

SOD333 - Rapport

Paul-Antoine Leveilley & Mila Rocco

Septembre 2022

Contents

1	TP1: Calcul d'une intégrale par méthode de Monte-Carlo, importance sampling.	3
1.1	Introduction	3
1.2	Calcul par méthode de Monte Carlo naive	3
1.2.1	Calcul préliminaire : variance de l'estimateur naif.	3
1.2.2	Simulations	4
1.2.3	Théorème central limite : illustration	4
1.3	Echantillonnage pondéré et réduction de la variance	4
1.3.1	Méthode du rejet	5
1.3.2	Simulations de $\tilde{\mu}_N$	6
1.4	Conclusion	7
2	TP2 : Filtre de Kalman: poursuite d'un mobile	8
2.1	Introduction	8
2.1.1	Cadre général : filtre de Kalman	8
2.2	Application, estimation d'une trajectoire	8
2.3	Simulations de la trajectoire et estimation par le filtre de Kalman	9
2.4	Conclusion	10
3	TP3 : Borne de Cramer Rao.	12
3.1	Introduction	12
3.1.1	Borne de Cramer Rao pour un problème de filtrage	12
3.1.2	Application : poursuite de bateau	12
3.1.3	Calcul du gradient de h_k par rapport à X_k	12
3.2	Simulations des dynamiques de l'observateur et de l'émetteur	13
3.3	Optimisation de la valeur de φ	13
3.4	Conclusion	13
4	TP4: Navigation par corrélation de terrain - filtre particulaire	16
4.1	Introduction	16
4.2	Filtre optimal et filtre particulaire	16
4.2.1	Filtre optimal	16
4.2.2	Filtre particulaire	17
4.2.3	Dynamique déterministe	18
4.3	Navigation par corrélation de Terrain	18
4.4	Conclusion	20

1 TP1: Calcul d'une intégrale par méthode de Monte-Carlo, importance sampling.

1.1 Introduction

Dans ce TP, on se propose de calculer une intégrale par méthode de Monte Carlo. De manière générale, la méthode de Monte Carlo consiste à approcher l'intégrale :

$$\mu = \int g(x)q(x)dx$$

Où q est une densité, par l'estimateur :

$$\hat{\mu}_N = \frac{1}{N} \sum_{i=1}^N g(X_i) \quad (1)$$

En effet, d'après la loi forte des grands nombres,

$$\hat{\mu}_N \xrightarrow[p.s.]{} \mu$$

Pour illustrer cette méthode, nous allons estimer l'intégrale

$$\int_0^1 \cos\left(\frac{\pi x}{2}\right) dx$$

Dont la valeur se calcule analytiquement, et vaut $\frac{2}{\pi}$

Dans un premier temps, on mettra en oeuvre la méthode dite "naive" de Monte Carlo. Des considérations de minimisation de la variance de l'estimateur nous conduiront à employer une méthode plus précise. Ce sera l'objet de la deuxième partie.

1.2 Calcul par méthode de Monte Carlo naive

Une première idée est de choisir la loi uniforme sur $[0, 1]$, qui admet pour densité $q(x) = \mathbb{1}_{x \in [0,1]}$, ainsi que $g(x) = \cos(\frac{\pi x}{2})$

1.2.1 Calcul préliminaire : variance de l'estimateur naif.

Comme expliqué précédemment, la variance de l'estimation est un indicateur clé pour nous. On cherche à ce qu'elle soit la plus petite possible, de sorte à avoir une précision donnée sur le calcul de l'intégrale, en un nombre minimum d'itérations. Employons nous donc à calculer cette variance pour (1).

$$\begin{aligned} \text{Var}(\hat{\mu}_N) &= \frac{1}{N^2} \sum_{i=1}^N \text{Var}(g(X_i)) \\ &= \frac{1}{N^2} \times N \text{Var}\left(\cos\left(\frac{\pi X}{2}\right)\right) \\ &= \frac{1}{N} \left(\int_0^1 \cos^2\left(\frac{\pi x}{2}\right) dx - \left(\int_0^1 \cos\left(\frac{\pi x}{2}\right) dx \right)^2 \right) \\ &= \frac{1}{N} \left(\int_0^1 \frac{1 + \cos(\pi x)}{2} dx - \frac{4}{\pi^2} \right) \\ &= \frac{1}{N} \left(\frac{1}{2} - \frac{4}{\pi^2} \right) \end{aligned}$$

1.2.2 Simulations

Dans l'intégralité de ce TP, on prendra $N = 50$. Pour réaliser cette estimation de μ , nous simulons en machine le tirage d'un échantillon (X_0, \dots, X_N) iid de loi uniforme sur $[0, 1]$. Ensuite, nous calculons une estimation de μ en utilisant l'estimateur (1) Nous réalisons ce calcul 10 fois, avec des échantillons différents. Le tableau (1) présente les résultat.

Figure 1: Calcul de l'estimation naive de μ pour 10 échantillons différents

0.6685	0.67890	0.69391	0.59470	0.69625	0.6424	0.64359	0.61404	0.65725	0.63069
--------	---------	---------	---------	---------	--------	---------	---------	---------	---------

Nous observons que le résultat est assez proche de la valeur exacte de l'intégrale ($\frac{2}{\pi} \approx 0.6366$). Cela dit, on constate également une forte variance des résultats de l'estimation. Nous avons cherché à estimer cette variance pour un nombre de 1000 réalisations de l'estimation $\hat{\mu}_N$. Nous obtenons une variance empirique de $Var_{emp}(\hat{\mu}_N) = 0.00184$. Ceci est conforme à ce qui est attendu. Le calcul de la partie précédente prédit en effet : $Var(\hat{\mu}_N) \approx 1,89.10^{-3}$

1.2.3 Théorème central limite : illustration

Le théorème centrale limite prédit que l'erreur normalisée doit se comporter comme une loi normale centrée réduite pour $N \rightarrow \infty$. C'est à dire :

$$\frac{\hat{\mu}_N - \mu}{\sigma/\sqrt{N}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

Pour illustrer ceci, nous estimons la distribution de l'erreur renormalisée par l'histogramme de 1000 réalisations de l'erreur renormalisée empirique. Nous présentons ce résultat sur le graphique de la figure 2, en superposant l'histogramme avec la représentation graphique de la densité de la gaussienne centrée réduite.

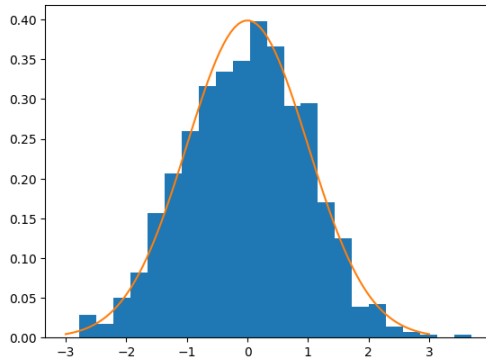


Figure 2: Distribution empirique de l'erreur renormalisée et densité de la gaussienne centrée réduite

Nous constatons que, comme le prédit le TCL, l'erreur renormalisée possède une distribution proche d'une gaussienne centrée réduite.

1.3 Echantillonnage pondéré et réduction de la variance

L'estimation de la partie précédente conduit à une grande variance du résultat. Ce n'est pas satisfaisant et l'on peut faire mieux. En effet, quand nous estimons μ avec (1), nous donnons autant d'importance à chaque réalisation de X_i dans l'intégrale. Mais certains termes ont une contribution plus importante dans le calcul. Ce sont les termes pour lesquels $g(X_i)$ est grand. Intuitivement, nous devons donc chercher à estimer "finement" les grandes valeurs de gq . Il faut donc tirer selon une densité la plus "proportionnelle" possible à gq . C'est l'idée de l'échantillonnage pondéré (préférentiel).

