

École Chercheurs en “traitement du signal” – 18 au 20 octobre 2004

filtrage particulaire

Introduction et application
à la poursuite de mobile dans un réseau cellulaire

Fabien Campillo

`Fabien.Campillo@inria.fr`

INRIA Rennes

introduction

- **système non observé**: on se donne le modèle (loi **a priori**)
 - **observations**: sont disponibles à des instants donnés
- traiter les observations au fur et à mesure de leur disponibilité afin d'affiner notre connaissance du système non observé (loi **a posteriori**)
- approche **bayésienne**: loi a priori + observations → loi a posteriori

plan

- estimation bayésienne
- système à espace d'état
- filtre non linéaire
- filtre de Kalman
- filtre de Kalman étendu
- monte carlo
- filtre particulaire
- géolocalisation
- obstacles

estimation bayésienne

loi a priori

X une **variable aléatoire** (v.a.) à valeurs dans \mathbb{R}^n de **loi** $p(X)$, i.e.

$$\mathbb{E}\phi(X) = \int_{\mathbb{R}^n} \phi(X) p(X) dX \quad \forall \phi \quad (\phi \text{ fonction continue bornée})$$

$$\mathbb{P}(X \in A) = \int_A p(X) dX \quad \forall A \quad (A \subset \mathbb{R}^n \text{ événement})$$

Notations de type “ingénierie”:

- X désigne à la fois la variable aléatoire et ses réalisations
- $p(X)$ désigne aussi bien la loi de X que sa densité

loi gaussienne

$$X \sim N(\mu, \Gamma)$$

- si la matrice de covariance Γ est inversible (i.e. $\det(\Gamma) \neq 0$) alors

$$p(X) = \frac{1}{\sqrt{(2\pi)^n \det(\Gamma)}} \exp \left(-\frac{1}{2} (X - \mu)^* \Gamma^{-1} (X - \mu) \right)$$

- (* désigne l'opérateur de transposition) si la matrice de covariance Γ est dégénérée (i.e. $\det(\Gamma) = 0$) alors

$$\mathbb{E} e^{i r^* X} = \exp \left(i r^* \mu - \frac{1}{2} r^* \Gamma r \right) \quad \forall r \in \mathbb{R}^n$$

(où $i^2 = -1$)

masse de Dirac

$$X \sim p(X) = \delta_{x_0}(X)$$

$$\mathbb{E}\phi(X) = \phi(x_0) \quad \forall \phi$$

$$\mathbb{P}(X \in A) = \mathbf{1}_A(x_0) \quad \forall A$$

cela signifie que $X = x_0$ p.s.

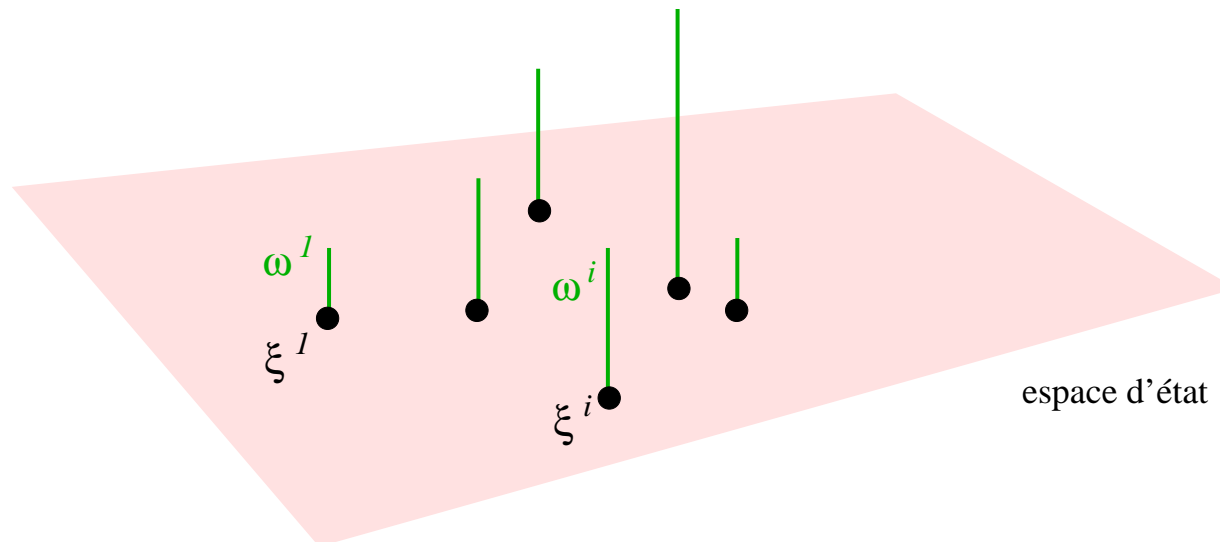
loi multinomiale

$$X \sim p(X) = \sum_{i=1}^N \omega^i \delta_{\xi^i}(X) \text{ avec } \omega^i \geq 0 \text{ et } \sum_{i=1}^N \omega^i = 1$$

$$\mathbb{E}\phi(X) = \sum_{i=1}^N \omega^i \phi(\xi^i) \quad \forall \phi$$

$$\mathbb{P}(X \in A) = \sum_{i=1}^N \omega^i \mathbf{1}_A(\xi^i) \quad \forall A$$

cela signifie que $X = \xi^i$ avec probabilité ω^i ($i = 1 \dots N$)



marginale

soit (X, Y) un v.a. à valeurs dans $\mathbb{R}^n \times \mathbb{R}^d$ et $p(X, Y)$ sa loi, la loi de X est la loi marginale

$$p(X) = \int_{\mathbb{R}^d} p(X, Y) dY$$

loi a posteriori

- (X, Y) à valeurs dans $\mathbb{R}^n \times \mathbb{R}^d$
 - X état d'un système (non observé)
 - Y observation

la loi $p(X, Y)$ est donnée

- on veut estimer X à partir de Y , i.e. on cherche une fonction $\psi^*(Y)$ telle que

$$\mathbb{E}|X - \psi^*(Y)|^2 = \min_{\psi(\cdot)} \mathbb{E}|X - \psi(Y)|^2$$

(i.e. un estimateur de X qui minimise le risque quadratique), cet estimateur est noté

$$\hat{X} = \mathbb{E}(X|Y)$$

(espérance conditionnelle de X sachant Y)

- de même: on veut estimer $\phi(X)$ à partir de Y , on note cet estimateur

$$\widehat{\phi(X)} = \mathbb{E}(\phi(X)|Y)$$

loi a posteriori (suite)

la loi conditionnelle de X sachant Y est

$$(LC) \quad p(X|Y) = \frac{p(X, Y)}{p(Y)} = \frac{p(X, Y)}{\int_{\mathbb{R}^n} p(X, Y) dX}$$

alors

$$\mathbb{E}(\phi(X)|Y) = \int_{\mathbb{R}^n} \phi(X) p(X|Y) dX$$

(et donc $\mathbb{E}(X|Y) = \int_{\mathbb{R}^n} X p(X|Y) dX$)

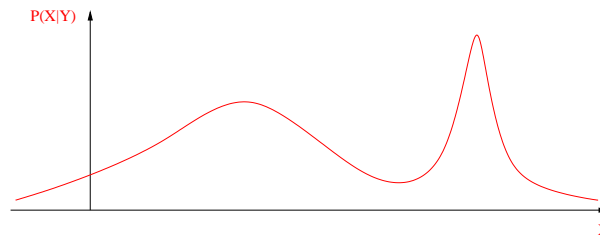
$p(X|Y)$ représente toute l'information sur X contenue dans Y

en pratique on calcule $p(X|Y)$ à partir de $p(X, Y)$

loi a posteriori (suite)

