variable $X^{n+1,f}$ based on the mean value and on the covariance matrix of the analysed random variable $X^{n,a}$.

Now we compute $\bar{x}^{n+1,a}$ by updating $\bar{x}^{n+1,f}$. Suppose the observation y^{n+1} is available and adopt the following update strategy

$$\bar{x}^{n+1,a} = \bar{x}^{n+1,f} + K^{n+1}\Big(y^{n+1} - \mathbb{E}(H(X^{n+1,f}, n+1))\Big), \tag{4.1.5}$$

where K^{n+1} is the Kalman gain matrix defined below. Again, since the full distribution of $X^{n+1,f}$ is not known, we do an approximation similar to the one of (4.1.2), namely

$$H(x^{n+1,f}, n+1) \cong H(\bar{x}^{n+1,f}, n+1) + [J_H(\bar{x}^{n+1,f}, n+1)](x^{n+1,f} - \bar{x}^{n+1,f}),$$

from which we get

$$\bar{x}^{n+1,a} = \bar{x}^{n+1,f} + K^{n+1}\Big(y^{n+1} - H(\bar{x}^{n+1,f}, n+1)\Big). \tag{4.1.6}$$

For the Kalman gain matrix, if H is linear, the expression proposed in Algorithm 3.2.2 can be used, and in this case $J_H(\cdot, n) = H(\cdot, n)$. This expression leads to the smallest covariance matrix $P^{n+1,a}$, as shown by a variational approach in [1, Sec. 3.4]. If H is non-linear, one uses the linear approximation of H, namely its Jacobian matrix. By collecting the various steps, one finally gets:

Algorithm 4.1.1 (Extended Kalman filter)**.** *The extended Kalman filter (EKF) consists in the following steps:*

$$\begin{aligned} \bar{x}^{n,f} &= f(\bar{x}^{n-1,a}, n-1) \\ P^{n,f} &= [J_f(\bar{x}^{n-1,a}, n-1)]P^{n-1,a}[J_f(\bar{x}^{n-1,a}, n-1)]^T + Q^{n-1} \\ K^n &= P^{n,f}[J_H(\bar{x}^{n,f}, n)]^T\Big([J_H(\bar{x}^{n,f}, n)]P^{n,f}[J_H(\bar{x}^{n,f}, n)]^T + R^n\Big)^{-1} \\ \bar{x}^{n,a} &= \bar{x}^{n,f} + K^n\Big(y^n - H(\bar{x}^{n,f}, n)\Big) \\ P^{n,a} &= P^{n,f} - K^n[J_H(\bar{x}^{n,f}, n)]P^{n,f}. \end{aligned}$$

The linearization is the main difference between EKF and KF. Contrary to KF, the distributions of $X^{n,f}$ and $X^{n,a}$ are not assumed to be Gaussian in EKF: only their mean values and their covariance matrices are evaluated.

4.2 Ensemble Kalman Filter (EnKF)

We summarize the algorithms we have discussed so far:

- Kalman filter (KF): the procedure is simple but strong assumptions are needed,
- Variational DA (4DVar): it requires a lot of analysis and minimization processes,
- Particle filter (PF): an intuitive method with almost no assumption, but the computations are heavy,
- Extended Kalman filter (EKF): weaker assumptions are required, but not suitable for highly non-linear systems.

Note that in the particle filter, all particles evolve independently, only weights depend on all particles. In the forthcoming developments, the particles will interact at every observation / assimilation step. In this section, we introduce an algorithm called ensemble Kalman filter (EnKF) that possess the advantages of both the particle filter and the Kalman filter. We shall present two main approaches of EnkF: the stochastic EnKF and the deterministic EnKF.

4.2.1 The Stochastic EnKF

The underlying model is same as the one used in Section 4.1, namely (4.1.1). The main idea is to consider a set of particles $\{x_j^{n,f}\}_{j=1}^J$ at time n, known as forecast ensemble, and to obtain the analysis ensemble $\{x_j^{n,a}\}_{j=1}^J$ by the Kalman update formula

$$x_j^{n,a} = x_j^{n,f} + K^n(y_j^n - H(x_j^{n,f}, n)),$$

where K^n is the Kalman gain matrix and y_j^n will be defined later. Note that without the assumption of linearity of the model, the formula of Kalman gain given by Algorithm 3.2.2 can not be applied, and has to be suitably adapted.

