Filtrage et Modélisation Stochastique

Loïc Maurin - *loic.maurin@meteo.fr*

Objectif du cours

Introduction au filtrage et à la modélisation stochastique

Objectif : construire une boîte à outils pour modéliser l'incertitude et les phénomènes aléatoires

- Applications :
  - Assimiler des données bruitées (ex : capteurs de vitesse, température, etc.)
  - Modéliser des phénomènes chaotiques (ex : dynamique atmosphérique)
  - Contrôler et optimiser des systèmes dont la dynamique est inconnue (boîte noire)
- En Météo :
  - Estimer les variances d'ébauche, à l'assimilation de données d'observations,
  - Construction des ensemblistes

Objectifs du cours

Démystifions l'aspect stochastique !

- Dans ce cours, nous nous intéresserons à modéliser la moyenne et la (co)-variance des processus
    - *en sachant que les méthodes peuvent être étendues aux moments d'ordres supérieurs*
- Ici, Stochastique = Gaussien (dans la mesure du possible)
    - *on modélisera les processus par rapport à la loi normale, en tirant partie du Théorème Centrale Limite,*
    - *en gardant à l'esprit d'autres lois peuvent supporter des modèles (ex: processus de poisson)*

Plan du cours

- Introduction
- Filtre de Kalman
  - TP : Application au filtrage d'un oscillateur harmonique
- Chaînes et processus de Markov
  - TP : Markov Chains Monte-Carlo avec Page-Rank
- (Un peu de) théorie sur les processus stochastiques et Ensemble Kalman Filter
  - TP : Assimilation de mesures de température et de salinité *in situ* pour un modèle d'océan
- Régression par Processus Gaussiens et Optimisation Bayésienne
  - TP : modèle de courant de surface par rapport aux données de bouées
  - TP : optimisation des hyper-paramètres d'un réseau de neurone
- Devoir

Processus Stochastiques - Applications

Processus stochastiques - Applications

Notions abordées :

- Marche aléatoire et Mouvement brownien
- Régression par processus gaussiens
- Processus de Markov : Transition d'états sans mémoire

Modèles et techniques développés :

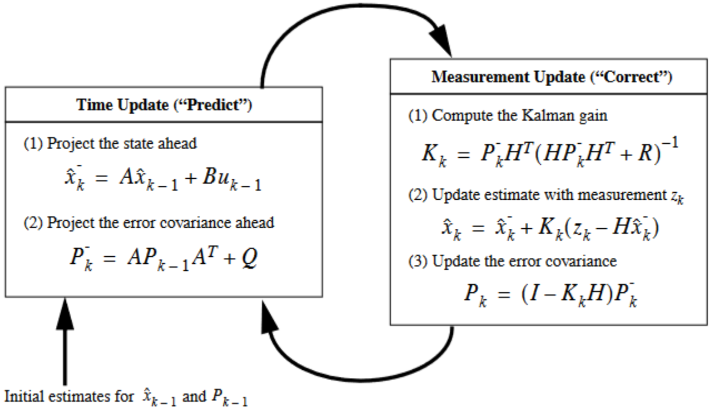
- Optimisation Bayésienne : échantillonage sur un processus gaussien
- Echantillonage par méthodes Monte-Carlo : Markov Chain Monte-Carlo
- Filtre Kalman : cas particulier de processus gaussien

Filtrage Kalman

Quelques exemples d'applications

Utiliser au mieux la dynamique d'un système pour recaler ses observations

- Localisation / Recalage GPS
- Filtrage de données robotiques
- Assimilation de données pour la Météo (Ensemble Kalman Filter - EnKF)



Markov Chain Monte-Carlo (MCMC)

Quelques exemples d'applications

Echantillonner et estimer les probabilités de transitions entre états d'un système

- Recherche de pages web (Page Rank)
  - Marche aléatoire sur le graphe des pages webs
- Estimation des récurrences et probabilités d'évènements extrêmes dans un modèle de climat
  - Marche aléatoire sur des séries temporelles de pluies : comparaison de données historiques et projetées par le modèle

(Sedlmeier, K., Mieruch, S., Schädler, G., and Kottmeier, C.: Compound extremes in a changing climate – a Markov chain approach, Nonlin. Processes Geophys., 23, 375–390, https://doi.org/10.5194/npg-23-375-2016, 2016.)

