Filtrage et Modélisation Stochastique

Loïc Maurin - *loic.maurin@meteo.fr*

2 - Filtrage Kalman

Au cours précédent

Processus stochastiques

Au cours précédent

Processus stochastiques

Processus stochastique : famille de variables aléatoires $(X_t)_{t \in \mathcal{T}}$.

Temps discret : $t \in \mathbb{N}$, le processus se modélise avec des **probabilités de transition**.

Temps continu : $t \in \mathbb{R}$, le processus se modélise avec des **densités de probabilités** ou **taux de transitions** notés $W(n, m)$ pour la transition entre de l'état n vers l'état m.

Equation Maîtresse : équation différentielle qui représente l'évolution temporelle de la densité de probabilité. On a vu qu'elle s'établissait de manière analogue à un bilan physique.

Relations sur la moyenne et la variance : à partir de l'équation maîtresse, on peut déduire l'évolution temporelle de la moyenne et de la variance du processus. Pour rappel,

$$\frac{d\langle n\rangle}{dt} = \langle W^+(n) - W^-(n)\rangle, \frac{dV}{dt} = 2\langle (n - \langle n\rangle)(W^+(n) - W^-(n))\rangle + \langle W^+(n) + W^-(n)\rangle$$

Au cours précédent

Précisions sur le Mouvement Brownien (ou Processus de Wiener)

On a vu que pour une marche aléatoire, avec des sauts de ± 1 équiprobables, la coupe temporelle à un pas de temps donné $t \in \mathbb{N}$, la distribution des positions se rapproche d'une gaussienne.

On peut alors formaliser le mouvement brownien, en temps continu, par rapport à cette observation.

Définition Soit $B = (B_t)_{t \geq 0}$ une famille de variables aléatoires indexées dans le temps. On dit que B est un mouvement brownien si c'est un processus à trajectoires continues telles que

1. $\forall t \geq 0 : B_t \sim \mathcal{N}(0, t)$
2. Pour tout $0 \geq t_1 \geq ... \geq t_n$, les variables aléatoires $B_{t_1}, B_{t_2} - B_{t_1}, ..., B_{t_n} - B_{t_{n-1}}$, sont indépendantes.

Remarque La seconde propriété signifie que le mouvement brownien n'a pas de mémoire du passé.

Note : Wiener a formalisé le mouvement brownien, d'où le nom de processus de Wiener et les notations $W(n, m)$ vues précédement

Au cours précédent

Correction - Cinétique chimique

Rappels de probabilités

Lois usuelles

Rappels de probabilité

Loi uniforme : X suit une **loi uniforme** sur $[a,b] \in \mathbb{R}$ si elle admet la densité de probabilité $f : \mathbb{R} \mapsto \mathbb{R}$

$$\forall x \in \mathbb{R}, f(x) = \begin{matrix} \frac{1}{b-a} & si\, x \in [a,b] \\ 0 & sinon \end{matrix}$$

Loi exponentielle : X suit une **loi exponentielle** $\mathcal{E}(\lambda)$ ($\lambda > 0$) si elle admet la densité de probabilité f : $\mathbb{R} \mapsto \mathbb{R}$

$$\forall x \in \mathbb{R}, f(x) = \begin{matrix} \lambda e^{-\lambda x} & si\, x \geq 0 \\ 0 & sinon \end{matrix}$$

On a alors $\mathbb{E}[X] = 1/\lambda, \mathbb{V}[X] = 1/\lambda^2$

Lois usuelles

Rappels de probabilité

Loi Normale : X suit une **loi normale** $\mathcal{N}(m, \sigma^2)$, $((m, \sigma) \in \mathbb{R} \times \mathbb{R}_+)$

$$\forall x \in \mathbb{R}, f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-\frac{(x-m)^2}{2\sigma^2})$$