Let's start with the forecast ensemble $\{x_j^{n,f}\}_{j=1}^J$ and define the *ensemble forecast mean* at time n by

$$\bar{x}^{n,f} := \frac{1}{J}\sum_{j=1}^{J} x_j^{n,f}.$$

Assume that the observation error ε^n follows a normal distribution $N(0, R^n)$, and let us draw a sample $\{\varepsilon_j^n\}_{j=1}^J$ following this distribution $N(0, R^n)$. Then the *sample mean* is $\bar{\varepsilon}^n := \frac{1}{J}\sum_{j=1}^J \varepsilon_j^n$. When the observation y_{obs}^n at time n is available, we define a family of observations

$$y_j^n := y_{\text{obs}}^n + \varepsilon_j^n - \bar{\varepsilon}^n,$$

which clearly satisfy

$$\bar{y}^n := \frac{1}{J}\sum_{j=1}^{J} y_j^n = y_{\text{obs}}^n.$$

In the sequel, each particle will be updated by using one of these observations y_j^n. This approach is consistent with the structure of the system where the observation error ε^n is added to the observation y_{obs}^n, and this trick also leads to a more stable algorithm.

For a second, let us assume that $H(\cdot, n)$ is linear, and denote it by H^n. Then, recall that the forecast *sample covariance matrix* is given by

$$P^{n,f} = \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})(x_j^{n,f} - \bar{x}^{n,f})^T$$

and observe that

$$\begin{aligned} P^{n,f}(H^n)^T &= \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})(x_j^{n,f} - \bar{x}^{n,f})^T(H^n)^T \\ &= \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})\Big(H^n x_j^{n,f} - \frac{1}{J}\sum_{k=1}^{J} H^n x_k^{n,f}\Big)^T \end{aligned} \tag{4.2.1}$$

and also

$$H^n P^{n,f}(H^n)^T = \frac{1}{J-1}\sum_{j=1}^{J}\Big(H^n x_j^{n,f} - \frac{1}{J}\sum_{k=1}^{J} H^n x_k^{n,f}\Big)\Big(H^n x_j^{n,f} - \frac{1}{J}\sum_{k=1}^{J} H^n x_k^{n,f}\Big)^T. \tag{4.2.2}$$

The point is to observe that (4.2.1) and the r.h.s. of (4.2.2) are well defined even if the observation operator $H^n \equiv H(\cdot, n)$ is not linear. Indeed, the observation operator H^n is always suitably associated with the forecast ensemble members, Thus we can re-write $P^{n,f}(H^n)^T$ and $H^n P^{n,f}(H^n)^T$ as

$$\frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})\Big(H(x_j^{n,f}, n) - \frac{1}{J}\sum_{k=1}^{J} H(x_k^{n,f}, n)\Big)^T$$

and as

$$\frac{1}{J-1}\sum_{j=1}^{J}\Big(H(x_j^{n,f}, n) - \frac{1}{J}\sum_{k=1}^{J} H(x_k^{n,f}, n)\Big)\Big(H(x_j^{n,f}, n) - \frac{1}{J}\sum_{k=1}^{J} H(x_k^{n,f}, n)\Big)^T.$$

These expressions can be understood as a cross-variance matrix and a covariant matrix.

For a fixed n, given a forecast ensemble $\{x_j^{n,f}\}_{j=1}^J$, the algorithm of stochastic EnKF reads as follows:

Algorithm 4.2.1 (Stochastic EnKF)**.**

(1) Draw $\{\varepsilon_j^n\}_{j=1}^J$ *and set* $y_j^n := y_{\text{obs}}^n + \varepsilon_j^n - \bar{\varepsilon}^n$,

(2) Set $\mathcal{X}_j^{n,f} := \frac{1}{\sqrt{J-1}}(x_j^{n,f} - \bar{x}^{n,f})$ *and* $\mathcal{H}_j^{n,f} := \frac{1}{\sqrt{J-1}}\Big(H(x_j^{n,f}, n) - \frac{1}{J}\sum_{k=1}^{J} H(x_k^{n,f}, n)\Big)$*, and consider the matrices* $\mathcal{X}^{n,f} \in M_{N\times J}(\mathbb{R})$ *and* $\mathcal{H}^{n,f} \in M_{M\times J}(\mathbb{R})$ *with the* $j-$*th columns* $\mathcal{X}_j^{n,f}$ *and* $\mathcal{H}_j^{n,f}$ *respectively,*

(3) Compute the Kalman gain matrix

$$K^n := \mathcal{X}^{n,f}(\mathcal{H}^{n,f})^T\Big(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n\Big)^{-1}, \tag{4.2.3}$$

(4) Compute the analysis ensemble $\{x_j^{n,a}\}_{j=1}^J$ *by*

$$x_j^{n,a} := x_j^{n,f} + K^n\Big(y_j^n - H(x_j^{n,f}, n)\Big),$$

(5) Compute the forecast ensemble $\{x_j^{n+1,f}\}_{j=1}^J$ *at time* $n+1$ *by*

$$x_j^{n+1,f} := f(x_j^{n,f}, n) + \xi^n.$$

Remark 4.2.2. *Step (3) can be modified by the following alternative expressions:*

(a) Replace R^n *by the sample covariance matrix* $\frac{1}{J-1}\sum_{j=1}^J \big((\varepsilon_j^n - \bar{\varepsilon}^n)(\varepsilon_j^n - \bar{\varepsilon}^n)^T\big)$.