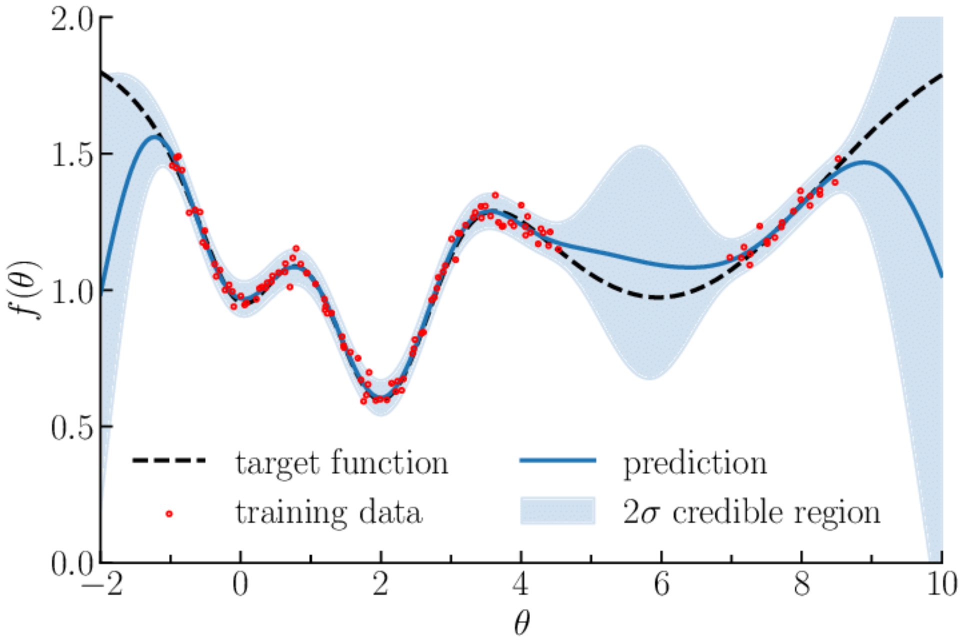
Régression par Processus Gaussiens

Quelques exemples d'applications

Modéliser les incertitudes lors de l'exploration d'un champ / d'un espace de paramètres

- Résolution de problèmes inverses :
  - Contrôle de bras robotiques,
  - Cartographie des sols : recherche pétrolière, géophysique
- Optimisation
  - Tuning des hyperparamètres d'un réseau de neurone
  - Réglage de procédés industriels par plans d'expérience

Régression par Processus Gaussiens



Définitions

Rappels de Probabilités et Statistiques

Pour être à l'aise sur les notions du cours

- Loi Normale
- Loi jointe, Loi marginale
- Théorème Centrale Limite et Loi des Grand Nombres
- Théorème de Bayes
- Estimateur du Maximum Vraisemblance

Processus stochastiques

Définition

On considère un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, un espace mesurable $(E, \mathcal{E})$ et un ensemble T.

Définition On appel processus stochastique, ou processus aléatoire, une famille $(X_t)_{t\in\mathcal{T}}$ de variables aléatoires à valeurs dans E.

Autrement dit, pour tout $t \in \mathcal{T}$, l'application $\omega \mapsto X_t(\omega)$ est une application mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{E})$. On appelle E l'espace d'état du processus.

Remarque

- X_t est souvent l'expression d'une variable à un temps t,
- $\mathcal{T}$ représente l'ensemble des dates possibles

Processus stochastiques

Définition

- Lorsque $\mathcal{T} = \mathbb{N}$ ou $\mathcal{T} = \mathbb{Z}$, on parle de **processus à temps discret**
- Lorsque $\mathcal{T} = \mathbb{R}$, on parle de **processus à temps continu**

Exemple

- Processus à temps discret : évolution du PIB de la France par année où :

  X_t représente le PIB, $t \in \mathbb{N}$ est l'année

- Processus à temps continue : évolution du cours d'une action où :

  X_t est la valeur de l'action, $t \in \mathbb{R}_+$ est le temps, considéré comme continu étant donnée la fréquence de rafraîchissement des valeurs.