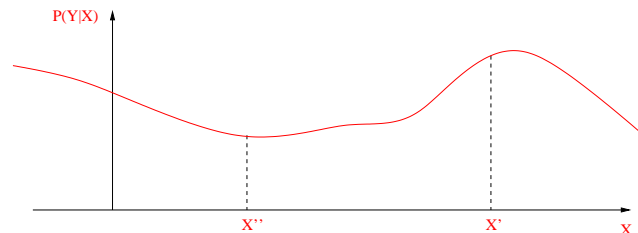
deux lectures:

- $X \mapsto p(X|Y)$ à Y fixé: c'est une **densité**



- $X \mapsto p(Y|X)$ à Y fixé: c'est une **vraisemblance**

i.e. une fonction positive telle que si $p(Y|X') > p(Y|X'')$ alors Y a été plus vraisemblablement généré par X' que par X''



propriétés

(P1) $p(X, Y) = p(X|Y) p(Y)$

(P2) $p(X, Y, Z) = p(X|Y, Z) p(Y|Z) p(Z)$

(P3) $p(X) = \int p(X|Y) p(Y) dY$

(P4) $p(X|Z) = \int p(X|Y, Z) p(Y|Z) dY$

indépendance

X et Y sont dits **indépendants** (noté $X \perp Y$) lorsque

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B) \quad \forall A, B$$

propriétés

- (i) $X \perp Y$ ssi $p(X, Y) = p(X) p(Y)$
- (ii) $X \perp Y$ ssi $p(X|Y) = p(X)$

formule de Bayes

comme $p(X, Y) = p(X|Y) p(Y) = p(Y|X) p(X)$, donc:

$$(FB) \quad p(X|Y) = \frac{p(Y|X) p(X)}{p(Y)} = \frac{p(Y|X) p(X)}{\int p(Y|X) p(X) dX}$$

donc, à Y fixé, la **loi a posteriori** $p(X|Y)$ est égale (à une constante multiplicative près) au produit de la **loi a priori** $p(X)$ et de la **vraisemblance** $p(Y|X)$,

$$\text{loi a posteriori} = \frac{1}{\text{cte de normalisation}} \times \text{vraisemblance} \times \text{loi a priori}$$

la constante de normalisation est simplement $\int p(Y|X) p(X) dX$
on note

$$p(X|Y) \propto p(Y|X) p(X)$$

loi a posteriori: cas gaussien

soit $Z = \begin{pmatrix} X \\ Y \end{pmatrix}$ de loi gaussienne $N(\bar{Z}, Q_Z)$ sur \mathbb{R}^{n+d} avec:

$$\bar{Z} = \begin{pmatrix} \bar{X} \\ \bar{Y} \end{pmatrix} \quad Q_Z = \begin{pmatrix} Q_{XX} & Q_{XY} \\ Q_{XY}^* & Q_{YY} \end{pmatrix}$$

alors $p(X|Y)$ est gaussien de moyenne et covariance:

$$\hat{X} = \bar{X} + Q_{XY} Q_{YY}^{-1} (Y - \bar{Y})$$

$$R = Q_{XX} - Q_{XY} Q_{YY}^{-1} Q_{XY}^*$$

loi a posteriori: cas gaussien (suite)

preuve: supposons $Q_Z > 0$, la densité conditionnelle de $X|Y = y$ est

$$p(X | Y) = \frac{p(X, Y)}{p(Y)} = \frac{\frac{1}{(\sqrt{2\pi})^{n+d} \sqrt{\det Q_Z}} \exp \left\{ -\frac{1}{2} (Z - \bar{Z})^* Q_Z^{-1} (Z - \bar{Z}) \right\}}{\frac{1}{(\sqrt{2\pi})^d \sqrt{\det Q_Y}} \exp \left\{ -\frac{1}{2} (Y - \bar{Y})^* Q_Y^{-1} (Y - \bar{Y}) \right\}}$$

on a

$$\begin{pmatrix} I & -Q_{XY} Q_Y^{-1} \\ 0 & I \end{pmatrix} Q_Z \begin{pmatrix} I & 0 \\ -Q_Y^{-1} Q_{XY}^* & I \end{pmatrix} = \begin{pmatrix} Q_X - Q_{XY} Q_Y^{-1} Q_{YX} & 0 \\ 0 & Q_Y \end{pmatrix}$$

donc $1 \times \det Q_Z \times 1 = \det R \times \det Q_Y$ et aussi

$$\begin{pmatrix} I & 0 \\ -Q_Y^{-1} Q_{XY}^* & I \end{pmatrix}^{-1} Q_Z^{-1} \begin{pmatrix} I & -Q_{XY} Q_Y^{-1} \\ 0 & I \end{pmatrix}^{-1} = \begin{pmatrix} R^{-1} & 0 \\ 0 & Q_Y^{-1} \end{pmatrix}$$

i.e. $Q_Z^{-1} = \begin{pmatrix} I & 0 \\ -Q_Y^{-1} Q_{XY}^* & I \end{pmatrix} \begin{pmatrix} R^{-1} & 0 \\ 0 & Q_Y^{-1} \end{pmatrix} \begin{pmatrix} I & -Q_{XY} Q_Y^{-1} \\ 0 & I \end{pmatrix}$

loi a posteriori: cas gaussien (suite)

Donc

$$(Z - \bar{Z})^* Q_Z^{-1} (Z - \bar{Z}) = (X - \hat{X})^* R^{-1} (X - \hat{X}) + (Y - \bar{Y})^* Q_Y^{-1} (Y - \bar{Y})$$

et

$$p(X | Y) = \frac{1}{(\sqrt{2\pi})^n \sqrt{\det R}} \exp \left\{ -\frac{1}{2} (X - \hat{X})^* R^{-1} (X - \hat{X}) \right\}$$

si Q_Z n'est pas inversible \rightarrow fonctions caractéristiques \blacklozenge

systeme à espace d'état

processus

- un **processus aléatoire** $\{X_k\}_{k \geq 0}$ (à valeurs dans \mathbb{R}^n) est une suite de variables aléatoires (à valeurs dans \mathbb{R}^n)
notation: $X_{0:k} = [X_0, \dots, X_k]$
- sa **loi** est du processus $\{X_k\}_{k \geq 0}$ est la donnée de $p(X_{0:k})$ pour tout k
par définition de la loi conditionnelle (cf. propriété **(P2)**), on a

$$p(X_{0:k}) = p(X_0) p(X_1|X_0) p(X_2|X_{0:1}) \cdots p(X_k|X_{0:k-1})$$

chaîne de Markov

$\{X_k\}_{k \geq 0}$ est une chaîne de Markov à valeurs dans \mathbb{R}^n , i.e. un processus aléatoire t.q.

$$p(X_{k+1}|X_{0:k}) = p(X_{k+1}|X_k) \quad \forall k \quad (\text{propriété de Markov})$$

alors

$$p(X_{0:k}) = p(X_0) \prod_{\ell=1}^k p(X_\ell|X_{\ell-1})$$

donc

la loi d'un processus de Markov est entièrement déterminée par

- (i) sa loi initiale $p(X_0)$
- (ii) son opérateur de transition $p(X_\ell|X_{\ell-1})$

chaîne de Markov: propriétés

- équation de Chapman–Kolmogorov

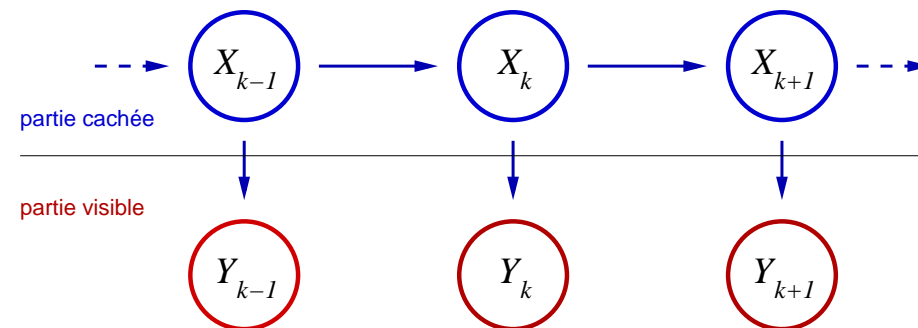
$$\begin{aligned} p(X_k) &= \int p(X_k | X_{k-1}) p(X_{k-1}) dX_{k-1} \\ &= \int dX_{k-1} p(X_k | X_{k-1}) p(X_{k-1}) \\ &\vdots \\ &= \int dX_{k-1} p(X_k | X_{k-1}) \cdots \int dX_{k-\ell} p(X_{k-\ell+1} | X_{k-\ell}) p(X_{k-\ell}) \end{aligned}$$

système à espace d'état

- **état (non observé)**: $\{X_k\}_{k \geq 0}$ une chaîne de Markov de loi initiale $p(X_0)$ et d'opérateur de transition $p(X_k|X_{k-1})$
- **observations**: $\{Y_k\}_{k \geq 1}$ observations avec une hypothèse de **canal sans mémoire** c-à-d:

$$p(Y_{1:k}|X_{0:k}) = \prod_{\ell=1}^k p(Y_\ell|X_\ell)$$

où $p(Y_\ell|X_\ell)$ désigne la “fonction de vraisemblance locale”