(b) Replace $\mathcal{H}_j^{n,f}$ *by* $\frac{1}{\sqrt{J-1}}\Big(H(x_j^{n,f}, n) - \frac{1}{J}\sum_{k=1}^J H(x_k^{n,f}, n) + \varepsilon_j^n - \bar{\varepsilon}^n\Big)$ *and omit* R^n *in the expression of* K^n.

Note that these expressions for K^n *are not equal, but they converge to the same expression in the limit* $J \to \infty$. *Note also that the particles are no more independent since* K^n *depends on all of them. The stochastic EnKF adopts the Kalman update formula without the assumption of the linearity of the system. A big difference between EnKF and the particle filter introduced in Section 3.4 is that the evolution of each particle of EnKF takes the observations into account, while this is not the case for the particle filter approach. As a result, the particles in EnKF are constraint by the observations, and the algorithm is efficient already for a relatively small number of particles. This is rarely the case for the particle filter approach.*

4.2.2 The Deterministic EnKF

The philosophy of the subsequent approach is similar to the stochastic EnKF, but the observations y_{obs}^n are not perturbed. We adopt the notations $\mathcal{X}_j^{n,f}$ and $\mathcal{H}_j^{n,f}$ introduced in Algorithm 4.2.1.

Given a forecast ensemble $\{x_j^{n,f}\}_{j=1}^J$, suppose that the analysis mean is computed by the formula

$$\begin{aligned}
\bar{x}^{n,a} &:= \bar{x}^{n,f} + K^n(y_{\text{obs}}^n - H(\bar{x}^{n,f}, n)) \\
&= \bar{x}^{n,f} + \mathcal{X}^{n,f}(\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}(y_{\text{obs}}^n - H(\bar{x}^{n,f}, n)) \\
&= \bar{x}^{n,f} + \mathcal{X}^{n,f}\Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1}(\mathcal{H}^{n,f})^T(R^n)^{-1}(y_{\text{obs}}^n - H(\bar{x}^{n,f}, n))
\end{aligned} \tag{4.2.4}$$

where the expression for K^n is borrowed from (4.2.3) of Algorithm 4.2.1. The last equality uses the Woodbury identity, namely given appropriate matrices F, G and H,

$$F(G^{-1} + HF)^{-1} = (\mathbb{1} + FGH)^{-1}FG. \tag{4.2.5}$$

Note that the inverse in the expression (4.2.4) is taking place in $\mathbb{R}^J$, which is usually a low-dimensional space.

Now, recall that the forecast sample covariance matrix is defined by

$$P^{n,f} := \frac{1}{J-1}\sum_{j=1}^J (x_j^{n,f} - \bar{x}^{n,f})(x_j^{n,f} - \bar{x}^{n,f})^T = \mathcal{X}^{n,f}(\mathcal{X}^{n,f})^T.$$

Our aim is to construct an analysis sample covariance matrix $P^{n,a}$ of a similar form, and to compute the corresponding analysis ensemble. Thus, we shall consider

$$P^{n,a} = \mathcal{X}^{n,a}(\mathcal{X}^{n,a})^T,$$

where $\mathcal{X}^{n,\text{a}}$ is some matrix which has to be defined. We shall also adopt the relation $P^{n,a} = (\mathbb{1} - K^nH^n)P^{n,f}$ used by KF and EKF. This relation is at the root of determining the best Kalman gain matrix, see [1, pp. 92-95].

By the formula of K^n in Algorithm 4.2.1, one has

$$\begin{aligned} P^{n,a} &= (\mathbb{1} - K^n H^n) P^{n,f} \\ &= P^{n,f} - K^n H^n P^{n,f} \\ &= \mathcal{X}^{n,f}(\mathcal{X}^{n,f})^T - \mathcal{X}^{n,f}(\mathcal{H}^{n,f})^T \Big(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n\Big)^{-1} \mathcal{H}^{n,f}(\mathcal{X}^{n,f})^T \\ &= \mathcal{X}^{n,f}\Big(\mathbb{1}_{J\times J} - (\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}\mathcal{H}^{n,f}\Big)(\mathcal{X}^{n,f})^T. \end{aligned} \tag{4.2.6}$$

Note that $0 \leq (\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}\mathcal{H}^{n,f} \leq \mathbb{1}$, and consequently $\mathbb{1}_{J\times J} - (\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}\mathcal{H}^{n,f} \geq 0$. It follows that the square root of $\mathbb{1}_{J\times J} - (\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}\mathcal{H}^{n,f}$ in the following definition will be well-defined. Be aware that one is dealing with the square root of a matrix, namely for any $A \geq 0$ there exists $B \geq 0$ with $B^2 = A$.