Processus stochastiques

Définition

Définition On appelle **filtration** une suite $(\mathcal{F}_t)_{t\in\mathcal{T}}$ σ-algèbres vérifiant

$$s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$$

Remarque On utilise la notion de filtration pour représenter l'information disponible à date t.

Quand on observe un processus au cours du temps, on connaîot les valeurs de X_s pour $s \leq t$, mais pas encore pour $s > t$. On sera donc souvent amenés à conditionner par les variables $(X_s)_{s\leq t}$.

Définition Le processus $(X_t)_{t\in\mathcal{T}}$ est dit adapté à la filtration $(\mathcal{F}_t)_{t\in\mathcal{T}}$, si pour tout $t \in \mathcal{T}$, X_t est $\mathcal{F}_t$-mesurable.

Processus stochastiques

Exemple

Processus Auto-Régressifs (AR) Dans cet exemple, $\mathcal{T} = \mathbb{N}$. Soit $(\epsilon_t)_{t\in\mathbb{N}}$ une suite de variables aléatoires i.i.d de loi $\mathcal{N}(0, 1)$, $(\alpha, \beta) \in \mathbb{R}^2$, $X_0 = 0$ et

$$X_{t+1} = \alpha X_t + \beta + \epsilon_t$$

On définit $\mathcal{F}_t = \sigma(\epsilon_s, s \leq t)$. On peut vérifier que le processus $(X_t)_{t\in\mathbb{N}}$ est adapté à la filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$.

Dans ce cours

On étudiera :

- Les **processus de Markov**, dont les lois de transitions ne dépendent que de l'état précédent,

$$\mathcal{P}(X_{t+1}|\mathcal{F}_t) = \mathcal{P}(X_{t+1}|X_t)$$

Dans ce cours

On étudiera :

- Les **processus gaussiens** $(X_t)_{t\in\mathcal{T}}$, pour lesquels toutes les lois fini-dimensionnelles $\mathcal{L}(X_{t_1}, .., X_{t_n})$ sont gaussiennes.

On définit les processus gaussiens par une moyenne au cours du temps $m(t) = \mathbb{E}[X_t]$ et un opérateur de covariance $K(s,t) = Cov(X_s, X_t)$

Dans ce cours

On étudiera :

- Le **filtre Kalman**, un exemple particulier de processus gaussiens, pour lequel l'état d'un système dynamique est estimé, conditionné à des observations :

$$\hat{X}_k = \mathbb{E}[X_k]$$

$$\mathrm{P}_k = \mathbb{E}[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T]$$

$$p(X_k | Z_k) \sim \mathcal{N}(\hat{X}_k, \mathrm{P}_k)$$

Dans ce cours

On étudiera :

- Les **processus de Markov**, dont les transitions ne dépendent que de l'état précédent,

$$\mathcal{P}(X_{t+1}|\mathcal{F}_t) = \mathcal{P}(X_{t+1}|X_t)$$

- Les **processus gaussiens** $(X_t)_{t\in\mathcal{T}}$, pour lesquels toutes les lois fini-dimensionnelles $\mathcal{L}(X_{t_1}, .., X_{t_n})$ sont gaussiennes. On définit les processus gaussiens par une moyenne au cours du temps $m(t) = \mathbb{E}[X_t]$ et un opérateur de covariance $K(s,t) = Cov(X_s, X_t)$

- Le **filtre Kalman**, un exemple particulier de processus gaussiens, pour lequel l'état d'un système dynamique est estimé, conditionné à des observations :

$$\hat{X}_k = \mathbb{E}[X_k]$$

$$\mathrm{P}_k = \mathbb{E}[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^T]$$

$$p(X_k|Z_k) \sim \mathcal{N}(\hat{X}_k, \mathrm{P}_k)$$

Processus stochastique - Théorie

Comment décrire les transitions entre états, y compris en temps continu ?