On a alors $\mathbb{E}[X] = m, \mathbb{V}[X] = \sigma^2$

Loi Gamma : X suit une **loi gamma** $\gamma(p, \lambda)$, $(p > 0, \lambda > 0)$ si sa densité est

$$\forall x \in \mathbb{R}, f(x) = \begin{matrix} \frac{\lambda}{\Gamma(p)}(\lambda x)^{p-1}e^{-\lambda x} & si & x \geq 0 \\ 0 & sinon & \end{matrix}$$

On a alors $\mathbb{E}[X] = p/\lambda, \mathbb{V}[X] = p/\lambda^2$

Lois usuelles

Applications à la modélisation en microphysique

En microphysique (modélisation des nuages et interactions entre hydrométéores), la distribution des gouttes par rapport à leur diamètre est donnée par une loi exponentielle (Loi de Marshall-Palmer).

Loi de Marshall-Palmer

$$N(D)dD = N_0 e^{-\lambda D} dD, où\, f(D) = e^{-\lambda D}$$

où la densité de probabilité est $f(D) = e^{-\lambda D}$.

Remarque

Une densité de loi Gamma (plus générique que la loi exponentielle) est choisie sur certains schémas microphysiques.

Lois usuelles

Applications à la modélisation en microphysique

Loi de Marshall-Palmer

$$N(D)dD = N_0 e^{-\lambda D} dD$$

On peut alors lier les moments de la distribution avec les caractéristiques physiques des gouttes :

- **Concentration de goutelettes** N_c et **moment d'ordre 0** : $N = \int_0^\infty f(D)dD$
- **Contenu en eau** q_c et **moment d'ordre 3** : $L = \frac{\pi \rho_w}{6} \int_0^\infty D^3 f(D)dD$ *(lié au volume des gouttes)*
- **Réflectivité radar** Σ_c et **moment d'ordre 6** : $R \propto \int_0^\infty D^6 f(D)dD$ *(lié à la loi de Rayleigh)*

En modélisation microphysique, "le jeu" est de choisir le nombre de moments à modéliser (**fermeture**), la discrétisation sur l'espace des diamètres (**échantillonage**), pour représenter les interactions possibles entre les gouttes (de pluie, neige, etc.).

Loi forte des Grands Nombres

Théorèmes fondamentaux

Théorème Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles **indépendantes** et **identiquement distribuées (i.i.d.)** et définies sur le même espace de probabilité $(\Omega, \mathcal{F}, \mathbf{P})$.

En posant $\mu = \mathbb{E}[X_i]$ et $\sigma^2 = \mathbb{V}(X_i) < \infty$.

La variable aléatoire $S_n = \sum_{i=1}^{n} X_i$ vérifier

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu, p.s.$$

p.s. : presque sûr = "à une infinité dénombrable de points près"

Application Méthode de Monte-Carlo : calcul d’une intégrale par des tirages de variables aléatoires.

Théorème Centrale Limite

Théorème Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires réelles **i.i.d.**

On note $\mu = \mathbb{E}[X]$ et $\sigma^2 = \mathbb{V}(X)$. On suppose $0 < \sigma^2 < \infty$.

En posant $S_n = \sum_{i=1}^{n} X_i$,

$$Y_n = \frac{S_n - n\mu}{\sigma\sqrt(n)} \rightarrow Y$$

Remarque Il s'agit d'une convergence en loi : la suite des lois Y_n tend vers Y qui est aussi une loi de probabilité.

Application Le Thèorème Centrale Limite donne des indications sur les vitesses de convergence vers la loi assymptotique.

Théorèmes fondamentaux

LGN vs TCL

- Loi des Grand Nombres : convergence des moyennes empiriques vers la moyenne de la loi des v.a.
- Théorème Centrale Limite : convergence asymptotique de la loi sur la moyenne empirique (donne des indications sur une loi limite).