Set

$$\mathcal{X}^{n,a} := \mathcal{X}^{n,f}\Big[\mathbb{1}_{J\times J} - (\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1}\mathcal{H}^{n,f}\Big]^{\frac{1}{2}} U,$$

and one gets $P^{n,a} = \mathcal{X}^{n,a}(\mathcal{X}^{n,a})^T$ by (4.2.6) if $U \in M_{J\times J}(\mathbb{R})$ satisfies $UU^T = \mathbb{1}$. This relation means precisely that U is an orthogonal matrix. By Woodbury identity (4.2.5), one infers that

$$(\mathcal{H}^{n,f})^T(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n)^{-1} = (\mathbb{1}_{J\times J}(\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f})^{-1}(\mathcal{H}^{n,f})^T(R^n)^{-1}$$

which leads to

$$\begin{aligned} &\mathbb{1}_{J\times J} - (\mathcal{H}^{n,f})^T\Big(\mathcal{H}^{n,f}(\mathcal{H}^{n,f})^T + R^n\Big)^{-1}\mathcal{H}^{n,f} \\ &= \mathbb{1}_{J\times J} - \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1}(\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f} \\ &= \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1}\Big[\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f} - (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big] \\ &= \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1} \end{aligned}$$

and thus

$$\mathcal{X}^{n,a} = \mathcal{X}^{n,f}\Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-\frac{1}{2}} U.$$

Let us now set

$$T := \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1},$$

and then we can re-write the formula of $\mathcal{X}^{n,a}$ as

$$\mathcal{X}^{n,a} = \mathcal{X}^{n,f} T^{\frac{1}{2}} U.$$

From the general expression for the sample covariance matrix $P^{n,a} = \mathcal{X}^{n,a}(\mathcal{X}^{n,a})^T = \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,a} - \bar{x}^{n,a})(x_j^{n,a} - \bar{x}^{n,a})^T$, one is led to define

$$x_j^{n,a} := \bar{x}^{n,a} + \sqrt{J-1}(\mathcal{X}^{n,a})_j$$

where $(\mathcal{X}^{n,a})_j$ is the j-th column of the matrix $\mathcal{X}^{n,a}$.

The previous construction corresponds to the *ensemble transform Kalman filter* (ETKF) which is one variant of EnKF. The algorithm of ETKF is summarized as below:

Algorithm 4.2.3 (Ensemble transform Kalman filter)**.** *Given a set of forecast ensemble members* $\{x_j^{n,f}\}_{j=1}^J$ *at time n, the Ensemble transform Kalman filter (ETKF) corresponds to:*

(1) Set $\bar{x}^{n,f} := \frac{1}{J}\sum_{j=1}^J x_j^{n,f}$, *and* $\mathcal{X}_j^{n,f} := \frac{1}{\sqrt{J-1}}(x_j^{n,f} - \bar{x}^{n,f})$, *and* $\mathcal{H}_j^{n,f} := \frac{1}{\sqrt{J-1}}\Big(H(x_j^{n,f},n) - \frac{1}{J}\sum_{k=1}^J H(x_k^{n,f},n)\Big)$, *and set* $\mathcal{X}^{n,f} := (x_1^{n,f},\dots,x_J^{n,f}) \in M_{N\times J}(\mathbb{R})$, $\mathcal{H}^{n,f} := (\mathcal{H}_1^{n,f},\dots,\mathcal{H}_J^{n,f}) \in M_{M\times J}(\mathbb{R})$.

(2) Set $T := \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T (R^n)^{-1}\mathcal{H}^{n,f}\Big)^{-1} \in M_{J\times J}(\mathbb{R})$.

(3) Choose $U \in M_{J\times J}(\mathbb{R})$, *orthogonal and satisfying*

$$U\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix} = \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$$

and set

$$x_j^{n,a} := \bar{x}^{n,f} + \mathcal{X}^{n,f}T(\mathcal{H}^{n,f})^T(R^n)^{-1}(y_{\text{obs}}^n - H(\bar{x}^{n,f},n)) + \sqrt{J-1}\Big[\mathcal{X}^{n,f}T^{\frac{1}{2}}U\Big]_j. \tag{4.2.7}$$

(4) One has $P^{n,a} = \mathcal{X}^{n,a}(\mathcal{X}^{n,a})^T$, *and one sets* $x_j^{n+1,f} := f(x_j^{n,a},n) + \xi_j^n$.