L'échantillonnage pondéré consiste à remplacer $\hat{\mu}_N$ par :

$$\tilde{\mu}_N = \frac{1}{N} \sum_{i=1}^N g(X_i) \frac{q(X_i)}{\tilde{q}(X_i)}$$

Où, \tilde{q} est une densité et $(X_i)_{1 \leq i \leq n}$ est distribué selon une loi de densité \tilde{q}

En effet, dans ce cas, par loi forte des grands nombres, $\tilde{\mu}_N$ converge presque sûrement vers :

$$\mathbb{E}^{\mathbb{Q}}[g(X) \frac{q(X)}{\tilde{q}(X)}] = \int_0^1 \frac{g(x)q(x)}{\tilde{q}(x)} \tilde{q}(x) dx = \int_0^1 g(x)q(x) dx = \mu$$

Où $\mathbb{E}^{\mathbb{Q}}$ désigne l'espérance sous la nouvelle probabilité \mathbb{Q} . Pour cette raison, on appelle également cette technique "changement" de probabilité.

Maintenant, puisque c'est cela que nous cherchons à optimiser, calculons la variance de ce nouvel estimateur.

$$\begin{aligned} Var(\tilde{\mu}_N) &= \frac{1}{N^2} \sum_{i=1}^N Var\left(g(X_i) \frac{q(X_i)}{\tilde{q}(X_i)}\right) \\ &= \frac{1}{N} \left(\int_0^1 \left(\frac{g(x)q(x)}{\tilde{q}(x)} \right)^2 \tilde{q}(x) dx - \left(\int_0^1 \frac{g(x)q(x)}{\tilde{q}(x)} \tilde{q}(x) dx \right)^2 \right) \\ &= \frac{1}{N} \left(\int_0^1 \frac{g^2(x)q^2(x)}{\tilde{q}(x)} dx - \left(\int_0^1 g(x)q(x) dx \right)^2 \right) \\ &= \frac{1}{N} \left(\int_0^1 \frac{g^2(x)q^2(x)}{\tilde{q}(x)} dx - \mu^2 \right) \end{aligned}$$

On voit donc qu'en faisant le choix $\tilde{q} = gq / \int g(x)q(x)dx$, on obtient une variance nulle. C'est surprenant mais en fait c'est attendu, car l'intuition du changement de probabilité est d'échantillonner notre variable aléatoire selon une densité la plus "proportionnelle" à gq . Le meilleur choix est donc sans surprise celui là. Seulement, en pratique, on ne peut pas faire ce choix, car pour échantillonner selon $gq / \int gq$, il faut être capable de calculer l'intégrale de gq or c'est exactement ce que l'on cherche et que l'on suppose difficile à calculer. Dans la pratique, on choisit donc une approximation de $gq / \int gq$, un développement limité par exemple. Maintenant une autre question se pose. Comment tirer une variable aléatoire ayant une loi à densité \tilde{q} ?

1.3.1 Méthode du rejet

On voudrait simuler une variable aléatoire X selon une densité f . On suppose qu'il existe une densité g telle que le rapport $\frac{f}{g}$ soit borné, disons, par C et qu'on sait simuler Y selon g .

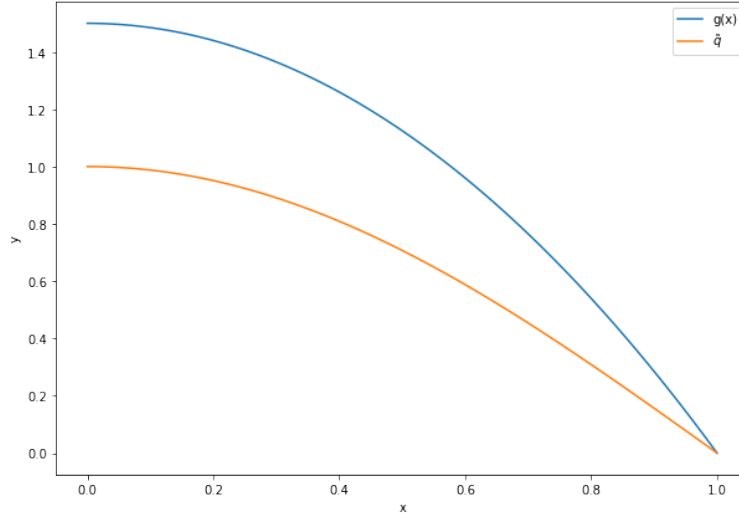
La méthode du rejet consiste à :

- tirer Y selon g et U uniforme sur $[0, 1]$, ce tant que $U > \frac{f(Y)}{Cg(Y)}$.
- Des que ce n'est plus le cas (i.e. $U \leq \frac{f(Y)}{Cg(Y)}$), on pose $Z = Y$.

On peut prouver que Z ainsi défini est une variable aléatoire à densité f . On appelle probabilité d'acceptation le nombre :

$$P_a(C) = P(U \leq \frac{f(Y)}{Cg(Y)})$$

Figure 3: Représentation graphique de g et de \tilde{q}



1.3.2 Simulations de $\tilde{\mu}_N$

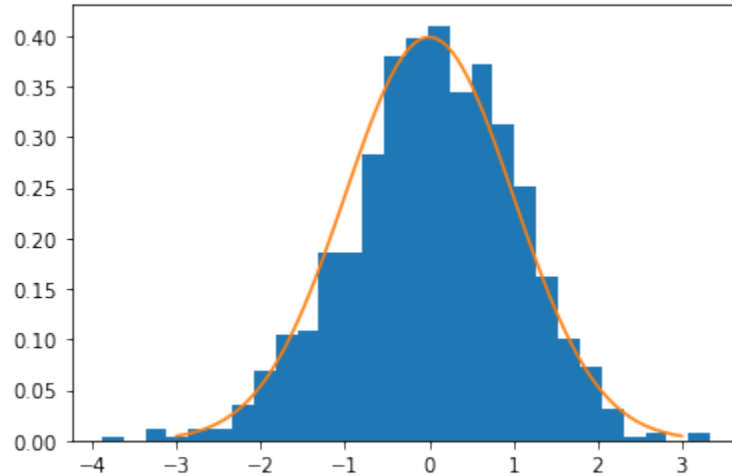
Pour simuler μ grâce à $\tilde{\mu}_N$, on génère par méthode du rejet un échantillon de N réalisation de X selon la loi à densité $\tilde{q} = \frac{3}{2}(1 - x^2)$. On a choisi cette fonction à partir de $g(x) = \cos(\frac{\pi x}{2})$ en prenant un développement limité en zero que l'on a translaté et renormalisé pour obtenir une densité à support dans $[0, 1]$.

Cette fois ci, nous obtenons la valeur $\tilde{\mu}_N = 0.63582$. Nous constatons une variance moindre.

Comme dans la partie précédente, nous estimons empiriquement la variance de $\tilde{\mu}_N$. Nous obtenons $Var_{emp}(\tilde{\mu}_N) = 1.93 \cdot 10^{-5}$, pour une valeur théorique de $Var_{\tilde{\mu}_N} \approx 1.98 \cdot 10^{-5}$.

Nous avons réalisé le même travail de comparaison de l'erreur renormalisée avec la loi normale centrée réduite. Nous présentons cela en figure ??.

Figure 4: Comparaison de la densité de la loi $\mathcal{N}(0, 1)$ et de l'erreur renormalisée de $\tilde{\mu}_N$



Nous nous sommes intéressé à la probabilité d'acceptation dans notre méthode du rejet. En effet, la complexité de notre algorithme d'estimation dépend de cette probabilité. On peut montrer qu'elle vaut $P_a = \frac{1}{C} = \frac{1}{\tilde{q}(0)} = \frac{2}{3}$ dans notre cas. Par le calcul, nous obtenons une probabilité d'acceptation empirique de 0,667, ce qui est attendu. Cela signifie qu'en moyenne, pour simuler 2 échantillons de X selon \tilde{q} , il nous faut de l'ordre de 3 opérations.