- les données de ce modèle sont donc: $p(X_0)$, $p(X_k|X_{k-1})$, $p(Y_k|X_k)$

système à espace d'état

loi du système

$$p(X_{0:k}, Y_{1:k}) = p(Y_{1:k}|X_{0:k}) p(X_{0:k})$$

on utilise les propriétés suivantes

$$p(Y_{1:k}|X_{0:k}) = \prod_{\ell=1}^k p(Y_{\ell}|X_{\ell}) \quad (\text{canal sans mémoire})$$

$$p(X_{0:k}) = p(X_0) \prod_{\ell=1}^k p(X_{\ell}|X_{\ell-1}) \quad (\text{Markov})$$

donc

$$p(X_{0:k}, Y_{1:k}) = p(X_0) \prod_{\ell=1}^k p(X_{\ell}|X_{\ell-1}) p(Y_{\ell}|X_{\ell})$$

(c'est toute l'information a priori que l'on possède sur ce modèle)

système à espace d'état: exemple 1

système

$$X_{k+1} = f_k(X_k) + W_k$$

$$Y_k = h_k(X_k) + V_k$$

où X_0, W_k, V_k sont indépendants et

$$W_k \sim p(W_k) = q_k^W(W_k) \quad V_k \sim p(V_k) = q_k^V(V_k)$$

système à espace d'état: exemple 1 (suite)

calcul du noyau

$$\begin{aligned}\mathbb{E}[\phi(X_k)|X_{k-1}] &= \mathbb{E}[\phi(f_{k-1}(X_{k-1}) + W_{k-1})|X_{k-1}] \\ &= \int \phi(f_{k-1}(X_{k-1}) + w) q_{k-1}^W(w) dw \\ &= \int \phi(x') q_{k-1}^W(x' - f_{k-1}(X_{k-1})) dx'\end{aligned}$$

donc

$$p(X_k|X_{k-1}) = q_{k-1}^W(X_k - f_{k-1}(X_{k-1}))$$

ystème à espace d'état: exemple 1 (suite)

calcul de la vraisemblance locale

$$\begin{aligned}\mathbb{E}[\phi(Y_k)|X_k] &= \mathbb{E}[\phi(h_k(X_k) + V_k)|X_k] \\ &= \int \phi(h_k(X_k) + v) q_k^V(v) dv \\ &= \int \phi(y) q_k^V(y - h_k(X_k)) dy\end{aligned}$$

donc

$$p(Y_k|X_k) = q_k^V(Y_k - h_k(X_k))$$

système à espace d'état: exemple 2

système

$$\begin{aligned}X_{k+1} &= F_k X_k + f_k + G_k W_k & X_0 &\sim N(\bar{X}_0, Q_0) \\Y_k &= H_k X_k + h_k + V_k\end{aligned}$$

où X_0, W_k, V_k sont indépendants et $W_k \sim N(0, Q_k^W)$, $V_k \sim N(0, Q_k^V)$, alors

$$\begin{aligned}p(X_k | X_{k-1}) &= N(F_{k-1} X_{k-1} + f_{k-1}, G_k Q_k^W G_k^*) \\p(Y_k | X_k) &= N(H_k X_k + h_k, Q_k^V)\end{aligned}$$

loi gaussienne

$(X_{0:k}, Y_{1:k})$ est gaussien car c'est une transformation linéaire de $(X_0, W_{0:k}, V_{1:k})$ qui est gaussien.

filtrage non linéaire

le filtre

on pose

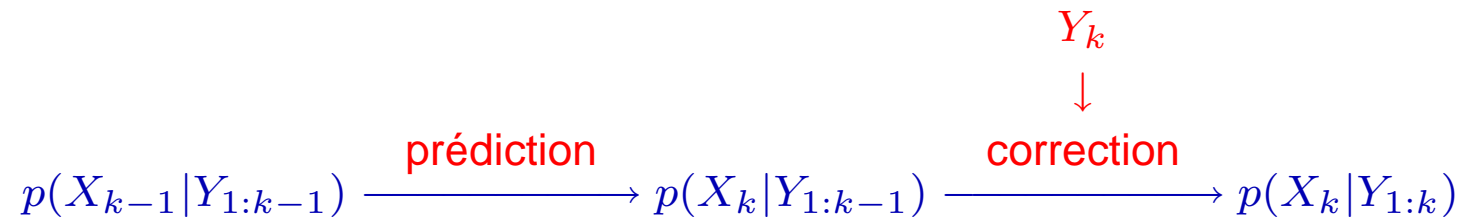
$$p(X_k | Y_{1:k})$$

(filtre)

$$p(X_k | Y_{1:k-1})$$

(filtre prédit)

le schéma classique est



prédiction

une simple application de l'équation de Chapman–Kolmogorov

$$\begin{aligned} p(X_k | Y_{1:k-1}) &= \int p(X_k | X_{k-1}, Y_{1:k-1}) p(X_{k-1} | Y_{1:k-1}) dX_{k-1} \\ &= \int p(X_k | X_{k-1}) p(X_{k-1} | Y_{1:k-1}) dX_{k-1} \quad \blacklozenge \end{aligned}$$

correction

avec la formule de Bayes on a les deux décompositions

$$(\blacktriangle) \quad p(X_{0:k}, Y_{1:k}) = p(X_{0:k} | Y_{1:k}) p(Y_{1:k})$$

$$(\blacktriangle\blacktriangle) \quad p(X_{0:k}, Y_{1:k}) = p(Y_k | X_{0:k}, Y_{1:k-1}) p(X_{0:k}, Y_{1:k-1})$$

et (tj avec Bayes)

$$\begin{aligned} (\blacktriangle) \quad & p(X_{0:k} | Y_{1:k}) p(Y_{1:k}) \\ &= p(X_{0:k} | Y_{1:k}) p(Y_k | Y_{1:k-1}) p(Y_{1:k-1}) \end{aligned}$$

$$\begin{aligned} (\blacktriangle\blacktriangle) \quad & p(Y_k | X_{0:k}, Y_{1:k-1}) p(X_{0:k}, Y_{1:k-1}) \\ &= p(Y_k | X_k) p(X_{0:k} | Y_{1:k-1}) p(Y_{1:k-1}) \end{aligned}$$

on pose $(\blacktriangle) = (\blacktriangle\blacktriangle)$, on simplifie par $p(Y_{1:k-1})$, on intègre en $\int \cdots \int dX_{0:k-1}$, on obtient:

$$p(X_k | Y_{1:k}) p(Y_k | Y_{1:k-1}) = p(Y_k | X_k) p(X_k | Y_{1:k-1})$$

correction (suite)

donc

$$\begin{aligned} p(X_k | Y_{1:k}) &= \frac{p(Y_k | X_k) p(X_k | Y_{1:k-1})}{p(Y_k | Y_{1:k-1})} \\ &= \frac{p(Y_k | X_k) p(X_k | Y_{1:k-1})}{\int p(Y_k | X_k, Y_{1:k-1}) p(X_k | Y_{1:k-1}) dX_k} \\ &= \frac{p(Y_k | X_k) p(X_k | Y_{1:k-1})}{\int p(Y_k | X_k) p(X_k | Y_{1:k-1}) dX_k} \quad \blacklozenge \end{aligned}$$

filtre non linéaire

- prédiction $p(X_{k-1}|Y_{1:k-1}) \rightarrow p(X_k|Y_{1:k-1})$

$$p(X_k|Y_{1:k-1}) = \int p(X_k|X_{k-1}) p(X_{k-1}|Y_{1:k-1}) dX_{k-1}$$

- correction $p(X_k|Y_{1:k-1}) \rightarrow p(X_k|Y_{1:k})$

$$p(X_k|Y_{1:k}) = \frac{p(Y_k|X_k) p(X_k|Y_{1:k-1})}{\int p(Y_k|X_k) p(X_k|Y_{1:k-1}) dX_k}$$