Remark 4.2.4. *For the analysis ensemble* $\{x_j^{n,a}\}_{j=1}^J$ *computed by ETKF, one would like to check that it satisfies*

$$\frac{1}{J}\sum_{j=1}^J x_j^{n,a} = \bar{x}^{n,a}$$

or equivalently

$$\sum_{j=1}^J \Big(\mathcal{X}^{n,f}T^{\frac{1}{2}}U\Big)_j = \boldsymbol{0}.$$

If we set $\boldsymbol{I} = \begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}$, *this reads* $\mathcal{X}^{n,f}T^{\frac{1}{2}}U\boldsymbol{I} = \boldsymbol{0}$. *By the assumption on U, this means* $\mathcal{X}^{n,f}T^{\frac{1}{2}}\boldsymbol{I} = \boldsymbol{0}$. *By the definition of* $\mathcal{X}^{n,f}$, *one observes that* $\mathcal{X}^{n,f}\boldsymbol{I} = \boldsymbol{0}$, *then it is enough to show that* $T^{\frac{1}{2}}\boldsymbol{I} = \boldsymbol{I}$. *It means that* 1 *is an eigenvalue of the symmetric matrix* $T^{\frac{1}{2}}$ *with the eigenvector* $\boldsymbol{I}$. *However, by diagonalization, this is equivalent to* $T^{-1}\boldsymbol{I} = 1\boldsymbol{I}$. *Namely* $\Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T(R^n)^{-1}\mathcal{H}^{n,f}\Big)\boldsymbol{I} = 1\boldsymbol{I}$. *Since* $\mathcal{H}^{n,f}\boldsymbol{I} = \boldsymbol{0}$ *(as* $\mathcal{X}^{n,f}\boldsymbol{I} = \boldsymbol{0}$*), one ends up with* $\mathbb{1}_{J\times J}\boldsymbol{I} = 1\boldsymbol{I}$, *which is obviously true.*

4.3 Inflation and Localisation

In many applications, the components of the random variables X^n or Y^n represent a certain quantity at a certain point p_ℓ in space and at time n. For example, the discretization of a continuous model leads typically to $10^6 \sim 10^{10}$ such points $\{p_\ell\}_{\ell=1}^L$ with $p_\ell \in \mathbb{R}^d$ for $d = 2$ or 3. For comparison, the ensemble size J is usually of order 10 or 100.

An intuitive question is then raised: Can one really represent the covariance matrix of the physical system with so few members? In other words, can one represent the spread of all possible evolutions with so few simulations? If the spread of the ensemble members (measured with the covariance) becomes too small, then the observations can not be suitably assimilated. This situation is called a *filter divergence*. For solving this problem, one applies a so-called *covariance inflation*. There exist two main types of covariance inflation:

1. *Multiplicative inflation*: Right after the forecast step, we replace $x_j^{n,f}$ by $\bar{x}^{n,f} + \varrho(x_j^{n,f} - \bar{x}^{n,f})$ for some $\varrho > 1$. As a result, the new forecast mean is not changed, but the covariance matrix is multiplied by ϱ^2.

2. *Additive inflation*: This corresponds to adding a larger model error ξ_j^n. Thus, after the forecast step, we add to $x_j^{n+1,f} := f(x_j^{n,a}, n) + \xi_j^n$ a new term $\xi_j'^n$ which is a realization of variable ξ'^n with $\mathbb{E}(\xi'^n) = 0$. Or we add a term $\xi_j'^n - \bar{\xi}'^n$ where the mean value $\bar{\xi}'^n$ on the sample $\{\xi_j'^n\}_{j=1}^J$ takes the value 0. As a result, this does not change the mean value of the forecast sample, but it increases its covariance matrix.

Note that the two inflation methods can be applied separately or simultaneously. Note also that the role of ξ_j^n and of $\xi_j'^n$ is not exactly the same: ξ_j^n should correct a model error, which means that it can encode some prior knowledge and in particular it can be biased (mean different from 0) if necessary. The new term $\xi_j^{n'}$ is added for increasing the covariance matrix and has no particular meaning. As a result, it is natural to impose that it is unbiased. There exists another subtle difference between the two inflation methods: the multiplication inflation preserves the subspace generated by the ensemble $\{x_j^{n,f}\}_{j=1}^J$, while the additive inflation might add some contributions in any directions of $\mathbb{R}^N$.

Another problem which often takes place for large system is related to localisation, as presented now. So far, the Kalman gain matrix is a global object, defined simultaneously with all components of X^n, often through the term $P^{n,f}$. However, this construction creates some spurious correlation between quantities at related to different points p_ℓ which can be far apart in space. For solving this problem, we need to introduce the notion of space localisation, and thus provide a more suitable indexation for the components of X^n and of Y^n.

Thus, let us consider a set $\Lambda := \{p_\ell\}_{\ell=1}^L$ of points in space, and let $\mathcal{I} := \{1, \dots, I\}$ denote an indexation of all possible information related to each point p_ℓ of Λ. More concretely, the set Λ could be a grid on earth, and the index $\mathcal{I}$ could be used for indexing information like temperature, three components of wind, humidity, *etc*, at each point of Λ. In the sequel, we shall consider Λ_X and $\mathcal{I}_X$ as the set of all points and all information used for describing the model, while the set $\Lambda_Y \subset \Lambda_X$ and $\mathcal{I}_Y \subset \mathcal{I}_Y$ correspond respectively to the points and to the information related to the observations. Typically, the observations are taken at fewer points, compared to the grid necessary for running the model, and the measurements can not be obtained on all variables necessary for the simulations. Observe that if we set $|\Omega|$ for the cardinality of the set Ω (namely the number of its elements), then the following relations hold

$$N = |\Lambda_X||\mathcal{I}_X| = LI \qquad \text{and} \qquad M = |\Lambda_Y||\mathcal{I}_Y|. \tag{4.3.1}$$

Note also that these sets Λ_X, Λ_Y, $\mathcal{I}_X$, and $\mathcal{I}_Y$, could be time dependent (with an additional index n) but that we do not consider this additional dependence here, mainly for simplicity.