1.4 Conclusion

Pour conclure, nous avons testé deux méthodes pour calculer notre intégrale par méthode de Monte Carlo. Une méthode dite naïve qui donne lieu à une grande variance des résultats puis, une méthode plus raffinée qui fait appel à un changement de probabilité. Grâce à cette méthode, nous avons pu réduire la variance d'un facteur presque 100 (83,7 pour être précis.) C'est donc très satisfaisant. Comparons le budget des deux méthodes. Pour rappel, nous avons $N = 50$.

- Méthode, naïve : pour obtenir une estimation, nous réalisons de l'ordre de N opérations puisqu'il s'agit simplement de tirer N réalisations d'une variable aléatoire et d'en faire la moyenne.
- Importance sampling : Cette fois-ci, pour obtenir une estimation de μ , nous avons besoin de N réalisations de X selon \tilde{q} mais pour tirer X comme ceci, on fait en moyenne $\frac{3}{2}$ essais (inverse de la probabilité d'acceptation). In fine, on fait de l'ordre de $\frac{3}{2}N$ opérations, ce qui n'est pas beaucoup plus compte tenu du gain en variance.

On conclut donc que l'échantillonnage préférentiel est la méthode de choix lorsque l'on veut estimer une intégrale par méthode de Monte Carlo. Ceci a des applications dans le calcul d'intégrales sur des espaces de grande dimension. Cela sert en le filtrage bayésien, lorsque l'on propage des distributions conditionnelles par exemple.

2 TP2 : Filtre de Kalman: poursuite d'un mobile

2.1 Introduction

Dans ce TP, nous allons simuler la trajectoire d'un mobile, et tenter de l'estimer à partir d'observations partielles et bruitées.

2.1.1 Cadre général : filtre de Kalman

On se donne un système linéaire du type :

$$\begin{cases} X_k = F_k X_{k-1} + f_k + W_k \\ Y_k = H_k X_k + h_k + V_k \end{cases}$$

Où les variables aléatoires suivent les hypothèses de la partie 3.1 de [1], ce qui assure que le couple (X_k, Y_k) forme un vecteur gaussien. X_k représente un système dont la dynamique est stochastique. Y_k est une mesure de l'état du système. Le problème du filtrage consiste à estimer la loi de X_k conditionnellement à l'information disponible à l'instant k : $Y_{0:k}$. Pour cela, comme les vecteurs sont gaussiens, il suffit de caractériser l'espérance et la covariance des lois conditionnelles :

$$\hat{X}_k = \mathbb{E}[X_k | Y_{0:k}]$$
$$P_k = \mathbb{E} \left[\left(X_k - \hat{X}_k \right) \left(X_k - \hat{X}_k \right)^T | Y_{0:k} \right]$$

Il est pratique de définir également les quantités suivantes :

$$\hat{X}_k^- = \mathbb{E}[X_k | Y_{0:k-1}]$$
$$P_k^- = \mathbb{E} \left[\left(X_k - \hat{X}_k^- \right) \left(X_k - \hat{X}_k^- \right)^T | Y_{0:k-1} \right]$$

On montre alors que ces quantités obéissent aux relations de récurrence suivantes :

Etape de prédiction :

$$\hat{X}_k^- = F_k \hat{X}_{k-1} + f_k$$

$$P_k^- = F_k P_{k-1} F_k^T + Q_k^W \quad Q_k^W \text{ est la matrice de covariance de } (W_k)$$

Etape de correction :

$$\hat{X}_k = \hat{X}_k^- + K_k \left(Y_k - \left(H_k \hat{X}_k^- + h_k \right) \right)$$

$$P_k = (I - K_k H_k) P_k^-$$

Où K_k est le gain de Kalman.

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + Q_k^V)^{-1}$$

Avec Q_k^V la matrice de covariance de (V_k) .

2.2 Application, estimation d'une trajectoire

Dans cette partie, on se donne le système suivant : un mobile de masse m se déplace dans le plan. Il est décrit par son vecteur d'état,

$$X_t = \begin{pmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \end{pmatrix} \quad \forall t \in 0, \dots, T$$

Le mobile est soumis à une force aléatoire $(W_t)_{t \in 0, \dots, T}$. Les équations de Newton nous donnent sa dynamique :

$$m \begin{pmatrix} \ddot{x}_t \\ \ddot{y}_t \end{pmatrix} = cW_t = c \begin{pmatrix} w_k^1 \\ w_k^2 \end{pmatrix} \quad \forall t \in [t_k, t_{k+1}]$$

Où $c > 0$ et W_t est un bruit blanc gaussien centré réduit.

Après intégration en temps, la dynamique de X_t s'écrit sous la forme matricielle suivante :

$$X_t = \begin{pmatrix} 1 & 0 & \delta t & 0 \\ 0 & 1 & 0 & \delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \\ \dot{x}_{t-1} \\ \dot{y}_{t-1} \end{pmatrix} + \begin{pmatrix} \frac{c(\delta t)^2}{2m} & 0 \\ 0 & \frac{c(\delta t)^2}{2m} \\ \frac{c\delta t}{m} & 0 \\ 0 & \frac{c\delta t}{m} \end{pmatrix} \begin{pmatrix} w_t^1 \\ w_t^2 \end{pmatrix} = FX_{t-1} + GW_t$$

Avec condition initiale:

$$m_0 = \mathbb{E}[X_0] = \begin{pmatrix} 5000 \text{ m} \\ 5000 \text{ m} \\ -20 \text{ m/s} \\ 20 \text{ m/s} \end{pmatrix}, \quad P_0 = \begin{pmatrix} (2000 \text{ m})^2 & 0 & 0 & 0 \\ 0 & (2000 \text{ m})^2 & 0 & 0 \\ 0 & 0 & (5 \text{ m/s})^2 & 0 \\ 0 & 0 & 0 & (5 \text{ m/s})^2 \end{pmatrix}$$

Pour ce TP, nous prendrons les valeurs suivantes :

$$\begin{cases} \delta t = 1 \text{ seconde} \\ T = N\delta t = 200 \text{ s} \\ c = 2 \text{ m/s}^2 \end{cases}$$

A chaque instant, on réalise une mesure partielle de l'état X_t , on ne mesure que la position. De plus, on suppose notre appareil de mesure bruité :

$$Y_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_t \\ y_t \\ \dot{x}_t \\ \dot{y}_t \end{pmatrix} + \begin{pmatrix} v_t^1 \\ v_t^2 \end{pmatrix}$$

Où $V_t = (v_t^1, v_t^2)$ est un bruit blanc gaussien de matrice de covariance :

$$Q_k^V = \begin{pmatrix} (50 \text{ m})^2 & 0 \\ 0 & (50 \text{ m})^2 \end{pmatrix}$$

2.3 Simulations de la trajectoire et estimation par le filtre de Kalman

Dans cette partie, nous simulons une trajectoire du mobile ainsi qu'une mesure bruitée de son état à chaque étape. Nous présentons cela en figure 5.