- utilisation: $\mathbb{E}[\phi(X_k)|Y_{1:k}] = \int \phi(X_k) p(X_k|Y_{1:k}) dX_k$

$$p(X_{k-1}|Y_{1:k-1}) \xrightarrow[p(X_k|X_{k-1})]{\text{prédiction}} p(X_k|Y_{1:k-1}) \xrightarrow[p(Y_k|X_k)]{\text{correction}} p(X_k|Y_{1:k})$$

Y_k
 \downarrow

filtre non linéaire (suite)

deux cas:

- on peut intégrer explicitement ces équations → (presque) uniquement dans le cas **linéaire/gaussien**
- on ne sait pas intégrer explicitement ces équations → alors **c'est dur...**
 - kalman étendu
 - simple à mettre en œuvre
 - problèmes de robustesse
 - éléments finis, différences finis
 - robuste
 - très contraignant et difficile à mettre en œuvre
 - uniquement pour les dimensions petites ($n \lesssim 4$)
 - méthodes particulières
 - simple à mettre en œuvre (+ simple que les kalmaneries)
 - robuste

filtrage de Kalman

modèle linéaire gaussien

modèle à espace d'état

$$X_{k+1} = F_k X_k + f_k + G_k W_k \quad (\text{éq. état})$$

$$Y_k = H_k X_k + h_k + V_k \quad (\text{éq. obs.})$$

$X_0 \sim N(\bar{X}_0, Q_0)$, $W_k \sim N(0, Q_k^W)$, $V_k \sim N(0, Q_k^V)$ indépendants — hyp: $Q_k^V > 0$

filtre

$(X_{0:k}, Y_{1:k})$ est gaussien donc:

$$\text{loi}(X_k | Y_{1:k}) = N(\hat{X}_k, R_k) \quad \text{où} \begin{cases} \hat{X}_k = \mathbb{E}[X_k | Y_{1:k}] \\ R_k = \mathbb{E}[(X_k - \hat{X}_k)(X_k - \hat{X}_k)^*] \end{cases}$$

$$\text{loi}(X_k | Y_{1:k-1}) = N(\hat{X}_k^-, R_k^-) \quad \text{où} \begin{cases} \hat{X}_k^- = \mathbb{E}[X_k | Y_{1:k-1}] \\ R_k^- = \mathbb{E}[(X_k - \hat{X}_k^-)(X_k - \hat{X}_k^-)^*] \end{cases}$$

les covariances R_k^- , R_k sont déterministes

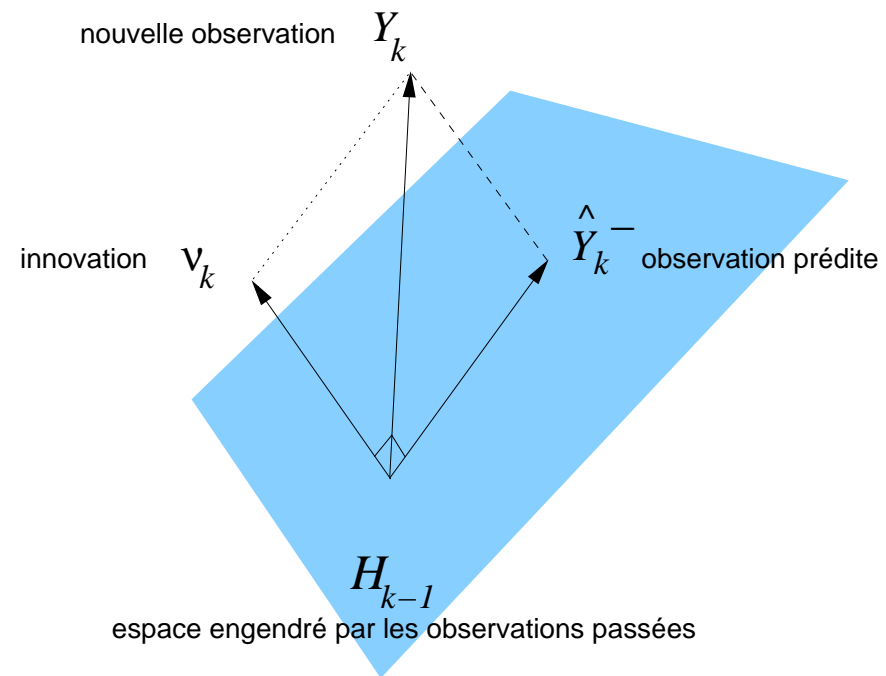
idée

si $\text{loi}(X_{k-1} | Y_{1:k-1})$ est connue, on détermine $\text{loi}(X_k | Y_{1:k})$ en deux étapes:

- *prédiction* — on calcule $\text{loi}(X_k | Y_{1:k-1})$ à l'aide de l'équation d'état
 - *correction* — on corrige la prédiction en tenant compte de la nouvelle observation Y_k et de l'équation d'observation
- qu'apporte la nouvelle observation Y_k par rapport aux observations passées $Y_{1:k-1}$?

innovation

$$\nu_k = Y_k - \mathbb{E}[Y_k | Y_{1:k-1}] = Y_k - \hat{Y}_k^-$$



innovation (suite)

$$\nu_k = Y_k - \left(H_k \mathbb{E}[X_k | Y_{1:k-1}] + h_k + \underbrace{\mathbb{E}[V_k | Y_{1:k-1}]}_{=0} \right) = Y_k - (H_k \hat{X}_k^- + h_k)$$

l'innovation ν_k vérifie

- $\nu_k \sim N(0, Q_k^\nu)$ avec $Q_k^\nu = H_k R_k^- H_k^* + Q_k^V$
- $\nu_k \perp\!\!\!\perp Y_{1:k-1}$

donc

- information $Y_1 \cdots Y_{k-1} Y_k = \text{information } Y_1 \cdots Y_{k-1} \nu_k$
- $\nu_k = \text{information "fraîche" contenue dans } Y_k$

preuve: $\nu_k = Y_k - (H_k \hat{X}_k^- + h_k) = H_k (X_k - \hat{X}_k^-) + V_k$ ♦

- $\mathbb{E}[(X_k - \hat{X}_k^-) \nu_k^*] = R_k^- H_k^*$

filtre de Kalman

initialisation	$\hat{X}_0 = \bar{X}_0 = \mathbb{E}[X_0]$ $R_0 = Q_0 = \text{cov}(X_0)$
prédiction	$\hat{X}_k^- = F_{k-1} \hat{X}_{k-1} + f_{k-1}$ $R_k^- = F_{k-1} R_{k-1} F_{k-1}^* + G_{k-1} Q_{k-1}^W G_{k-1}^*$
correction	$K_k = R_k^- H_k^* [H_k R_k^- H_k^* + Q_k^V]^{-1}$ (gain) $\hat{X}_k = \hat{X}_k^- + K_k [\mathbf{Y}_k - (H_k \hat{X}_k^- + h_k)]$ $R_k = [I - K_k H_k] R_k^-$