With the notations introduced above, we shall index the components of X^n by $X_{p,i}^n \in \mathbb{R}$ with $p \in \Lambda_X$ and $i \in \mathcal{I}_X$, meaning that this component of X^n corresponds to the information i at the point p. Similarly, the components of Y^n are denoted by $Y_{p,i}^n$ with $p \in \Lambda_Y$ and $i \in \mathcal{I}_Y$. We are now ready for introducing two methods of localisation, the *covariance localisation* and the *R-localisation*.

4.3.1 Covariance localisation

The covariance matrix for the forecast ensemble is defined by

$$P^{n,f} = \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})(x_j^{n,f} - \bar{x}^{n,f})^T$$

which means that the element $(p,i)(p',i')$ of the matrix is given by

$$(P^{n,f})_{(p,i)(p',i')} = \frac{1}{J-1}\sum_{j=1}^{J}(x_j^{n,f} - \bar{x}^{n,f})_{(p,i)}(x_j^{n,f} - \bar{x}^{n,f})_{(p',i')},$$

with $p, p' \in \Lambda_X$ and $i, i' \in \mathcal{I}_X$. We also introduce a correlation matrix providing the information about the correlation, namely, $\rho \in M_{N\times N}(\mathbb{R})$ with

$$\rho_{(p,i)(p',i')} = C_{ii'}(\|p - p'\|) \tag{4.3.2}$$

where $C_{ii'} = C_{i'i} : \mathbb{R}_+ \to [0,1]$ describes the range of correlation. Here $\|p - p'\|$ denotes the Euclidean distance between the points p and p'. Usually, we impose that the functions $C_{ii'}$ vanishes rather quickly, implying later that two points far apart are not going to have any correlation. Note also that we can choose the same function for all pairs (i,i') if the correlation does not depend on the precise information. In the general case the decay of correlation would depend on the information (for example different correlation decays for the wind, the temperature, or the humidity, ...). Very importantly, we impose that $\rho \geq 0$, meaning that ρ is a positive semi-definite matrix, see for example [6, Sec. 2.2].

We can then define a new covariance matrix

$$P_\rho^{n,f} := \rho \circ P^{n,f} = \left(\rho_{(p,i)(p',i')}P^{n,f}_{(p,i)(p',i')}\right)_{(p,i),(p',i')}.$$

The Hadamard product $\circ$, also called pointwise product, has the property that given two matrices A and B with $A \geq 0$ and $B \geq 0$, then one has $A \circ B \geq 0$. As a result, since ρ is positive semi-definite and since $P^{n,f}$ is positive definite, the resulting matrix $P_\rho^{n,f}$ is also positive semi-definite. As a result of this construction, the correlation coefficient $(P_\rho^{n,f})_{(p,i)(p',i')}$ is going to be 0 or very small if p and p' are far apart.

If we assume now that H is linear, then the expression $P_\rho^{n,f}$ can be introduced in the expression of the Kalman gain matrix, namely one gets

$$K^n = P_\rho^{n,f}H^T\left(HP_\rho^{n,f}H^T + R\right)^{-1} = (\rho \circ P^{n,f})H^T\left(H(\rho \circ P^{n,f})H^T + R\right)^{-1}. \tag{4.3.3}$$

In addition, if $C_{ii'}$ does not depend on i and i' (which means that $\rho_{(p,i)(p',i')=C(\|p-p'\|)\equiv\rho_{pp'}}$) and if H is local, then one can go one step further. Here H local means that $H_{(p,i)(p',i')} = h_{p,i,i'}\delta_{pp'}$ for all $p \in \Omega_Y$, $p' \in \Omega_X$, $i \in \mathcal{I}_Y$, $i' \in \mathcal{I}_X$, and $\delta_{pp'}$ represents the Kronecker delta function. In other words, the entry $H_{(p,i)(p',i')}$ of he matrix H is 0 whenever p is not equal to p', and when $p = p'$ the entry is the number $h_{p,i,i'}$. With this assumption, one gets:

$$\begin{aligned}\left((\rho \circ P^{n,f})H^T\right)_{(p,i)(p',i')} &= \sum_{q,j}\rho_{pq}P^{n,f}_{(p,i)(p',i')}h_{p',i',j}\delta_{p'q} \\ &= \rho_{pq}\sum_j P^{n,f}_{(p,i)(p',i')}h_{p',i',j}\end{aligned}$$

$$= \Big(\rho \circ (P^{n,f} H^T)\Big)_{(p,i)(p',i')}.$$

Similarly, one gets

$$H(\rho \circ P^{n,f})H^T = \rho \circ (HP^{n,f}H^T)$$

which leads to the final expression

$$K^n = \rho \circ (P^{n,f} H^T)\Big(\rho \circ (HP^{n,f}H^T) + R^n\Big)^{-1} \tag{4.3.4}$$

which is often the formula provided for the covariance localisation in the literature.