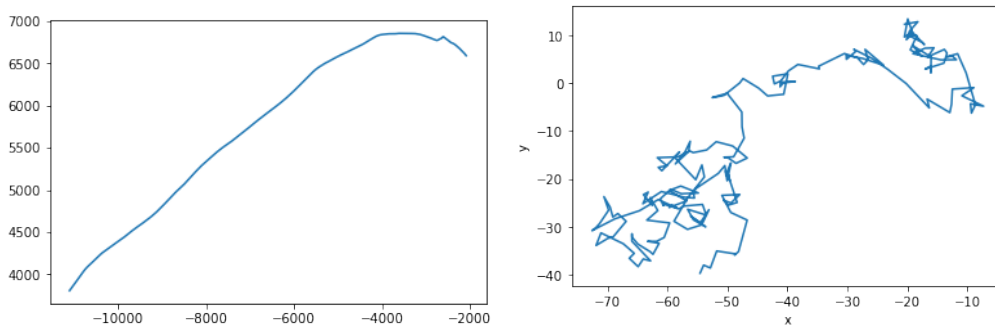
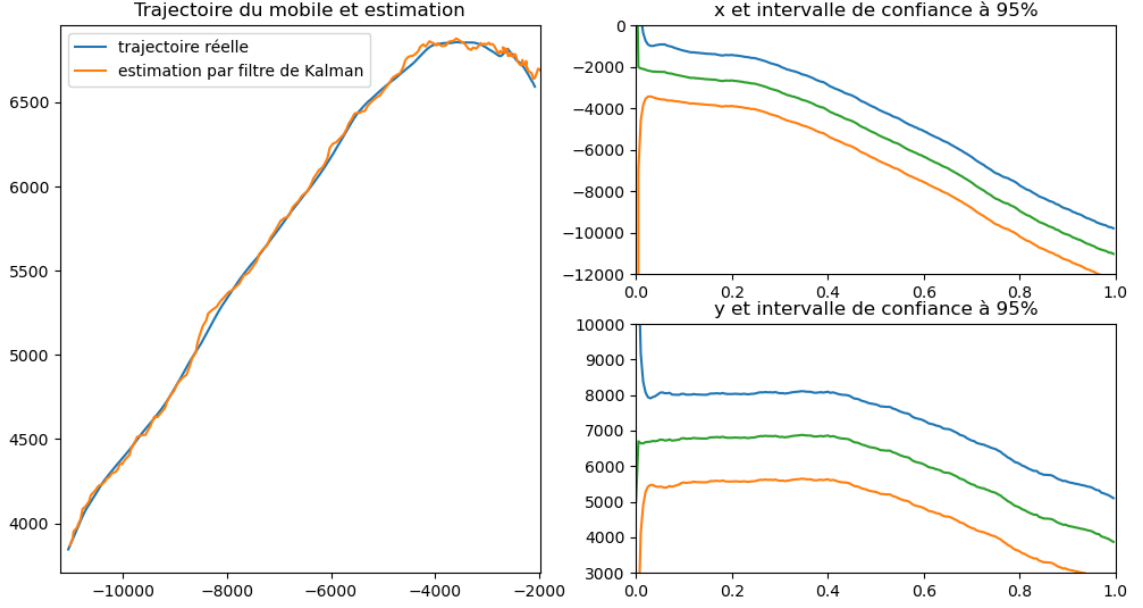


Figure 5: Simulation d'une trajectoire du mobile. Position à gauche et vitesse à droite.

Figure 6: Trajectoire estimée par le filtre de Kalman et intervalles de confiance à 95% associés.



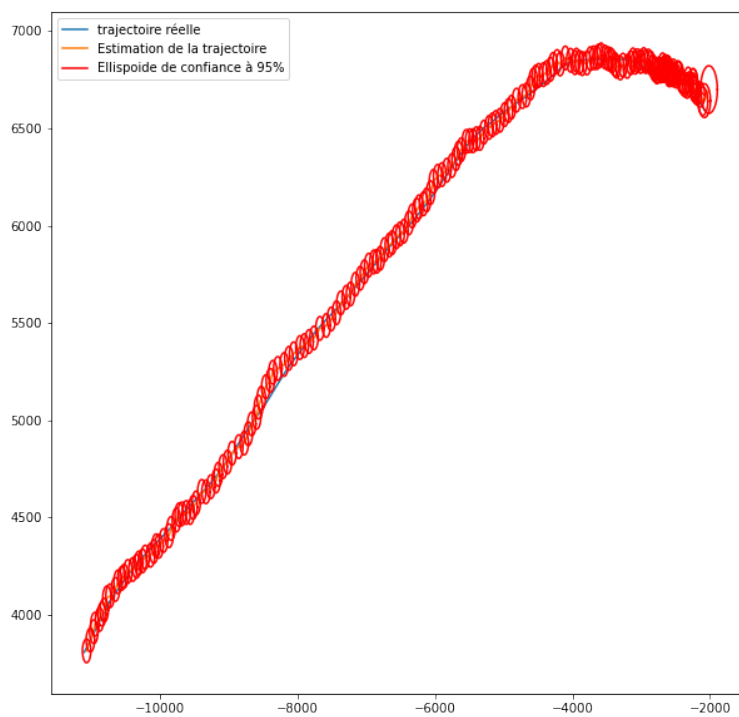
Ce qui est intéressant avec le filtre de Kalman est que l'on connaît la covariance de la distribution conditionnelle à chaque étape. On peut donc tracer des intervalles de confiance à 95%. C'est ce que nous présentons en figure 6.

D'ailleurs, puisque la méthode du filtre de Kalman fournit des matrices de covariance, on peut tracer des ellipses de confiance à 95% autour de l'estimation de la trajectoire de X c'est ce que nous présentons en figure 7.

On constate qu'au départ, il existe une grande incertitude sur la position du mobile. Cette incertitude décroît pendant les premières itérations, jusqu'à ce qu'elle se stabilise autour d'une certaine valeur. Il est intéressant de noter que les ellipses de confiance à 95% autour de l'estimation de la trajectoire contiennent bien les points de la trajectoire réelle, comme attendu.

2.4 Conclusion

Figure 7: Trajectoire estimée par le filtre de Kalman et ellipses de confiance à 95% associées.



3 TP3 : Borne de Cramer Rao.

3.1 Introduction

3.1.1 Borne de Cramer Rao pour un problème de filtrage

Dans ce TP, on se propose de calculer une borne de Cramer Rao pour un problème de filtrage. Le cadre général est le suivant. On se donne un système $(X_k)_{k \geq 0}$ sujet à une certaine dynamique, ainsi qu'une série de mesures bruitées $(y_k)_{k \geq 0}$ (bruit gaussien de variance σ_k^2) :

$$\begin{cases} X_{k+1} = \phi(k, k+1)X_k \\ y_k = h_k(X_k) + \epsilon_k \end{cases}$$

Avec $X_0 \sim \mathcal{N}(X_0, P_0)$. On peut montrer que la matrice d'information relative à l'instant k vérifie la relation de récurrence suivante :

$$J_k = \frac{1}{\sigma_k} \left(\frac{\partial h_k}{\partial X_k} \right) \left(\frac{\partial h_k}{\partial X_k} \right)^T + \left(\frac{\partial X_{k-1}}{\partial X_k} \right)^T J_{k-1} \left(\frac{\partial X_{k-1}}{\partial X_k} \right)$$

Avec $J_0 = P_0^{-1}$ et $\phi(k, k+1) = \frac{\partial X_{k+1}}{\partial X_k}$. La borne de Cramer Rao à l'instant k est alors égale à l'inverse de la matrice d'information :

$$BCR_k = J_k^{-1}$$

3.1.2 Application : poursuite de bateau

Nous étudions le système suivant : un bateau émetteur de bruit se déplace dans le plan de manière rectiligne uniforme. Un bateau observateur mesure la direction d'où lui parvient le bruit, plus précisément, il mesure l'angle θ_k que fait cette direction avec l'horizontale, ce toute les secondes pendant 100s. Le bateau observateur se déplace également en ligne droite à vitesse constante. L'état de l'émetteur est donné par sa position et sa vitesse : $X_k = (x_k, y_k, \dot{x}_k, \dot{y}_k)$, sa matrice de transition est :

$$\phi(k, k+1) = \phi = \begin{pmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Pour pouvoir discriminer entre différentes trajectoires possibles, le bateau doit virer de bord à un certain moment. Si ça n'est pas le cas, on peut vérifier que des trajectoires différentes peuvent donner lieu à des mesures d'angle identiques. Pour simplifier l'étude, on suppose que le bateau vire de bord toujours au même moment : au milieu du trajet. Le bateau observateur choisit l'angle φ duquel il tourne. Chaque choix de φ conduit à une précision différente de l'estimation de la trajectoire de l'émetteur. La question est donc la suivante : quel choix de φ est le meilleur pour estimer la trajectoire de l'émetteur?