Modes de convergences

Rappels et théorèmes fondamentaux

Modes de convergence

- Convergence presque sûre (**p.s.**) : $P(\{\omega \in \Omega : \lim_{n\to\infty\}} X_n(\omega) = X(\omega)) = 1$
- Convergence dans L^P : $\lim_{n\to\infty} \mathbb{E}[|X_n - X|^p] = 0$
- Convergence en probabilité **P**: $\lim_{n\to\infty} P(|X_n - X| > \epsilon) = 0$

Convergence $\mathrm{L}^p \Rightarrow$ Convergence **P**, Convergence **p.s.** $\Rightarrow$ Convergence **P**

Filtre Kalman

Introduction

Filtre de Kalman

Filtre de Kalman (1960), Rudolf Kalman : utilisé pour la première fois pour l'estimation de trajectoire des programmes Apollo

$\rightarrow$ Comment concilier au mieux l'information disponible (capteurs), et les équations de la dynamique pour contrôler un système ?

Applications Contrôle, Calage GPS, Filtrage de données.

Système dynamique

Filtre de Kalman

Système dynamique discret

$$x_{n+1} = Ax_n + Bu_n + \phi_n$$

$$y_{n+1} = Cx_{n+1} + \psi_{n+1}$$

où, à chaque instant n:

- x_n est l'état du système. *Exemple : Température d'un moteur d'avion*.
- u_n est la commande du système. *Exemple : Débit de carburant dans le moteur*.
- y_n est la mesure de l'état du système. *Exemple : Mesure renvoyée par le thermomètre*.

Remarque En météo, on appellerait C un **opérateur d'observation**. C'est le lien entre **l'espace des mesures** (ex : la tension au bornes thermomètre), et **l'espace d'état** (ex : la température effective mesurée).

Bruits associés au système

Filtre de Kalman

Système dynamique discret

$$x_{n+1} = Ax_n + Bu_n + \phi_n$$

$$y_{n+1} = Cx_{n+1} + \psi_{n+1}$$

Bruits ϕ_n, ψ_n: on modélise un bruit d'état ϕ_n un bruit de mesure ψ_n.

1. **Bruit d'état ϕ_n**: représente notre méconnaissance de la physique du système. *Exemple : Variations de la richesse du mélange.*
2. **Bruit de mesure ψ_n**: représente un bruit de mesure. *Exemple: bruit électronique de la sonde de température.*

Hypothèse : Les bruits ϕ et ψ sont supposés blancs, gaussiens, centrés, stationnaires et indépendants l'un de l'autre. Ces bruit sont chacun associé à une matrice de covariance Φ, et Ψ.

Construction d’un estimateur

Filtre de Kalman

On cherche à construire un estimateur qui dépende de l’état estimé à l’instant précédent, de la mesure renoyée par le capteur et de la comande imposée.

Nous construisons alors un estimateur de la forme :

$$\hat{x}_{n+1} = A_f \hat{x}_n + B_f u_n + K_{n+1} y_{n+1}$$

$\rightarrow$ Comment construire A_f, B_f, K_{n+1} ?

Note : On cherche à construire un estimateur de la forme générique $\hat{x}_{n+1} = f(\hat{x}_n, y_{n+1}, u_n)$

Nous nous concentrons sur des systèmes linéaires, et verrons plus tard comment l’étendre à des systèmes non-linéaires.

Estimateur (assymptotiquement) sans biais

Erreur d'estimation : On cherche à minimiser l'erreur d'estimation, définie ci-dessous,

$$e = \hat{x} - x$$

Pour un estimateur parfait, on aurait $\forall n \in \mathbb{N}, e_n = 0$.

Sans accès aux vraies valeurs de x et face à un processus stochastique, annuler l'erreur est généralement impossible. Nous décrivons alors (non sans mal) l'erreur e_n par récurrence:

$$e_{n+1} = (I - K_{n+1}C)Ae_n + (A_f + K_{n+1}CA - A)\hat{x}_n + (B_f + K_{n+1}CB - B)u_n + (K_{n+1}C - I)\phi_n + K_{n+1}\psi_{n+1}$$

Estimateur (assymptotiquement) sans biais

Filtre de Kalman

Espérance de l'erreur Sachant que les bruits blancs gaussiens sont d'espérance nulle :