Remark 4.3.1.
1. *If H is not linear, we can either consider its linearization, as in the EKF approach, or adapt the approach provided in Algorithm 4.2.1.*
2. *If H is not local but only approximately local, meaning that $H_{(p,i)(p',i')} = 0$ as soon as $\|p - p'\|$ is bigger than a fixed constant, then formula (4.3.4) can be adapted.*
3. *In ETKF, in which $P^{n,f}$ does not play any role, this approach can not be applied.*

4.3.2 R-localisation

As mentioned in the last remark, the covariance localisation can not be used (of justified) when $P^{n,f}$ does not appear explicitly in the analysis step. On the other hand, even in Algorithm 4.2.3 about ETKF, the observation error covariance matrix R^n appears, and we shall use it for the localisation process. The name R-localisation comes from this matrix.

Recall from (4.2.7) that the analysis ensemble at time n for ETKF is given by

$$x_j^{n,a} := \bar{x}^{n,a} + \sqrt{J-1}\Big[\mathcal{X}^{n,f} T^{\frac{1}{2}} U\Big]_j. \tag{4.3.5}$$

with

$$\begin{aligned} \bar{x}^{n,a} &:= \bar{x}^{n,f} + \mathcal{X}^{n,f} T (\mathcal{H}^{n,f})^T (R^n)^{-1} (y^n_{\text{obs}} - H(\bar{x}^{n,f}, n)) \\ T &:= \Big(\mathbb{1}_{J\times J} + (\mathcal{H}^{n,f})^T (R^n)^{-1} \mathcal{H}^{n,f}\Big)^{-1} \in M_{J\times J}(\mathbb{R}). \end{aligned} \tag{4.3.6}$$

With the notations introduced above and for each $p \in \Lambda_X$ and $i \in \mathcal{I}_X$, let us also introduce a diagonal matrix $D(p,i) \in M_{M\times M}(\mathbb{R})$ with

$$D(p,i)_{(p',i')(p',i')} = C_{ii'}(\|p - p'\|) \tag{4.3.7}$$

with $C_{ii'} : \mathbb{R}_+ \to [0,1]$ a decreasing function. As before, in the simplest case, this function does not depend on i and i', which means that $D(p,i)_{(p',i')(p',i')} = C(\|p - p'\|)$ for some decreasing function $C : \mathbb{R}_+ \to [0,1]$. Note also that only the diagonal element of $D(p,i)$ are defined in (4.3.7) since the off-diagonal term are 0.

Clearly, the matrix $D(p,i)$ provides an information about the distance between a fixed point p and another point p'. If we consider $p, p' \in \Lambda_Y$, the set of points where observations are taken, and if $i, i' \in \mathcal{I}_Y$, the set related to the measured quantities at each observation points, then the matrix $D(p,i)$ is a $M \times M$ diagonal matrix, see also (4.3.1). Thus, when computing $(x_j^{n,a})_{(p,i)}$ (the component (p,i) of the random variable $X_j^{n,a}$ corresponding to the j^{th} member of the analysis ensemble at time n) with formula (4.3.5), one can replace $(R^n)^{-1}$ by $D(p,i)(R^n)^{-1}$ in the formula for $\bar{x}^{n,a}$ and T, see (4.3.6). Since R^n is often diagonal, it follows that the computations of $(R^n)^{-1}$ and of $D(p,i)(R^n)^{-1}$ are usually quite simple. As a result, when computing $(x_j^{n,a})_{(p,i)}$, only quantities close to the point p matter, and in particular, only the information of the *innovation term*

$y^n_{\text{obs}} - H(\bar{x}^{n,f}, n)$ close to the point p is taken into account. This happens to be true since this expression will be preceded by $D(p,i)(R^n)^{-1}$.

This R-localisation procedure leads to the so-called *local ensemble transform Kalman filter* or in short LETKF. Very briefly, it corresponds to a ETKF applied locally to each component of $X^{n,a}_j$.

4.4 Parameter estimation

Quite often a model involves some unknown parameters and one aim of data assimilation techniques is to evaluate them. Let $\Theta^n \in \mathbb{R}^d$ denote these parameters, which can be time (n) dependent or not. We review some of the algorithms considered so far, and emphasize how these parameters can be evaluated.