Pour répondre à cette question, nous allons simuler la dynamique de l'émetteur et de l'observateur. Grâce à ces dynamiques, nous allons calculer le gradient de la fonction d'observation et nous allons en déduire la matrice d'information. Dans notre cas, la fonction d'observation est l'angle que fait la direction d'observation avec l'horizontale. En terme de l'abscisse et de l'ordonnée de l'émetteur et de l'observateur, cette fonction s'écrit :

$$h_k(X_k) = \arctan \left(\frac{y_k - y_k^o}{x_k - x_k^o} \right)$$

En inversant la matrice d'observation, on obtiendra la borne de Cramer Rao, sous forme d'une matrice carrée d'ordre 4. On cherchera ensuite la valeur de φ qui minimise le critère : $\sqrt{\sigma_{x_n}^2 + \sigma_{y_n}^2}$

3.1.3 Calcul du gradient de h_k par rapport à X_k

Déjà, on voit facilement que :

$$\frac{\partial h_k}{\partial \dot{x}_k} = 0$$

et :

$$\frac{\partial h_k}{\partial \dot{y}_k} = 0$$

Ensuite, par le calcul, on a que :

$$\frac{\partial h_k}{\partial x_k} = \frac{y_k^o - y_k}{(x_k - x_k^o)^2 \left(1 + \left(\frac{y_k - y_k^o}{x_k - x_k^o} \right)^2 \right)}$$

et :

$$\frac{\partial h_k}{\partial y_k} = \frac{1}{(x_k - x_k^o) \left(1 + \left(\frac{y_k - y_k^o}{x_k - x_k^o} \right)^2 \right)}$$

3.2 Simulations des dynamiques de l'observateur et de l'émetteur

Nous avons simulé la dynamique du système pour différents changements de cap de l'observateur, avec les paramètres du tableau 8. Nous obtenons les tracés de la figure ??

Figure 8: Paramètres pour la simulation

Horizon temporel	$T = 1$
Nombre de mesures	$N = 100$
Ecart type du bruit	$\sigma_k = 1^\circ$
Vitesse de l'observateur	$V_o = 10m.s^{-2}$
Position initiale de l'observateur	$(x_0^o, y_0^o) = (0, 0)$ m
Vitesse de l'émetteur	$V_e = 5m.s^{-2}$
Cap de l'émetteur	$\alpha = -20^\circ$
Position initiale de l'émetteur	$(x_0, y_0) = (2000, 2000)$
Incertitude initiale sur la position de l'émetteur	$P_0 = diag(1000, 1000, 10, 10)^2$

Grâce à ces dynamiques, nous sommes capables de calculer la Borne de Cramer Rao associée à cet estimateur de la position de l'émetteur. Nous présentons ces résultat en figure 10.

3.3 Optimisation de la valeur de φ

Pour trouver la valeur de φ optimale, nous avons simplement testé toutes les valeurs entre 0° et 360° , avec un pas de 1° . Nous avons trouvé une valeur optimale de $\varphi = 109^\circ$, ce qui correspond à la trajectoire représentée en figure 11. La borne de Cramer Rao associée est représentée en figure 12. Le critère associé est un écart type sur la position de 128m.

3.4 Conclusion

En conclusion, ce TP nous aura permis de voir comment la Borne de Cramer Rao permet de réaliser le choix d'une stratégie pour optimiser la précision sur l'estimation de état d'un système, grâce à la seule connaissance d'un modèle à priori.

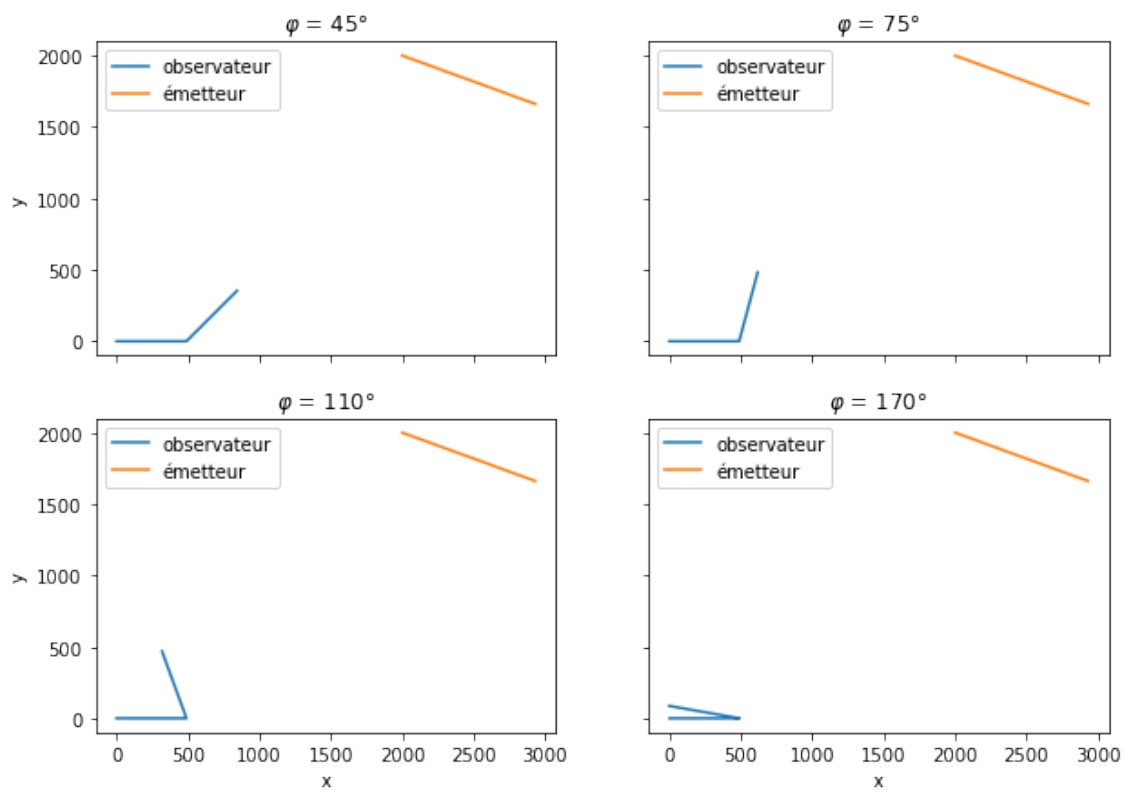


Figure 9: Trajectoires de l'émetteur et de l'observateur pour différentes valeurs de l'angle φ