$\forall n \in \mathbb{N},\ \mathbb{E}[\phi_n] = 0, \mathbb{E}[\psi_n]$, on décrit l'évolution, l'espérance de l'erreur.

$$\mathbb{E}[e_{n+1}] = (I - K_{n+1}C)A\,\mathbb{E}[e_n] + (A_f + K_{n+1}CA - A)\,\hat{x}_n + (B_f + K_{n+1}CB - B)u_n$$

Estimateur assymptotiquement sans biais

On cherche à construire un estimateur dont l'espérance de l'erreur tend vers 0.

Définition

$$\lim_{n \to \infty} \mathbb{E}[e_n] = 0$$

Estimateur (assymptotiquement) sans biais

Filtre de Kalman

Espérance de l'erreur Sachant que les bruits blancs gaussiens sont d'espérance nulle :

$\forall n \in \mathbb{N},\ \mathbb{E}[\phi_n] = 0, \mathbb{E}[\psi_n]$, on décrit l'évolution, l'espérance de l'erreur.

$$\mathbb{E}[e_{n+1}] = (I - K_{n+1}C)A\,\mathbb{E}[e_n] + (A_f + K_{n+1}CA - A)\,\hat{x}_n + (B_f + K_{n+1}CB - B)u_n$$

Estimateur assymptotiquement sans biais $\rightarrow \lim_{n\to\infty} \mathbb{E}[e_n] = 0$

Conditions suffisantes

1. $B_f + K_{n+1}\,CB - B = 0$ (*on annule le terme lié à la comande* u_n)
2. $A_f + K_{n+1}\,CA - A = 0$ (*on annule le terme dû à l'estimation* $\hat{x}_n$)
3. $(I - K_{n+1}C)$ est stable (*module inférieur à 1 pour que l'erreur puisse décroître*)

Construction des Matrices du filtre

Filtre de Kalman

Les conditions de stabilité fixent les matrices A_f et B_f.

$$A_f = (I - K_{n+1}C)A$$

$$B_f = (I - K_{n+1}C)B$$

Il reste à régler K_{n+1} pour que $(I - K_{n+1}C)$ soit stable.

Définition : K est appelé le **gain du filtre**.

Forme de l'estimateur : Nous établissons (et rencontrons souvent) la forme suivante du filtre.

$$\hat{x}_{n+1} = A\hat{x}_n + Bu_n + K_{n+1}[y_{n+1} - C(A\hat{x}_n + Bu_n)]$$

Remarque K peut être vu comme un compromis à régler entre la fidélité au modèle numérique $(A\hat{x}_n + Bu_n)$ et la fidélité aux valeurs de mesure (y_{n+1}).

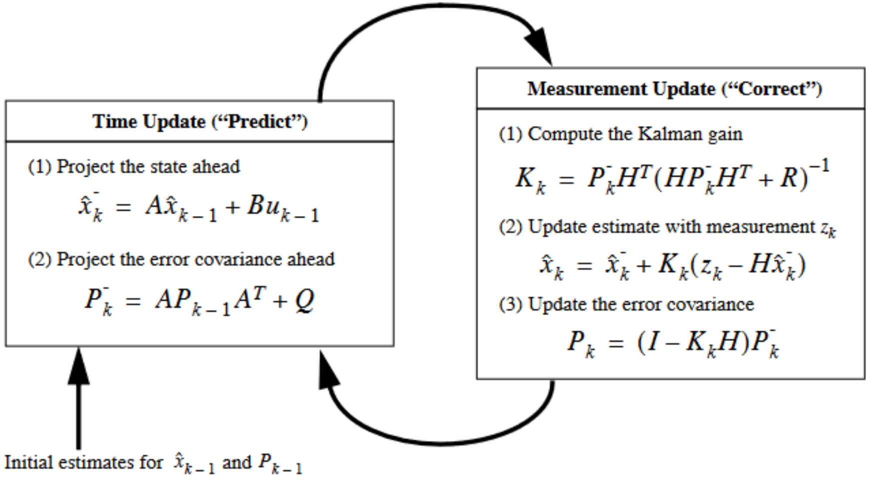
Filtre Prédicteur - Correcteur

Filtre de Kalman - Implémentation

Le filtre de Kalman est un "prédicteur-correcteur", l'estimation $\hat{x}$ de x se contruit en 2 temps :

1. **Prédiction** : Estimation à priori de l'état $\hat{x}^-$, comme si on n'avait que les équations du système à disposition.