Mouvement Brownien discret 1D

Vers une description formelle des transitions entre états

On prend un monde divisé en cases (numérotées chacune par un indice n). A chaque pas de temps, la particule fait un saut sur une case immédiatement à sa droite ou à sa gauche.

- Quelle est la probabilité $P(n, t)$ de trouver la particule dans la case n, au temps t ?

(sachant que la particule début à la case $n = 0$ à $t = 0$)

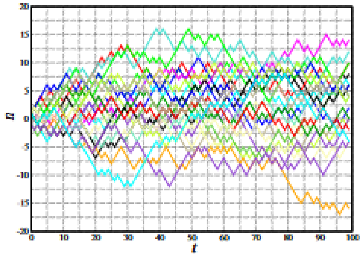


FIGURE 3.1 – 25 réalisations du mouvement brownien discret.

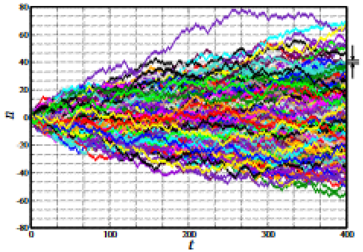


FIGURE 3.2 – 500 réalisations du mouvement brownien discret, pendant 400 tours. La proportion de trajectoires qui aboutissent à $n = 40$ au temps $t = 400$ est la probabilité $P(40, 400; 0, 0)$.

Mouvement Brownien discret 1D

Vers une description formelle des transitions entre états

On cherche à décrire $P(n, t_i)$, pour un temps t_i donné, ou autrement dit, on cherche à faire une coupe temporelle du mouvement brownien au temps t_i.

Pour cela, nous nous intéressons à extraire les moments statistiques, ici, la moyenne et la variance du processus en fonction du temps.

Ici, on remarque que (figure de droite), **la moyenne est nulle** à chaque pas de temps, et **la variance évolue linéairement** en fonction du temps.

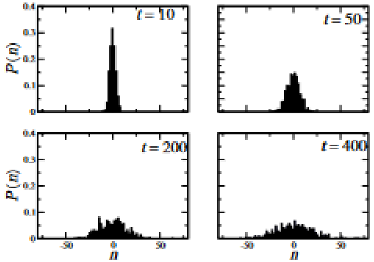


FIGURE 3.3 – Les probabilités $P(n, t_i)$ pour quatre temps $t_i = 10, 50, 200, 400$ obtenu à partir des trajectoire de la figure 3.2.

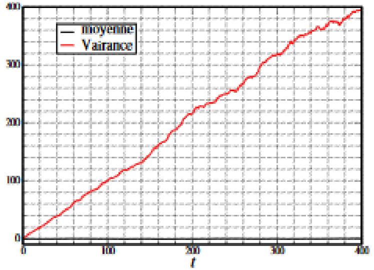


FIGURE 3.4 – Moyenne $\langle n(t) \rangle$ et variance $V(t)$ extraite des coupes temporelles des trajectoires de la figure 3.2.

Mouvement Brownien discret 1D

Vers une description formelle des transitions entre états

Démonstration Soit D la variable aléatoire de déplacement à chaque tour. D prend les valeurs ± 1 avec une probabilité de $0,5$.

Sa moyenne est $\mu = \langle D \rangle = 0$ et sa variance $\sigma^2 = \langle (D - \mu)^2 \rangle$.

La position de la particule est une variable aléatoire X qui s'exprime au bout de T pas de temps comme la somme de T déplacements D indépendants.

On obtient :

$$\langle X(T) \rangle = \langle \sum_{t=0}^{T} D_t \rangle = \sum_{t=0}^{T} \langle D_t \rangle = 0$$

$$Var(X(T)) = Var(\sum_{t=0}^{T} D_t) = \sum_{t=0}^{T} Var(D_t) = T\sigma^2$$

Mouvement Brownien - Généralisation

Vers une description formelle des transitions entre états

Généralisation à d'autres lois Le résultat précédent se généralise aisément à d'autres lois de probabilité de transition, mais le résultat est le même : la variance croit linéairement avec le temps.

- Quelle modélisation quand le pas de temps δt tend vers 0 ?