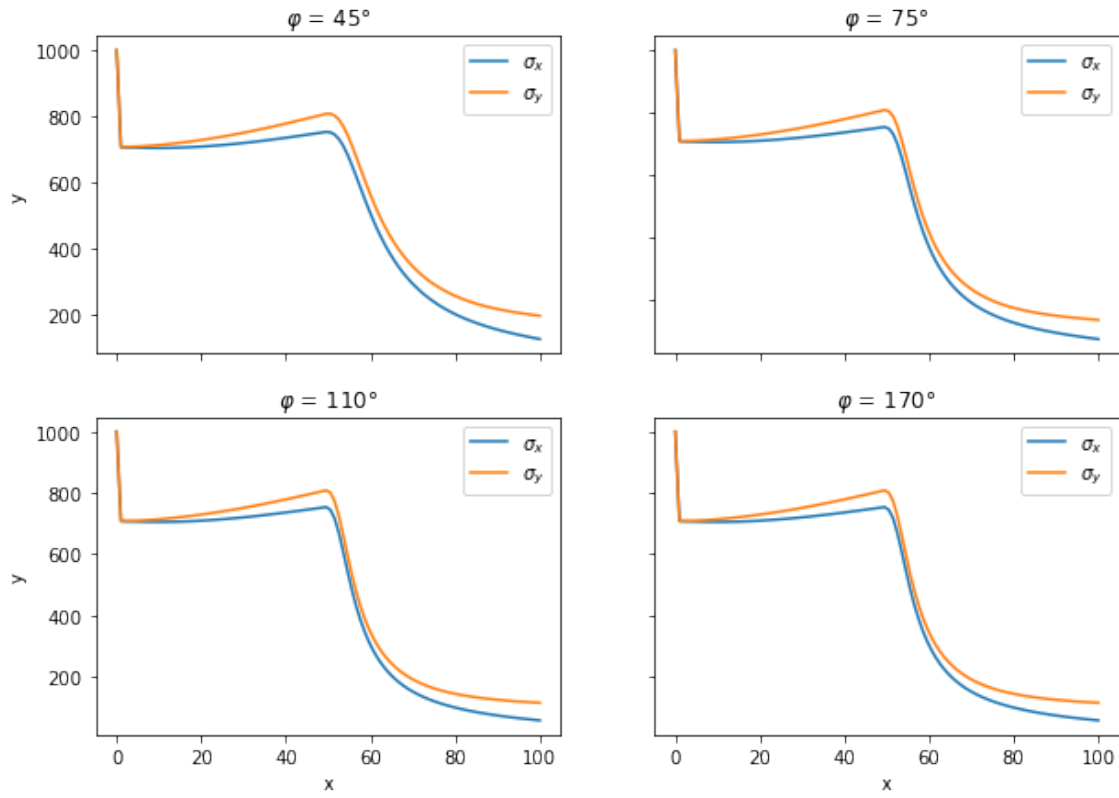


Figure 10: Borne de Cramer Rao pour l'abscisse et l'ordonnée de la position de l'émetteur au cours du temps, pour différentes valeurs de φ

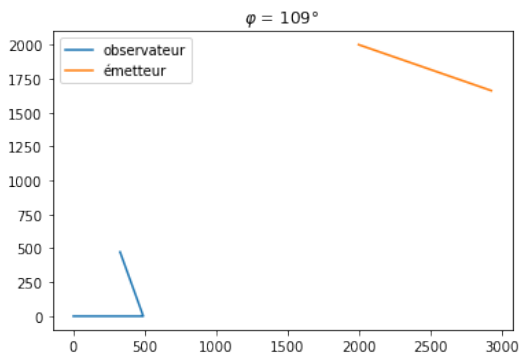


Figure 11: Trajectoire optimale pour l'observateur

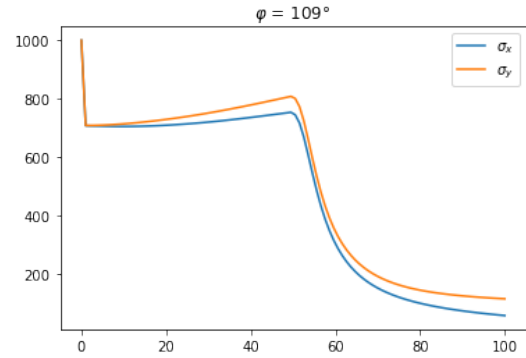


Figure 12: Borne de Cramer Rao pour la trajectoire optimale

4 TP4: Navigation par corrélation de terrain - filtre particulaire

4.1 Introduction

Ce TP est l'aboutissement du tout ce qui précède. Dans ce TP, nous allons implémenter le filtre particulaire pour un problème de filtrage non linéaire : la navigation par corrélation de terrain. La première partie présente ce qu'est le filtre particulaire. La deuxième partie discute de notre problème spécifique et de l'implémentation du filtre particulaire dans ce cas.

4.2 Filtre optimal et filtre particulaire

4.2.1 Filtre optimal

Rappelons les équations du filtrage :

On considère une chaîne de Markov $(X_n)_{n \geq 0}$ à valeurs dans \mathbb{R}^d , de loi initiale π_0 , i.e. $\pi_0(dx_{1:d}) = P(X_0 \in dx_{1:d})$, et de noyau de transition $Q_n(x_{n-1}, dx_{1:d}) = P(X_n \in dx_{1:d} | X_{n-1} = x_{n-1})$. On considère les observations bruitées $Y_n = h(X_n) + V_n \in \mathbb{R}^q$. On suppose que les V_n sont iid et a densité g par rapport à la mesure de Lebesgue. On suppose de plus que $V_n \perp X_0$. On introduit la fonction de vraisemblance $\psi_n(x) = g(Y_n - h(x))$ c'est donc une fonction aléatoire qui n'est connue que lorsque l'on a la n ème réalisation de Y . La fonction de vraisemblance mesure la cohérence entre le point de l'espace d'état x et l'observation courante.

On appelle filtre optimal la suite de mesures de probabilité suivante : $(\Pi_n)_{n \geq 0}$ ou $\Pi_0 = \pi_0$ et $\Pi_n(dx_{1:d}) = P(X_n \in dx_{1:d} | Y_{0:n})$. Il est également utile d'introduire la suite $\Pi_{n|n-1} = P(X_n \in dx_{1:d} | Y_{0:n-1})$. On montre que :

$$\begin{aligned} \Pi_n(dx_{1:d}) &= P(X_n \in dx_{1:d} | Y_{0:n}) \\ &= \frac{P(X_n \in dx_{1:d}, Y_n \in dy_{1:q} | Y_{0:n-1})}{P(Y_n \in dy_{1:q} | Y_{0:n-1})} \\ &= \frac{P(Y_n \in dy_{1:q} | X_n = x_n) P(X_n \in dx_{1:d} | Y_{0:n-1})}{P(Y_n \in dy_{1:q} | Y_{0:n-1})} \\ &= \frac{P(Y_n \in dy_{1:q} | X_n = x_n) P(X_n \in dx_{1:d} | Y_{0:n-1})}{\int_{x_n \in \mathbb{R}^d} P(Y_n \in dy_{1:q} | X_n = x_n) P(X_n \in dx_{1:d} | Y_{0:n-1})} \end{aligned}$$

On peut voir que :

$$\begin{aligned} P(Y_n \in dy_{1:q} | X_n = x_n) &= P(V_n \in dy_{1:q} - h(x_n) | X_n = x_n) \\ &= g(Y_n - h(x_n)) dy_{1:q} \end{aligned}$$

de sorte que :

$$\Pi_n(dx_{1:d}) = \frac{\psi_n(x_n) \Pi_{n|n-1}(dx_{1:d})}{\int_{x_n \in \mathbb{R}^d} \psi_n(x_n) \Pi_{n|n-1}(dx_{1:d})}$$

Cette opération s'appelle le produit projectif de $\Pi_{n|n-1}$ par ψ , et se note : $\Pi_n = \psi_n \cdot \Pi_{n|n-1}$

Il reste donc à voir comment passer de Π_{n-1} à $\Pi_{n|n-1}$.

$$\begin{aligned} \Pi_{n|n-1}((dx_{1:d})_n) &= P(X_n \in (dx_{1:d})_n | Y_{0:n-1}) \\ &= \int_{x_{n-1} \in \mathbb{R}^d} P(X_n \in (dx_{1:d})_n, X_{n-1} \in (dx_{1:d})_{n-1} | Y_{0:n-1}) \\ &= \int_{x_{n-1} \in \mathbb{R}^d} P(X_n \in (dx_{1:d})_n | X_{n-1}, Y_{0:n-1}) \Pi_n((dx_{1:d})_{n-1}) \end{aligned}$$

Mais X_n est une chaîne de Markov donc, en particulier, sa loi ne dépend pas des observations passées $Y_{0:n-1}$. on a alors :

$$\begin{aligned}\Pi_{n|n-1}((dx_{1:d})_n) &= \int_{x_{n-1} \in \mathbb{R}^d} P(X_n \in (dx_{1:d})_n | X_{n-1}) \Pi_n((dx_{1:d})_{n-1}) \\ &= \int_{x_{n-1} \in \mathbb{R}^d} \Pi_{n-1}((x_{1:d})_{n-1}) Q_n(x_{n-1}, (dx_{1:d})_n) \\ &= (\Pi_{n-1} Q_n)((dx_{1:d})_n)\end{aligned}$$

On appelle cette opération le "transport" de la mesure Π_{n-1} par le noyau Q_n .
On peut résumer le filtre optimal par le schéma suivant :

$$\Pi_{n-1} \xrightarrow{\text{prediction}} \Pi_{n|n-1} = (\Pi_{n-1} Q_n) \xrightarrow{\text{correction}} \Pi_n = \psi_n \cdot \Pi_{n|n-1}$$