In the framework of particle filters, as developed in Section 3.4, an initial value for Θ^0_j is provided according to an initial distribution or according to some prior knowledge. As a consequence, the evolution of the J particles begin with independent values for the unknown parameters Θ^0. At the forecast step, a new value for the parameters is obtained with the equation $\Theta^{n+1}_j = \Theta^n_j + \zeta^n_j$, where ζ^n_j usually follows a distribution of mean 0. Depending on the system and on the knowledge for the evolution of the parameters, some information can be encoded in the distribution for ζ^n. If J is large enough, "all" possible values of the parameters will be visited, and the weight ω^n_j will select the particles with the most suitable parameters. As developed in Section 3.4, a resampling process is often necessary for tracking the correct values. Finally, the set $\{\Theta^n_j, \omega^n_j\}_{j=1}^J$ provides a distribution of Θ^n and its mean value is given by $\sum_{j=1}^J \Theta^n_j \omega^n_j$.

For the Kalman filter or for its extensions, we consider the parameter Θ as a new random variable Θ^n with values in $\mathbb{R}^d$. If the correlation between the observations and the parameters is strong enough (this can usually be checked with a scatter plot) then the parameter can be estimated together with the data assimilation process. The suitable framework is the so-called "augmented state space" consisting in the system

$$\begin{cases} X^{n+1} = f(X^n, n) + \xi^n, \\ \Theta^{n+1} = \Theta^n + \zeta^n, \\ Y^n = H(X^n, n) + \varepsilon^n, \end{cases}$$

where ζ^n is a random variable following a prescribed distribution, usually of mean 0. As for particle filter, some information can be encoded in the distribution for ζ^n, depending on the system and on the knowledge for the evolution of the parameters, Even if the evolution of Θ^n is essentially trivial, it will be updated with the assimilation process. Note that ζ^n should neither be too small (in which case the evolution of Θ^n is not possible or too slow), nor too big (in which case arbitrary values will be taken at every time step). In other words, the role of ζ^n is to increase the spread of Θ^n as n increases. Note that this approach works for ensemble Kalman filter as well, an augmented state space can also be considered.

Let us finally and briefly come back to the general framework of filtering processes, as introduced in Section 3.1, and consider an additional random variable Θ^n. In this framework the analysis distribution (3.1.8) reads

$$\Pi_{X^{n,a},\Theta^{n,a}}(x^n, \theta^n) \equiv \Pi_{X^n,\Theta^n|Y^{1:n}}(x^n, \theta^n \mid y^{1:n})$$

from which one infers the marginal distribution for the parameters:

$$\Pi_{\Theta^{n,a}}(\theta^n) = \int_{\mathbb{R}^N} \Pi_{X^n,\Theta^n}(x^n, \theta^n \mid y^{1:n}) \, \mathrm{d}x^n.$$

This computation can be performed with a MCMC (Markov chain Monte Carlo) algorithm, and even more efficiently with a pseudo-marginal MCMC method, see [5, Sec. 9.2] for additional information.

Chapter 5

Conclusion

Data assimilation consists in a set of techniques for emulating models and observations. Both are necessary, there is no data assimilation without model, and reciprocally no data assimilation without data! There exist many variants of data assimilation techniques, they have been developed for different purposes and according to many physical or technical constraints. Applications in environmental sciences, atmosphere sciences, geosciences, biology, chemistry, social sciences are for example presented in [1, Ch. 8 - 12]. For understanding the material presented in these notes, and for getting used to the algorithms, it is now highly recommended to start implementing some of them and to solve some problems. Many examples and exercises are presented in the reference books provided in the bibliography.

Good luck, and thank you.

Index

Bibliography

[1] M. Asch, M. Bocquet, M. Nodet, *Data Assimilation: methods, algorithms, and applications*, Fundamentals of Algorithms 11, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2016.

[2] G. Evensen, *Data Assimilation, the ensemble Kalman filter*, Second edition. Springer-Verlag, Berlin, 2009.

[3] G. Grimmett, D. Welsh, *Probability: an introduction*, Second edition. Oxford University Press, Oxford, 2014.

[4] K. Law, A. Stuart, K. Zygalakis, *Data Assimilation, a mathematical introduction*, Texts in Applied Mathematics 62. Springer, Cham, 2015.

[5] S. Reich, C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation*, Cambridge University Press, New York, 2015.

[6] S. Roh. M. Jun, I. Szunyogh, M.G. Genton, *Multivariate localization methods for ensemble Kalman filtering*, Nonlin. Processes Geophys. 22, 723–735, 2015.

[7] S. Särkkä, *Bayesian filtering and smoothing*, Institute of Mathematical Statistics Textbooks 3. Cambridge University Press, Cambridge, 2013.

[8] C. Sun, S. Richard, T. Miyoshi, N. Tsuzu, *Analysis of COVID-19 spread in Tokyo through an agent-based model with data assimilation*, J. Clin. Med. 11. 2401, 2022,