2. **Correction** : Construction de l'estimation à posteriori avec l'information apportée par les mesures.



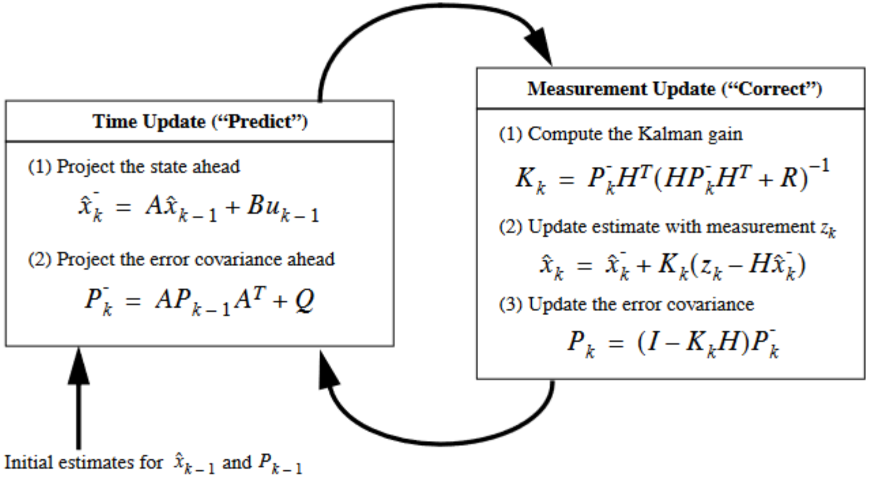
Filtre Prédicteur - Correcteur

Filtre de Kalman - Implémentation

Sans oublier la construction des matrices de covariances, liées au bruit.

1. **Prédiction** : Estimation de la matrice de covariance P_k^- , par rapport au **bruit d'état** Q (ou Φ comme noté précédemment).

2. **Correction** : Construction de la matrice de covariance P_k par rapport au **bruit de mesure** R (ou Ψ comme noté précédemment).



Filtre de Kalman - Implémentation

La recette de cuisine !

1. On initialise $\hat{x}$ à $\hat{x}_0$: on peut par exemple prendre la valeur y_0 renvoyée par le capteur.
2. On initialise P à P_0 : on peut prendre la valeur Ψ de la covariance de bruit du capteur.
3. On fait évoluer K selon :

$$K_{n+1} = (AP_nA^T + \Psi)C^T \times (CAP_nA^TC^T + C\Psi C^T + \Psi)^{-1}$$

4. On fait évoluer $\hat{x}$ selon :

$$\hat{x}_{n+1} = A\hat{x}_n + Bu_n + K_{n+1}[y_{n+1} - C(A\hat{x}_n + Bu_n)]$$

5. On fait évoluer P selon :

$$P_{n+1} = (I - K_{n+1}C)(APA^T + \Phi)$$

Est-ce que ça marche vraiment ?

Considérations sur l'évolution de la variance

On a établit que $\lim_{n \to \infty} \mathbb{E}[e_n] = 0$, ainsi que la formule pour P_{n+1}.

$\rightarrow$ Est-ce qu'on est obligé d'attendre l'∞ pour que le filtre commence à fonctionner ?

$\rightarrow$ Pourquoi on se balade avec une matrice de covariance P_{n+1} (si ce n'est par pure beauté mathématique) ?

$\rightarrow$ En pratique, comment fixer K_{n+1} le gain du filtre ?