4.2.2 Filtre particulaire

En pratique, le filtre optimal est très compliqué à calculer car les étapes de prédiction et de correction font intervenir des intégrales dans des espaces à (possiblement) grande dimension. On préfère alors utiliser le filtre particulaire. L'idée du filtre particulaire est d'approcher le filtre optimal par des méthodes de Monte Carlo. Le filtre particulaire repose sur l'idée de mesure empirique. On appelle mesure empirique associée à un échantillon (X_0, \dots, X_n) iid suivant la loi Π la mesure $S^N(\Pi) = \frac{1}{N} \sum_{i=1}^N \delta_{X_i}$ ou δ_x désigne la mesure de Dirac en x . Il faut bien voir qu'étant donné une mesure Π , $S^N(\Pi)$ est une mesure aléatoire. Elle dépend de l'échantillon (X_0, \dots, X_n) . Le filtre particulaire consiste alors en les opérations suivantes :

$$\Pi_{n-1}^N \xrightarrow{\text{prediction échantillonnée}} \Pi_{n|n-1}^N = S^N(\Pi_{n-1}^N Q_n) \xrightarrow{\text{correction}} \Pi_n^N = \psi_n \cdot \Pi_{n|n-1}^N$$

Regardons en détail les étapes de prédiction et de correction pour comprendre l'intérêt du filtre particulaire.

correction A cette étape, on dispose de $\Pi_{n|n-1}^N = \frac{1}{N} \sum_{i=1}^N \delta_{X_{n|n-1}^{i,N}}$, ou $(X_{n|n-1}^{1,N}, \dots, X_{n|n-1}^{N,N})$ est un échantillon iid distribué selon $\Pi_{n|n-1}^N Q_n$. On peut alors calculer Π_n^N :

$$\begin{aligned}\Pi_n^N(dx_{1:d}) &= \psi_n \cdot \Pi_{n|n-1}^N \\ &= \frac{\psi_n(x) \Pi_{n|n-1}^N(dx_{1:d})}{\int_{x \in \mathbb{R}^d} \psi_n(x) \Pi_{n|n-1}^N(dx_{1:d})} \\ &= \frac{\sum_{i=1}^N \psi_n(X_{n|n-1}^{i,N}) \delta_{X_{n|n-1}^{i,N}}}{\sum_{i=1}^N \psi_n(X_{n|n-1}^{i,N})} \\ &= \frac{1}{N} \sum_{j=1}^N \psi_n(X_{n|n-1}^{j,N})\end{aligned}$$

On pose alors :

$$\omega_n^i = \frac{\psi_n(X_{n|n-1}^{i,N})}{\frac{1}{N} \sum_{j=1}^N \psi_n(X_{n|n-1}^{j,N})}$$

De sorte que :

$$\Pi_n^N = \sum_{i=1}^N \omega_n^i \delta_{X_{n|n-1}^{i,N}}$$

On voit donc que l'étape de correction devient une étape de pondération de chaque particule par un poids proportionnel à sa vraisemblance.

prédiction échantillonnée A l'étape n , on a $\Pi_n^N = \sum_{i=1}^N \omega_n^i \delta_{X_{n|n-1}^{i,N}}$. On peut alors calculer le transport de la mesure Π_n^N par le noyau Q_{n+1} :

$$\begin{aligned} (\Pi_n^N Q_{n+1})(dx_{1:d}) &= \int_{y_{1:d} \in \mathbb{R}^d} \Pi_n^N(dy_{0:d}) Q_{n+1}(y_{1:d}, dx_{1:d}) \\ &= \int_{y_{1:d} \in \mathbb{R}^d} \sum_{i=1}^N \omega_n^i \delta_{X_{n|n-1}^{i,N}} Q_{n+1}(y_{1:d}, dx_{1:d}) \\ &= \sum_{i=1}^N \omega_n^i Q_{n+1}(X_{n|n-1}^{i,N}, dx_{1:d}) \end{aligned}$$

Ecris sous cette forme, on constate que $(\Pi_n^N Q_{n+1})$ est un mélange de probabilités. Or il est facile de générer un échantillon selon un mélange de probabilités. On peut faire comme ceci :

- Générer I selon $P(I = i) = \omega_n^i$
- Générer X selon $Q_{n+1}(X_{n|n-1}^{I,N}, \cdot)$

Cela revient à sélectionner les particules selon leur poids (plus le poids d'une particule est grand, plus on a de chance de la tirer), puis à faire évoluer la particule selon la dynamique de la chaîne de Markov.

4.2.3 Dynamique déterministe

Le cas d'une dynamique déterministe $X_{n+1} = f(X_n)$ est un peu embêtant. En effet, l'étape de prédiction échantillonnée fait une sélection des particules. Si aucun bruit n'est ajouté par l'évolution, la sélection va finir par "tuer" la diversité des particules. Dans ce cas (et c'est le cas qui se présente pour nous dans le TP), il est nécessaire de régulariser la mesure Π_n^N (à l'étape de prédiction). Concrètement, on remplace Π_n^N par :

$$\tilde{\Pi}_n^N = \sum_{i=1}^N \omega_n^i K\left(\frac{\cdot - X_{n|n-1}^{i,N}}{h}\right)$$

Où K est un noyau (Gaussien, Epanechnikov,...) et h est un paramètre qui contrôle la " finesse " de la régularisation. Il existe un choix de h qui minimise l'erreur quadratique moyenne entre la mesure discrète et la mesure régularisée, notons le h_{opt} .

L'étape de prédiction échantillonnée devient alors :

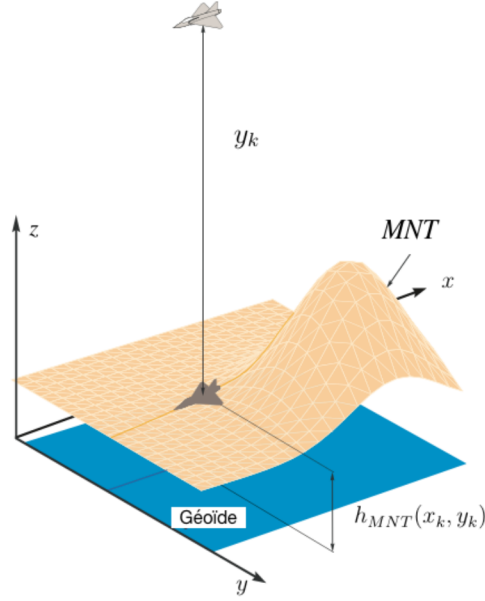
- Générer I selon $P(I = i) = \omega_n^i$
- Générer ϵ selon K .
- Poser $\tilde{X}_{n|n-1}^{i,N} = X_{n|n-1}^{I,N} + h_{opt} \Gamma_k \epsilon$. Avec, $\Gamma_k^T \Gamma_k = Cov\left((X_{n|n-1}^{i,N})_{i \in \{1, \dots, N\}}\right)$
- Faire évoluer la particule : $X_n^{i,N} = f(\tilde{X}_{n|n-1}^{i,N})$

4.3 Navigation par corrélation de Terrain

Le problème de la navigation par corrélation de terrain est le suivant : un avion se déplace en ligne droite à vitesse constante (horizontale). Disons qu'à un moment, il perd son signal GPS et se retrouve "perdu", c'est à dire qu'il y a une incertitude sur son état (position/vitesse). Dans ce cas, il existe un moyen pour l'avion de retrouver sa position avec une précision acceptable. En effet, les avions embarquent dans leur ordinateurs de bord des modèles numériques de terrain, c'est à dire une carte de l'altitude du sol en fonction de la position. En effectuant des mesures d'altitude et en "corrélant" ces mesures au modèle de terrain, l'avion peut se retrouver. Naturellement, ces mesures sont bruitées. On retrouve en fait un problème de filtrage avec les paramètres suivants :

$$X_k = (x_k, y_k, z_k, \dot{x}_k, \dot{y}_k, \dot{z}_k)^T$$

Figure 13: Navigation par corrélation de terrain, illustration



$$\begin{cases} X_k = \Phi X_{k-1} \\ Y_k = h(X_k) + \epsilon_k = Z_k - h_{MNT}(x_k, y_k) + \epsilon_k \end{cases}$$

Avec,

$$\Phi = \begin{pmatrix} 1 & 0 & 0 & dt & 0 & 0 \\ 0 & 1 & 0 & 0 & dt & 0 \\ 0 & 0 & 1 & 0 & 0 & dt \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Dans le TP, h_{MNT} est un tableau de nombres donné.