Généralisation au temps continu On décrit les probabilités de saut comme une densité qui dépend du pas de temps : $W(n)\delta t$.

On note :

- $W^-(n)$: **densité de probabilité** ou **taux de transition** de saut vers la gauche à partir de la case n.
- $W^+(n)$: **densité de probabilité** ou **taux de transition** de saut vers la droite.

Cette description peut être étendue à des sauts de plusieurs cases, en considérant $W(n, m)$: densité de probabilité de sauter de n à m. Et, éventuellement, à un espace d'états continus *(non traité formellement dans le cours)*.

Mouvement Brownien - Temps continu

Vers l'équation Maîtresse

En reprenant les sauts de ± 1 case, nous pouvons effectuer un bilan infinitésimal (analogue aux bilans de physique).

- Probabilité d'aboutir à la case n à $t + \delta t$?

3 possibilités :

1. La particule est en $n+1$ à t et saute vers n, avec une probabilité de $W^{-}(n+1)dt$
2. La particule est en $n-1$ à t et saute vers n, avec une probabilité de $W^{+}(n-1)dt$
3. La particule reste en n avec une probabilité de $1 - (W^{+}(n) + W^{-}(n))$

Mouvement Brownien - Temps continu

Vers l'équation Maîtresse

En reprenant les sauts de ± 1 case, nous pouvons effectuer un bilan infinitésimal (analogue aux bilans de physique).

- Probabilité d'aboutir à la case n à $t + \delta t$?

On obtient alors (avec des évènements indépendants):

$$P(n, t+dt) = \begin{aligned}[t] &P(n+1,t)W^-(n+1)dt \\ &+P(n-1,t)W^+(n-1)dt \\ &+P(n,t)(1-(W^+(n)+W^-(n)))dt \end{aligned}$$

Mouvement Brownien - Temps continu

Vers l'équation Maîtresse

En reprenant les sauts de ± 1 case, nous pouvons effectuer un bilan infinitésimal (analogue aux bilans de physique).

- Probabilité d'aboutir à la case n à $t + \delta t$?

Ce qui nous mène à établir **l'équation Maîtresse** :

$$\frac{\partial P(n,t)}{\partial t} = W^+(n-1)P(n-1,t) - W^-(n)P(n,t) + W^-(n+1,t)P(n+1,t) - W^+(n)P(n,t)$$

Nous obtenons en quelque sorte un **bilan de flux de probabilités** entre cellules adjacentes, telle qu'elle pourrait être établie en physique statistique, avec un flux sur la face gauche de la cellule, et un flux sur la face droite.

Equation Maîtresse - Temps continu

Généralisation et notations

On peut ainsi généraliser l'équation précédente à des transitions entre états m, n quelconques.

$$\frac{\partial P(n,t)}{\partial t} = \sum_m W(m \to n)P(m,t) - W(n \to m)P(n,t)$$

Formellement, on écrira :

$$\partial_t \left|P(t)\right\rangle = \mathcal{L} \left|P(t)\right\rangle$$

avec $\mathcal{L} = \mathcal{L}_m^n = W(m \to n)$ la matrice de transition entre états n, m et $\left|P(t)\right\rangle$ le vecteur colonne des cellules du système.

Remarque *On arrive ici au bout du travail de formalisation. L'exemple bilan sur un mouvement brownien avec des sauts élémentaires* ± 1 *pouvant être généralisé à d'autres problèmes avec des transitions plus complexes.*

Equation Maîtresse - Temps continu

Moyenne et Variance

Après insertion des moments d'ordre 1, $\langle n(t)\rangle = \sum_n nP(n,t)$,et d'ordre 2, $\langle n^2(t)\rangle = \sum_n n^2 P(n,t)$, dans l'équation maîtresse, nous obtenons,

Moyenne :

$$\frac{d\,\langle n(t)\rangle}{dt} = \langle W^+(n) - W^-(n)\rangle = \sum_n (W^+(n) - W^-(n))P(n,t)$$

Variance :

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

Remarque Ces formules (un peu ingrates), nous permettent d'extraire une équation différentielle pour chacun des moments de la distribution, et obtenir ainsi la moyenne et la variance du processus au cours du temps.