- on suppose $Q_k^V > 0$
- les covariances ne dépendent pas des observations
- c'est un théorème et un algorithme

filtre de Kalman (suite)

preuve de la prédiction: $(X_k, Y_1, \dots, Y_{k-1})$ est gaussien

$\Rightarrow \text{loi}(X_k | Y_{1:k-1}) = N(\hat{X}_k^-, R_k^-)$ avec:

$$\hat{X}_k^- = \mathbb{E}[X_k | Y_{1:k-1}] = F_{k-1} \underbrace{\mathbb{E}[X_{k-1} | Y_{1:k-1}]}_{=\hat{X}_{k-1}} + f_{k-1} + G_{k-1} \underbrace{\mathbb{E}[W_{k-1} | Y_{1:k-1}]}_{=0}$$

comme $X_k - \hat{X}_k^- = F_{k-1} (X_{k-1} - \hat{X}_{k-1}) + G_{k-1} W_{k-1}$ on a:

$$\begin{aligned} R_k^- &= \mathbb{E}[(X_k - \hat{X}_k^-) (X_k - \hat{X}_k^-)^*] \\ &= \mathbb{E}[(F_{k-1} (X_{k-1} - \hat{X}_{k-1}) + G_{k-1} W_{k-1}) (\quad)^*] \\ &= F_{k-1} R_{k-1} F_{k-1}^* + G_{k-1} Q_{k-1}^W G_{k-1}^* \end{aligned}$$

(car $(X_k - \hat{X}_k^-) \perp W_{k-1}$ donc $\mathbb{E}[(X_{k-1} - \hat{X}_{k-1}) W_{k-1}^*] = 0$) ♦

filtre de Kalman (suite)

preuve de la correction: (X_k, Y_1, \dots, Y_k) est gaussien $\Rightarrow \text{loi}(X_k | Y_{1:k}) = N(\hat{X}_k, R_k)$
avec:

$$\begin{aligned}\hat{X}_k &= \mathbb{E}[X_k | Y_{1:k}] = \hat{X}_k^- + \mathbb{E}[X_k - \hat{X}_k^- | Y_{1:k}] \\ &= \hat{X}_k^- + \underbrace{\mathbb{E}[X_k - \hat{X}_k^- | Y_{1:k-1}]}_{=0} + \mathbb{E}[X_k - \hat{X}_k^- | \nu_k]\end{aligned}$$

donc $X_k - \hat{X}_k = (X_k - \hat{X}_k^-) - \underbrace{(\hat{X}_k - \hat{X}_k^-)}_{=\mathbb{E}[X_k - \hat{X}_k^- | \nu_k]}$

on calcule la moyenne et covariance de $(X_k - \hat{X}_k^-) | \nu_k$.

$(X_k - \hat{X}_k^-, \nu_k)$ est gaussien, centré de covariance:

$$Q_k^\nu = \begin{pmatrix} R_k^- & R_k^- H_k^* \\ H_k R_k^- & H_k R_k^- H_k^* + Q_k^V \end{pmatrix}$$

filtre de Kalman (suite)

Q_k^V inversible $\Rightarrow Q_k^\nu$ inversible
d'après le lemme

$$\hat{X}_k = \hat{X}_k^- + R_k^- H_k^* [H_k R_k^- H_k^* + Q_k^V]^{-1} \nu_k$$

$$R_k = R_k^- - R_k^- H_k^* [H_k R_k^- H_k^* + Q_k^V]^{-1} H_k R_k^- \quad \blacklozenge$$

filtre de Kalman étendu

idée

modèle à espace d'état

$$X_{k+1} = f(X_k) + g(X_k) W_k \quad (\text{éq. état})$$

$$Y_k = h(X_k) + V_k \quad (\text{éq. obs.})$$

idée

on **linéarise** (autour de l'estimée courante) et rend tout **gaussien**

initialisation

on pose:

$$\hat{X}_0 = \mathbb{E}[X_0]$$

$$R_0 = \text{cov}(X_0)$$

(on n'a pas nécessairement X_0 gaussien)

prédiction

on “linéarise” l’équation d’état autour de \hat{X}_{k-1} :

$$\begin{aligned} X_k &= f(X_{k-1}) + g(X_{k-1}) W_{k-1} \\ &\simeq f(\hat{X}_{k-1}) + \underbrace{\nabla f(\hat{X}_{k-1}) (X_{k-1} - \hat{X}_{k-1})}_{F_{k-1}} + \underbrace{g(\hat{X}_{k-1})}_{G_{k-1}} W_{k-1} \end{aligned}$$

donc

$$\hat{X}_k^- = \mathbb{E}[X_k | Y_{1:k-1}] \simeq f(\hat{X}_{k-1})$$

et

$$R_k^- = \mathbb{E}[(X_k - \hat{X}_k^-) (X_k - \hat{X}_k^-)^*] \simeq F_{k-1}^* R_{k-1} F_{k-1} + G_{k-1} Q_{k-1}^W G_{k-1}^*$$

correction

on “linéarise” l’équation d’observation autour de \hat{X}_k^- :

$$Y_k = h(X_k) + V_k \simeq h(\hat{X}_k^-) + \underbrace{\nabla h(\hat{X}_k^-)}_{H_k} (X_k - \hat{X}_k^-) + V_k$$

$$K_k = R_k^- H_k^* [H_k R_k^- H_k^* + Q_k^V]^{-1}$$

$$\hat{X}_k = \hat{X}_k^- + K_k [Y_k - h(\hat{X}_k^-)]$$

$$R_k = [I - K_k H_k] R_k^-$$

fke

initialisation	$\hat{X}_0 = \bar{X}_0 = \mathbb{E}[X_0]$ $R_0 = Q_0 = \text{cov}(X_0)$
prédiction	$\hat{X}_k^- = f(\hat{X}_{k-1})$ $R_k^- = F_{k-1} R_{k-1} F_{k-1}^* + G_{k-1} Q_{k-1}^W G_{k-1}^*$ avec $F_{k-1} = \nabla f(\hat{X}_{k-1})$
correction	$K_k = R_k^- H_k^* [H_k R_k^- H_k^* + Q_k^V]^{-1}$ (gain) avec $H_k = \nabla h(\hat{X}_k^-)$ $\hat{X}_k = \hat{X}_k^- + K_k [\mathbf{Y}_k - h(\hat{X}_k^-)]$ $R_k = [I - K_k H_k] R_k^-$

- c'est un algorithme, *pas* un théorème
- les “covariances” *dépendent* des observations
- sont-ce des covariances ?

références (Kalman/Kalman étendu)

- Rudolf E. Kalman, A new approach to linear filtering and prediction problems, *Transactions of the ASME—Journal of Basic Engineering*, 82, Series D, 35-45, 1960.
<http://www.cs.unc.edu/~welch/media/pdf/Kalman1960.pdf>
- Rudolf E. Kalman and Richard S. Bucy, New results in linear filtering and prediction theory, *Transactions of the ASME—Journal of Basic Engineering*, 83, Series D, 95–108, 1961.
- A. H. Jazwinski, *Stochastic processes and filtering*, Academic Press, 1970.
- Arthur Gelb (ed), *Applied Optimal Estimation*, MIT Press, 1974.
- Brian D.O. Anderson and John B. Moore, *Optimal Filtering*, Prentice-Hall, 1978.
- Peter S. Maybeck, *Stochastic Models, Estimation and Control*, 3 volumes, Academic Press, 1979 à 1982.
- le site de Greg Welch: <http://www.cs.unc.edu/~welch/kalman/>

Monte Carlo

le principe

- estimer

$$\mathcal{I} = \int_{\mathbb{R}^n} \phi(x) q(x) dx$$

où $q(x)$ est la densité d'une loi $Q(dx)$ sur \mathbb{R}^n et $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$