Dispersion Nous allons travailler sur la dynamique de l'erreur e_n et sa dispersion, pour

$$P = \mathbb{E}[e_n \times e_n^T]$$

Dans un cas simple à une dimension, $P = \mathbb{E}[(\hat{x}_n - x_n)^2] = \mathbb{V}[e_n]$, *s'écrit bien comme la variance de l'erreur.*

Dynamique de l'erreur

Choix du gain K

Condition d'optimalité du gain du filtre On cherche K de telle sorte que la variance de l'erreur $P = \mathbb{E}[e \times e^T] \sim \mathbb{E}[e]$ soit minimale.

Un condition nécessaire est de chercher le gain K tel que $P(k)$ soit extrêmal, autrement dit que

$$\forall n \in \mathbb{N}, \frac{\partial P_{n+1}}{\partial K_{n+1}} = 0$$

Dynamique de l'erreur En maniuplant les équations du filtre, on obtient

$$e_{n+1} = (I - K_{n+1}C)Ae_n + (K_{n+1}C - I)\phi_n + K_{n+1}\psi_{n+1}$$

$$\frac{\partial e_{n+1}}{\partial K_{n+1}} = -CAe_n + C\phi_n + \psi_n$$

Dynamique de l'erreur

Choix du gain K

On cherche le gain K pour établir $\forall n \in \mathbb{N}, \frac{\partial P_{n+1}}{\partial K_{n+1}} = 0$

Condition suffisante

$$\mathbb{E}[\frac{\partial e_{n+1}}{\partial K_{n+1}} e_{n+1}^T] = 0$$

Formule de K à partir de P

$$P_{n+1} = (AP_n n A^T + \Psi)C^T \times (CAP_n A^T C^T + C\Phi C^T + \Psi)^{-1}$$

Formule de P à partir de K (en prenant en compte $P_0 = \Psi$)

$$P_{n+1} = (I - K_{n+1}C)(AP_n A^T + \Phi)$$

Remarque On ne peut pas obtenir de condition d'optimalité, mais simplement d'une relation de récurrence entre P et K. C'est elle qui nous permet d'implémenter K en pratique.

Synthèse

Synthèse et lien vers les processus stochastiques

Estimateur sans biais

On cherche à obtenir un estimateur qui satisfait $\lim_{n \to \infty} e_n = 0$ (assymptotiquement sans biais)

Gain optimal

On cherche à régler K pour que les variances d'erreurs soit les plus faibles à chaque pas de temps, c'est-à-dire $\frac{\partial P_{n+1}}{\partial K_{n+1}} = 0$

Récurrence

Dans la mesure où nous ne connaissons pas les valeurs vraies x_n, nous tirons partie des relations de récurrence sur l'erreur e_n pour progesser vers une erreur nulle.

C'est cela même qui fait la structure du Filtre Kalman. Et c'est bien pratique dans la mesure où le filtre ne dépend que des valeurs à l'état n pour estimer l'état $n+1$ (le filtre est robuste et facile à mettre en oeuvre).

Synthèse

Structure probabiliste

$\rightarrow$ Quel lien avec les probabilités ?

Point de départ

- Bruit d'état $\phi \sim \mathcal{N}(0, \Phi)$ (ex : Φ donné par la distibution de températures au point de mesure, i.e. la climatologie du lieu)
- Bruit de mesure $\psi \sim \mathcal{N}(0, \Psi)$ (ex : Ψ, précision donnée par la fiche technique du capteur de température)

Synthèse

Structure probabiliste

$\rightarrow$ Quel lien avec les probabilités ?