Comme discuté précédemment, la dynamique de ce problème est déterministe (l'avion se déplace en ligne droite, à vitesse constante). Il faut donc mettre en oeuvre la technique de régularisation. Nous avons choisi d'utiliser un noyau Gaussien. Simplement, cette étape est assez coûteuse. Pour s'épargner du temps de calcul, on ne la pratique que lorsque l'on détecte que la diversité de l'échantillon de particules est trop faible. Quand la condition $N_{eff} = \frac{1}{\sum_{i=1}^N \omega_i^2} \leq N_{th} = Coeff \times N$, on procède au rééchantillonnage. En pratique, un bon choix est $Coeff = 0.4$.

Les paramètres du problème sont résumés dans le tableau 14 :

Pour initialiser l'algorithme, on tire une position initiale "fausse", uniformément autour de X_0 , dans les intervalles $X_0 \pm \sigma_{x,y,z}$, puis on tire les particules initiales autour de cette fausse position, selon une distribution gaussienne de matrice de covariance P_0 .

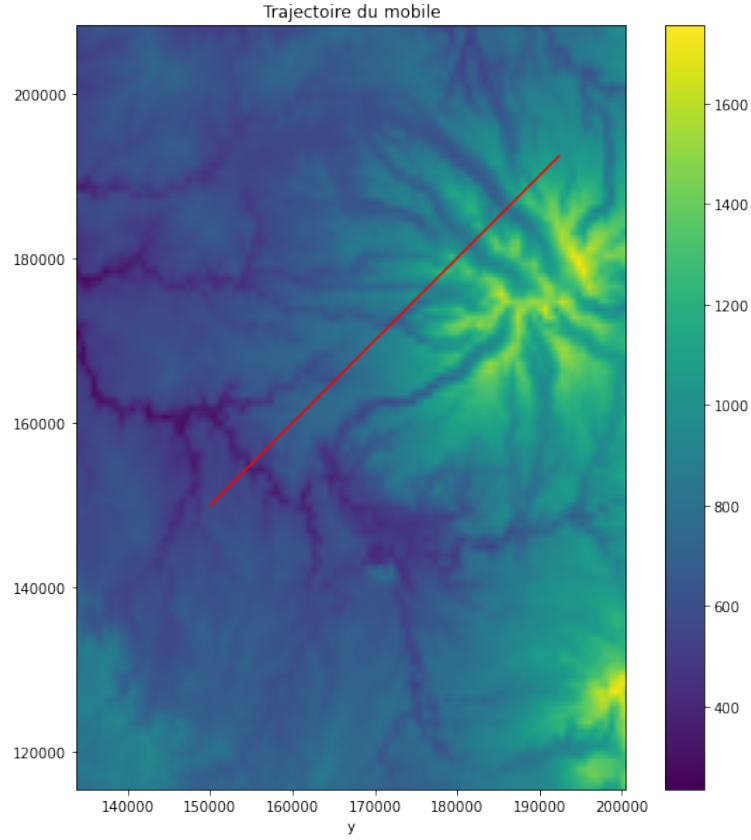
La figure 15 présente la trajectoire de l'avion. La figure 16 montre les mesures que fait l'avion le long d'une trajectoire (avec et sans le bruit). La figure 17 montre le tirage initial des particules.

Résultats Après quelques efforts, nous avons été capables de faire fonctionner le filtre particulaire. La figure ?? présente l'évolution des particules. On constate bien que le groupe de particules se concentre autour de la vraie position de l'avion. De plus, on observe clairement les étapes de redistribution, et l'on comprend bien comme elles agissent pour explorer l'espace des états à partir des positions pertinentes connues.

Figure 14: Paramètres pour la simulation

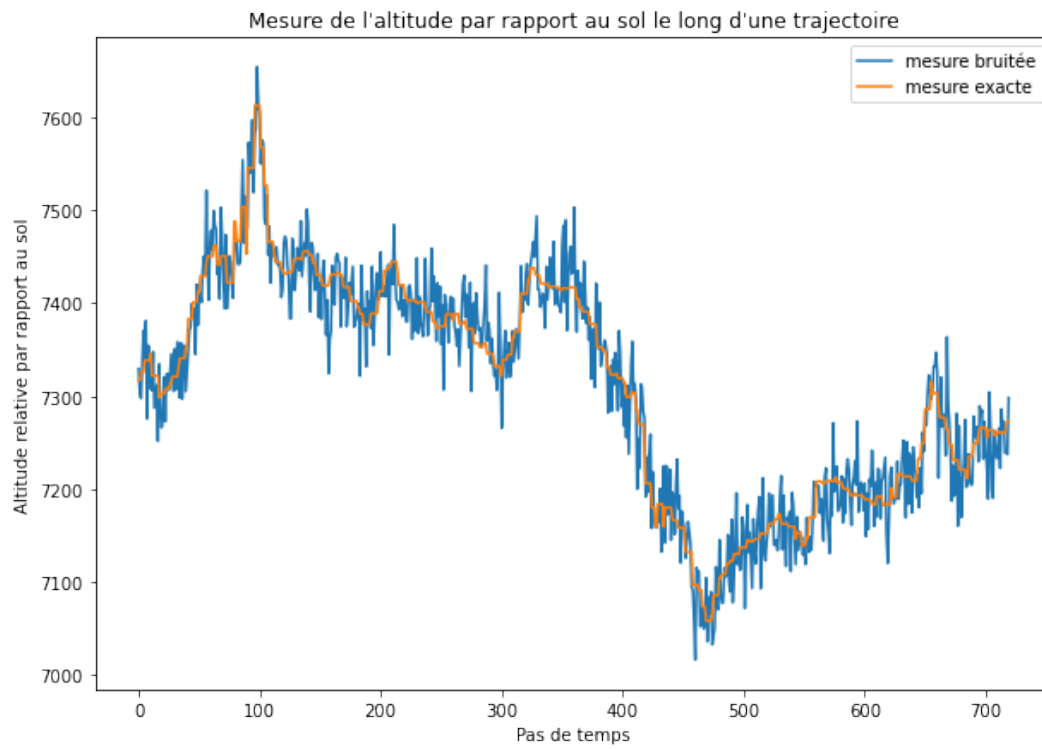
Horizon temporel	$T = 720s$
Vitesse avion	$v = 300km.h^{-1}$
Cap de l'avion	$\theta = 45^\circ$
Etat initial	$X_0 = (150, 150, 8, v \cos(\theta), v \sin(\theta), 0) km$
dt	$dt = 1s$
Ecart type bruit mesure	$\sigma_{mes} = 30m$
Incertitude initiale position(forme P_0)	$(\sigma_x, \sigma_y, \sigma_z) = (3000, 3000, 500)m$
Incertitude initiale vitesse (forme P_0)	$(\sigma_{v_x}, \sigma_{v_y}, \sigma_{v_z}) = (5, 5, 5)m.s^{-1}$
Nombre de particules	$N = 5000$

Figure 15: Trajectoire de l'avion



4.4 Conclusion

Figure 16: Trajectoire de l'avion



References

- [1] François LEGLAND. Filtrage bayésien optimal et approximation particulaire. *Les Presses de l'Ecole Nationale Supérieure de Techniques Avancées*, 2008.

Figure 17: Tirage initial des particules