- idée: $\mathcal{I} = \mathbb{E}\phi(X)$ où $X \sim Q(dx) = q(x) dx$
- si $\xi^1 \dots \xi^N \stackrel{\text{iid}}{\sim} Q$ alors $\mathcal{I} \simeq \mathcal{I}^N = \frac{1}{N} \sum_{i=1}^N \phi(\xi^i)$
 - convergence (loi des grands nombres): $\mathcal{I}^N \rightarrow \mathcal{I}$ p.s.
 - vitesse de convergence (théorème central limite): $\frac{\sqrt{N}}{\sigma}(\mathcal{I}^N - \mathcal{I}) \Longrightarrow N(0, 1)$
où $\sigma^2 = \text{var}(Q)$

Nicholas Metropolis and Stanislaw Ulam (1949). The Monte Carlo method, *Journal of the American Statistical Association*, 44 (247), 335-341.

convergence

$$\mathcal{I} \simeq \mathcal{I}^N = \frac{1}{N} \sum_{i=1}^N \phi(\xi^i) \quad \text{où} \quad \xi^1 \dots \xi^N \stackrel{\text{iid}}{\sim} Q$$

- convergence (loi des grands nombres): $\mathcal{I}^N \rightarrow \mathcal{I}$ p.s.
- vitesse de convergence (théorème central limite): $\frac{\sqrt{N}}{\sigma}(\mathcal{I}^N - \mathcal{I}) \Rightarrow N(0, 1)$ où $\sigma^2 = \text{var}(Q)$

$$\mathcal{I} \simeq \mathcal{I}^N \pm \sigma \frac{1.96}{\sqrt{N}} \quad \text{avec probabilité} \simeq 0.95$$

$$\text{et } \sigma^2 \simeq \frac{1}{N} \sum_{i=1}^N \phi^2(\xi^i) - \left(\frac{1}{N} \sum_{i=1}^N \phi(\xi^i) \right)^2$$

- vitesse de convergence $\frac{1}{\sqrt{N}}$ faible (un gain en précision de facteur 2 nécessite 4 fois plus de simulations)
- ne dépend pas de la dimension n du problème

deux points importants

on ne peut rien faire à la vitesse $\frac{1}{\sqrt{N}}$ mais:

- comment simuler efficacement la loi Q ?
- peut-on diminuer σ^2 ? (réduction de variance)

simuler

«générer une suite de nombres $\xi^1 \dots \xi^N$ qui “ressemblent” (pseudo-aléatoire) à N réalisations indépendantes de la loi Q » ?

loi uniforme

$$\xi^{1:N} \stackrel{\text{iid}}{\sim} U\{1, \dots, M\}$$

$$\xi^{1:N} \stackrel{\text{iid}}{\sim} U[0, 1]$$

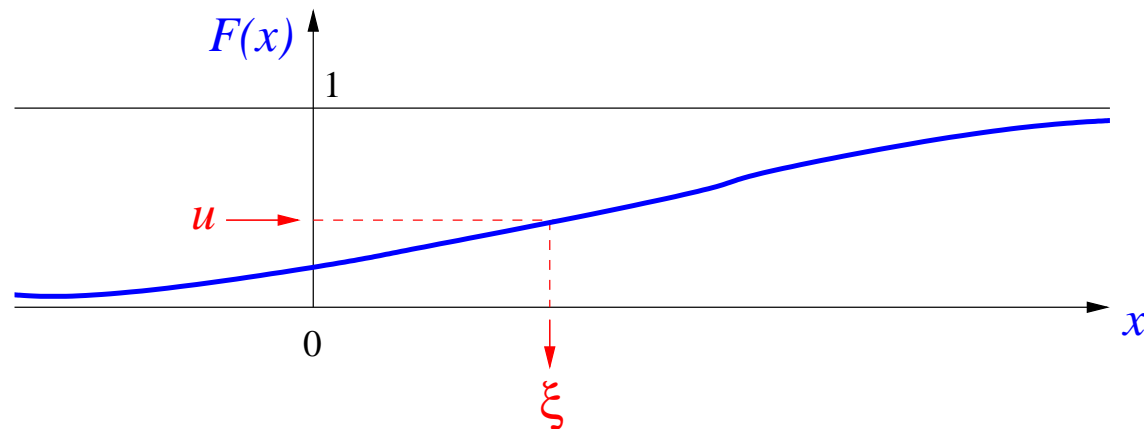
→ présent sur toute machine

simuler (suite)

méthode la fonction de répartition

$$F(x) = \int_{-\infty}^x Q(dy) = \mathbb{P}(X \leq x)$$

on génère $\xi^i = F^{-1}(u^i)$ avec $\begin{cases} u^{1:N} \stackrel{\text{iid}}{\sim} U[0, 1] \\ F^{-1}(u) = \inf\{x \in \mathbb{R}; F(x) \geq u\} \end{cases}$



simuler (suite)

application: la loi multinomiale $\mathbb{P}(X = x^i) = \omega^i$ (avec $\omega^i \geq 0$ et $\sum_{i=1}^N \omega^i = 1$) i.e.

$$\text{loi}(X) = Q(dx) = \sum_{i=1}^N \omega^i \delta_{x^i}(dx)$$

```
pour  $i = 1 : N$  faire  
   $u \sim U[0, 1]$   
   $j \leftarrow 1$   
  tant que  $\omega^1 + \dots + \omega^j < u$  faire  
     $j \leftarrow j + 1$   
  fin tant que  
   $\xi^i \leftarrow x^j$   
fin pour
```

filtrage particulaire

idée

$$p(X_k | Y_{1:k-1}) \simeq p^N(X_k | Y_{1:k-1}) = \sum_{i=1}^N \omega_{k-}^i \delta_{\xi_{k-}^i}(X_k)$$

$$p(X_k | Y_{1:k}) \simeq p^N(X_k | Y_{1:k}) = \sum_{i=1}^N \omega_k^i \delta_{\xi_k^i}(X_k)$$

- évaluer un nombre fini de paramètres
- certaines tâches sont simplifiées, e.g. $\mathbb{E}[\phi(X_k) | Y_{1:k}] \simeq \sum_{i=1}^N \omega_k^i \phi(\xi_k^i)$

prédiction

supposons que $p(X_{k-1}|Y_{1:k-1}) = \sum_{i=1}^N \omega_{k-1}^i \delta_{\xi_{k-1}^i}(X_{k-1})$ que devient $p(X_k|Y_{1:k-1})$?

$$\begin{aligned} p(X_k|Y_{1:k-1}) &= \int p(X_k|X_{k-1}) p(X_{k-1}|Y_{1:k-1}) dX_{k-1} \\ &= \sum_{i=1}^N \omega_{k-1}^i \int p(X_k|X_{k-1}) \delta_{\xi_{k-1}^i}(X_{k-1}) dX_{k-1} \\ &= \sum_{i=1}^N \omega_{k-1}^i p(X_k|X_{k-1} = \xi_{k-1}^i) \end{aligned}$$

on obtient donc un mélange des lois $p(X_k|X_{k-1} = \xi_{k-1}^i)$ (qui n'est pas sous forme particulière), on peut échantillonner selon cette loi, ou bien poser:

$$p(X_k|Y_{1:k-1}) = \sum_{i=1}^N \omega_{k-1}^i \delta_{\xi_{k-1}^i}(X_k) \quad \text{où} \quad \xi_{k-1}^i \sim p(X_k|X_{k-1} = \xi_{k-1}^i)$$

correction

supposons que $p(X_k|Y_{1:k-1}) = \sum_{i=1}^N \omega_{k-}^i \delta_{\xi_{k-}^i}(X_k)$ que devient $p(X_k|Y_{1:k})$?