Point de départ

- Bruit d'état $\phi \sim \mathcal{N}(0, \Phi)$ (ex : Φ donné par la distibution de températures au point de mesure, i.e. la climatologie du lieu)
- Bruit de mesure $\psi \sim \mathcal{N}(0, \Psi)$ (ex : Ψ, précision donnée par la fiche technique du capteur de température)

Point d'arrivée

- On modélise $P(x_n|z_n) \sim \mathcal{N}(\hat{x}_n, P_n)$, avec notre estimateur $\hat{x}_n$
- Avec les propriétés de l'estimateur, on progresse vers $\mathbb{E}[\hat{x}_n] = x_k$ en gardant une dispersion minimale $\mathbb{E}[(x_n - \hat{x}_n)(x_n - \hat{x_n})^T] = P_n$.

Synthèse pratique

Construction pratique d'un filtre de Kalman

Hypothèses structurantes Il reste en pratique à vérifier que les bruits de mesure et d'état sont effectivement des bruits blancs gaussiens, stationnaires.

1. Qu'on a "suffisament capté" la dynamique du système et son observation dans les matrices A, B et C, pour que les bruits Φ et Ψ soient effectivement centrés.

2. Que les bruits puissent être assimilés à des bruits blancs gaussiens. Cela implique des tests statistiques, et éventuellement un travail sur les équations pour centrer et réduire le bruit (lien avec le TCL).

3. Qu'il y ait effectivement indépendance entre le bruit de mesure et le bruit d'état. Exemple : le capteur de température n'influence pas son environnement en ralentissant le flux d'air dont il mesure la température.

En pratique

Il n'y a pas forcément de réponse systématique à ces questions, simplement un travail de mise au point du filtre, sur un problème donné $\rightarrow$ **Coeur du travail de l'ingénieur**

Exemple - Estimation d’une constante aléatoire

Filtre de Kalman

Extended Kalman Filter

Si la dynamique est non linéaire ?

Remplaçons :

$$x_{n+1} = Ax_n + Bu_n + \phi_n$$

$$y_{n+1} = Cx_{n+1} + \psi_{n+1}$$

Par :

$$x_{n+1} = f(x_n, u_n, \phi_n),\ (mod\grave{e}le\ non-lin\acute{e}aire)$$

$$y_{n+1} = h(x_{n+1}, \psi_{n+1}), (observateur\ non-lin\acute{e}aire)$$

Remarque C'est le cas en Météo où le modèle repose sur les équations de Navier-Stokes (non-linéaires), et les observateurs reposent sur les lois non-linéaire (ex : reflectivité radar $R \propto D^6$)

Extended Kalman Filter

Si la dynamique est non-linéaire ?

Solution : Linéariser les équations (modèle et observateurs), autour du point de fonctionnement $\hat{x}_n$ estimé.

Outil : les matrices jacobiennes des opérateurs $y = h(x)$ et modèles $x_{n+1} = f(x_n)$

On note alors

En Météo On ne s'étonnera pas de trouver les termes de **Tangent-Linéaire** (TL) et **Adjoint** en assimilation, il s'agit resp. de la **jacobienne** et de sa **transposée** *(exactement du conjugué de sa transposée si on travail sur un espace complexe)*

En pratique La encore, le **coeur du travail d'ingénieur** est d'obtenir les opérateurs adéquats (en respectant de manière empirique les hypothèses sur les distributions).

A retenir

Qu'est-ce qu'un filtre de Kalman déjà ?

Basique

- La recette et le fonctionnement de prédicteur-correcteur pour le filtre Kalman,
- (Théorie), c'est un estimateur assymptotiquement sans biais, et avec une variance à minimiser.

En fonction de l'espace de stockage disponible

- La démarche récursive de construction du filtre, *utile pour comprendre le filtre "pas-à-pas"*.

En pratique, avec l'algorithme à disposition

- Comment linéariser le modèle et les observateurs de mesure ?
- Comment respecter au mieux les hypothèses de bruit blanc gaussien ?

En gardant à l'esprit qu'il s'agit d'un mélange de pratique et de théorie, sur la base du Filtre de Kalman, qui est plutôt un filtre simple et robuste

La semaine prochaine

Spoiler alert !

TP : Filtre Kalman sur un oscillateur harmonique

$\rightarrow$ Apportez vos PC, TP sur Google Colab

END