Processus de Poisson

Exemple de processus stochastique

Le processus de Poisson est fondamental pour comprendre les processus stochastiques. On le rencontre sur plusieurs phénomènes :

- Nombre de particules radioactives qui se désintègrent pendant un temps t,
- Nombre de communications reçues par un central sur une période ΔT,
- Réaction chimique.

Exemple d'une particule radioactive :

On note αdt la probabilité de désintégration d'une particule sur un temps dt

- Quelle est la probabilité d'observer n évènements pendant un temps t ?

Processus de Poisson

Exemple d'une particule radioactive :

- Quelle est la probabilité d'observer n évènements pendant un temps t ?
- On note : $W^+(n) = \alpha,$ et $W^-(n) = 0,$ (la probabilité de "ré-intégration", ou de l'évènement contraire est nulle).

Selon l'équation Maîtresse :

$$\frac{\partial P(n,t)}{\partial t} = \alpha(P(n-1,t) - P(n,t)), n \geq 1$$

$$\frac{\partial P(n,t)}{\partial t} = \alpha P(0,t)$$

Sa résolution donne :

$$P(n,t) = e^{-\alpha t}\frac{(\alpha t)^n}{n!}$$

Processus de Poisson

Exemple d'une particule radioactive :

- Quelle est la probabilité d'observer n évènements pendant un temps t ?
- Quelles sont la moyenne et la variance du processus ?

Selon l'équation Maîtresse :

$\frac{\partial\langle n(t)\rangle}{\partial t} = \sum_n (W^+(n) - W^-(n))P(n;t) = \alpha$, avec $n(t=0) = 0$

donc $\langle n(t)\rangle = \alpha t$

et $\frac{\partial\langle n^2(t)\rangle}{\partial t} = 2\langle \alpha n\rangle + \langle\alpha\rangle = 2\alpha\langle n\rangle + \alpha$ avec variance nulle à l'origine $\langle n^2(0)\rangle = 0$

donc $\langle n^2(t)\rangle = \alpha^2 t + \alpha t.$

On obtient la moyenne et la variance d'un processus de Poisson :

$$M = \langle n\rangle = \alpha t, V = \langle n^2\rangle - \langle n\rangle^2 = \alpha t$$

Exercices

Exercice

Cinétique chimique

Nous posons $W^+(n) = \alpha, W^-(n) = \mu n$. L'équation de cinétique chimique est donnée par

$$\frac{d\langle n\rangle}{dt} = \alpha - \mu\langle n\rangle$$

Avec $n_0 = 0$, on a $\langle n(t)\rangle = (\alpha/\mu)(1 - e^{-\mu t})$

1. Montrer que la variance vaut $V(t) = (\alpha/\mu)(1 - e^{-\mu t})$
2. Vérifier que pour l'état stationnaire, quand $\partial_t P = 0$, la solution est une distribution de Poisson de paramètre $\lambda = \alpha/\mu$

Relations utiles

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

$$\langle n\rangle = \sum_n nP(n,t), \langle n^2(t)\rangle = \sum_n n^2 P(n,t), \langle f(n)\rangle = \sum_n f(n)P(n,t)$$

Sources

Liens utiles du pour les processus stochastiques

Processus stochastiques, B. HOUCHMANDZADEH

Introduction aux processus stochastiques - Notes de cours, N. CHOPIN

A retenir

- Contruction d'un processus stochastique à partir des **taux de transition**, ou **densités de probabilités** de transition,
- Analogie avec des bilans physiques,
- Après formulation d'un processus stochastique, les grandeurs qui nous intéressent sont la moyenne du processus (qui souvent correspond à la dynamique classique), et sa variance (ce qui nous permettra de quantifier des incertitudes).

Remarques

Ce cours pose les bases sur processus stochastiques, ce qui permettra d'aller vers les modèles et applications lors des prochains cours : Filtre de Kalman, Processus Gaussiens, Chaînes de Markov.

La semaine prochaine

Spoiler de la prochaine séance

TP Filtrage Kalman : Filtre Kalman sur un oscillateur harmonique

Apportez vos PC -> TP sur Google Colab

Questions ?