$$\begin{aligned} p(X_k|Y_{1:k}) &= \frac{p(Y_k|X_k) p(X_k|Y_{1:k-1})}{\int p(Y_k|X_k) p(X_k|Y_{1:k-1}) dX_k} \\ &= \sum_{i=1}^N \frac{\omega_{k-}^i p(Y_k|X_k)}{\sum_{j=1}^N \int \omega_{k-}^j p(Y_k|X_k) \delta_{\xi_{k-}^j}(X_k) dX_k} \delta_{\xi_{k-}^i}(X_k) \\ &= \sum_{i=1}^N \frac{\omega_{k-}^i p(Y_k|X_k = \xi_{k-}^i)}{\sum_{j=1}^N \omega_{k-}^j p(Y_k|X_k = \xi_{k-}^j)} \delta_{\xi_{k-}^i}(X_k) = \sum_{i=1}^N \omega_k^i \delta_{\xi_{k-}^i}(X_k) \end{aligned}$$

avec

$$\omega_k^i = \frac{\omega_{k-}^i p(Y_k|X_k = \xi_{k-}^i)}{\sum_{j=1}^N \omega_{k-}^j p(Y_k|X_k = \xi_{k-}^j)}$$

c'est bon: $p(X_k|Y_{1:k})$ est particulière

un premier filtre particulaire

- prédiction

$$p(X_k | Y_{1:k-1}) = \sum_{i=1}^N \omega_{k-}^i \delta_{\xi_{k-}^i}(X_k) \quad \text{où} \quad \begin{cases} \omega_{k-}^i = \omega_{k-1}^i \\ \xi_{k-}^i \sim p(X_k | X_{k-1} = \xi_{k-1}^i) \end{cases}$$

- correction

$$p(X_k | Y_{1:k}) = \sum_{i=1}^N \omega_k^i \delta_{\xi_k^i}(X_k) \quad \text{où} \quad \begin{cases} \omega_k^i = \frac{\omega_{k-}^i p(Y_k | X_k = \xi_{k-}^i)}{\sum_{j=1}^N \omega_{k-}^j p(Y_k | X_k = \xi_{k-}^j)} \\ \xi_k^i = \xi_{k-}^i \end{cases}$$

dégénérescence des poids

- pb: en quelques itérations en k , tous les ω_k^i sont nuls sauf un
- la variance des poids doit rester petite, idéalement $\omega_k^i = \frac{1}{N}$
- critère

$$N_k^{\text{eff}} = \frac{1}{\sum_{i=1}^N (\omega_k^i)^2} \in [1, N] \quad \left\{ \begin{array}{ll} \simeq 1 & \text{mauvais} \\ \simeq N & \text{bon} \end{array} \right.$$

i.e. il faut faire quelque chose dès que N_k^{eff} est trop petit

→ redistribution

redistribution

on tire au hasard selon la loi multinomiale:

$$\sum_{i=1}^N \omega_k^i \delta_{\xi_k^i}(X)$$

i.e. on choisit au hasard des points de $\{\xi_k^1, \dots, \xi_k^N\}$ selon la loi $\{\omega_k^1, \dots, \omega_k^N\}$, on obtient de nouveaux points de poids $\frac{1}{N}$

$$\xi_k^{1:N} \leftarrow \text{resample}(\xi_k^{1:N}, \omega_k^{1:N})$$

filtre bootstrap

```
 $\xi^{1:N} \stackrel{\text{iid}}{\sim} p(X_0)$   
 $\omega^{1:N} \leftarrow 1/N$   
pour  $k = 1, 2, 3 \dots$  faire  
   $\tilde{\xi}^i \sim p(X_k | X_{k-1} = \xi^i)$  pour  $i = 1 : N$  % mutation  
   $\tilde{\omega}^i \leftarrow \omega^i p(Y_k | X_k = \tilde{\xi}^i)$  pour  $i = 1 : N$  % vraisemblance  
   $\tilde{\omega}^i \leftarrow \tilde{\omega}^i / \text{sum}(\tilde{\omega}^{1:N})$  pour  $i = 1 : N$  % normalisation  
   $N^{\text{eff}} \leftarrow (\sum_{i=1}^N (\tilde{\omega}^i)^2)^{-1}$   
  si  $N^{\text{eff}}/N \leq 0.75$  alors  
     $\xi^{1:N} \leftarrow \text{resample}(\tilde{\omega}^{1:N}, \tilde{\xi}^{1:N})$  % sélection  
     $\omega^{1:N} \leftarrow 1/N$   
  sinon  
     $\xi^{1:N} \leftarrow \tilde{\xi}^{1:N}$   
     $\omega^{1:N} \leftarrow \tilde{\omega}^{1:N}$   
  fin si  
  sortie  $(\xi^{1:N}, \omega^{1:N})$   
fin pour
```

filtre bootstrap (suite)

idem avec redistribution systématique

```
 $\xi^{1:N} \stackrel{\text{iid}}{\sim} p(X_0)$  % initialisation
pour  $k = 1, 2, 3 \dots$  faire
   $\xi^i \sim p(X_k | X_{k-1} = \xi^i)$  pour  $i = 1 : N$  % mutation
   $\omega^i \leftarrow p(Y_k | X_k = \xi^i)$  pour  $i = 1 : N$  % vraisemblance
   $\omega^i \leftarrow \omega^i / \text{sum}(\omega^{1:N})$  pour  $i = 1 : N$  % normalisation
   $\xi^{1:N} \leftarrow \text{resample}(\omega^{1:N}, \xi^{1:N})$  % sélection
  sortie  $\xi^{1:N}$ 
fin pour
```

redistribution

la procédure `resample`($\tilde{\omega}^{1:N}, \tilde{\xi}^{1:N}$) est trop lente \rightarrow à repenser

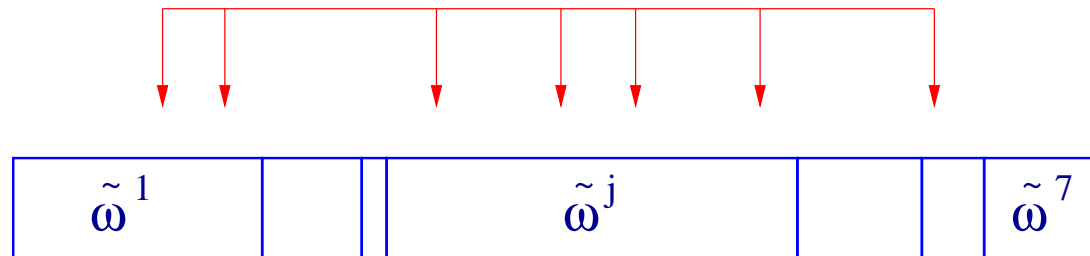
$$\xi^{1:N} \leftarrow \text{resample}(\tilde{\omega}^{1:N}, \tilde{\xi}^{1:N})$$

redistribution (suite)

rééchantillonnage multinomial

$$u^1, \dots, u^N \stackrel{\text{iid}}{\sim} U[0, 1]$$

u^i



$$\xi^i = \xi^j$$

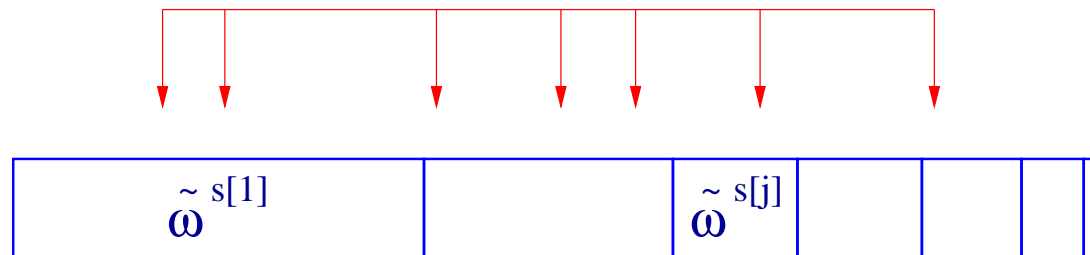
$$O(N \log(N))$$

redistribution (suite)

rééchantillonnage multinomial (bis)

$$u^1, \dots, u^N \stackrel{\text{iid}}{\sim} U[0, 1]$$
$$\tilde{\omega}^{s[1]} > \tilde{\omega}^{s[2]} > \dots > \tilde{\omega}^{s[N]}$$

u^i



$$\xi^i = \tilde{\xi}^{s[j]}$$

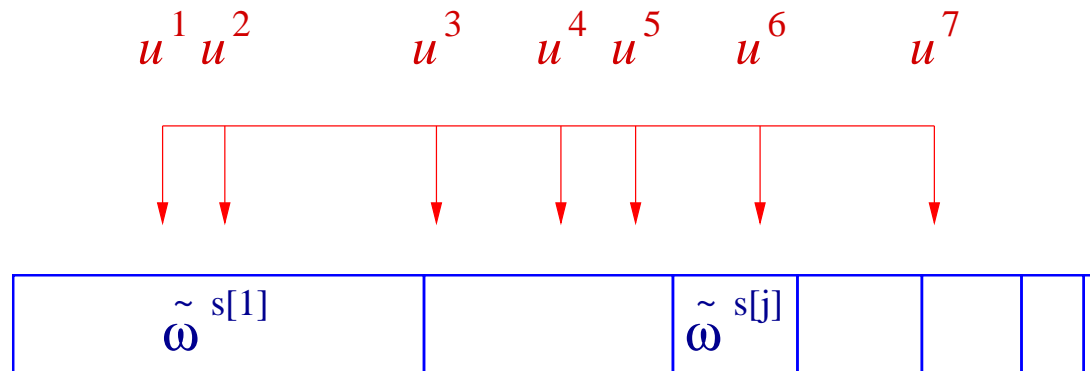
$O(N \log(N))$ mais plus rapide (moins de tests)

redistribution (suite)

rééchantillonnage multinomial (ter)

statistique d'ordre $u^1 < u^2 < \dots < u^N \stackrel{\text{iid}}{\sim} U[0, 1]$

$$\tilde{\omega}^{s[1]} > \tilde{\omega}^{s[2]} > \dots > \tilde{\omega}^{s[N]}$$



$$\xi^i = \tilde{\xi}^{s[j]}$$

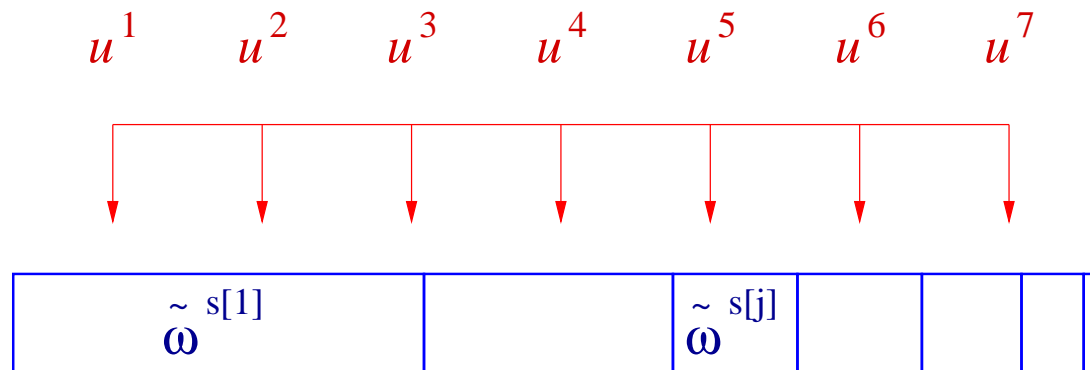
$O(N)$ mais il faut savoir simuler des statistiques d'ordre

redistribution (suite)

le «peigne de Kitagawa»

$$u_1 \sim U[0, \frac{1}{N}] \quad u^i = u^1 + \frac{i}{N}, \quad i = 2 : N$$

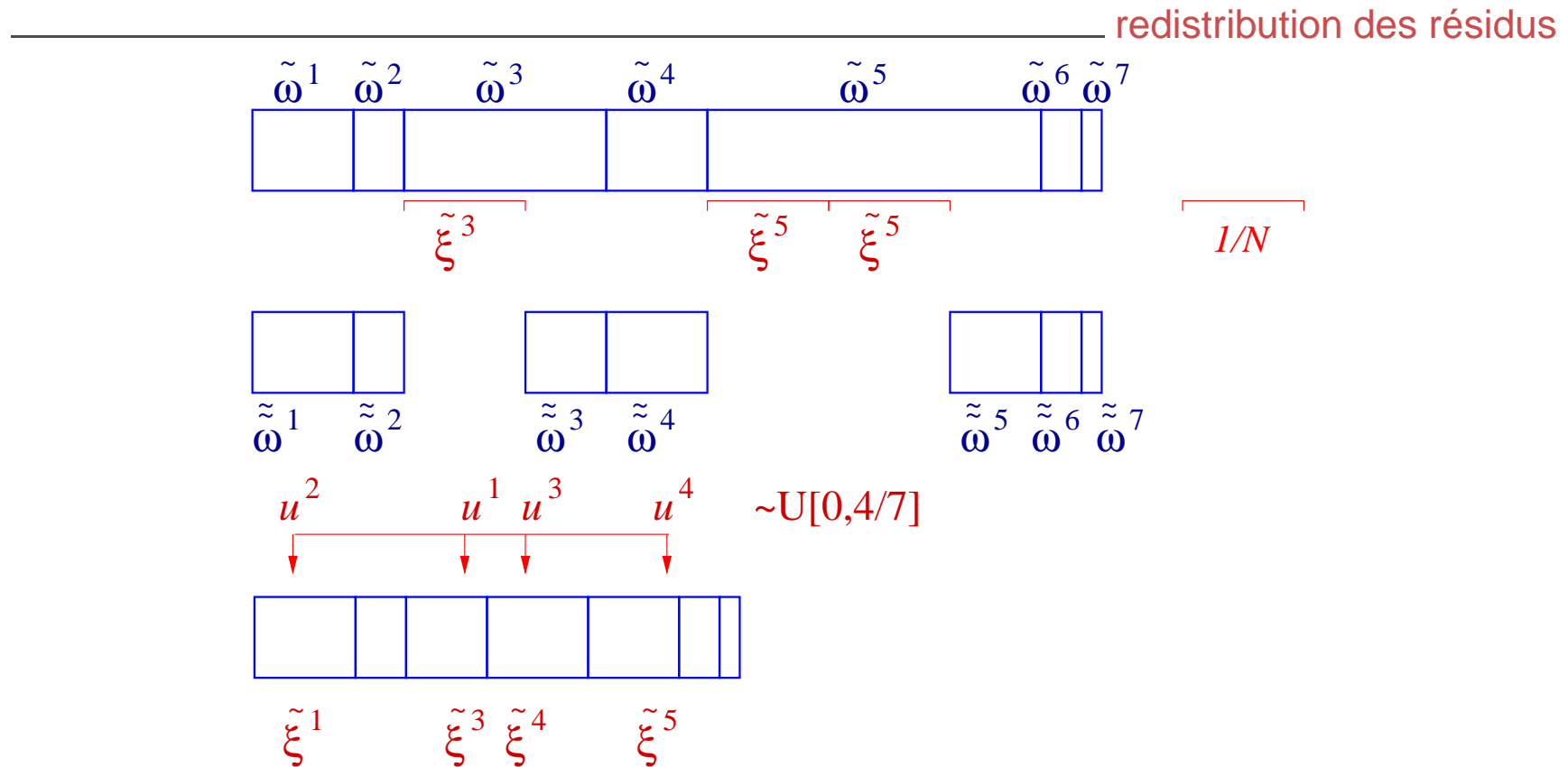
$$\tilde{\omega}^s[1] > \tilde{\omega}^s[2] > \dots > \tilde{\omega}^s[N]$$



$$\xi^i = \tilde{\xi}^s[j]$$

$O(N)$ mais ça ne marche pas toujours !

redistribution (suite)



n'a d'intérêt que si la plupart des particules sont sélectionnée dans la première phase déterministe: c'est le cas quand N_k^{eff} est petit

ce qu'il faut...

...savoir faire

- simuler selon $p(X_0)$
- simuler selon $p(X_{k+1} | X_k = \xi)$ quelque soit ξ
- calculer la vraisemblance $p(Y_k | X_k)$ pour tout X_k (à Y_k fixé)

...utiliser

- une routine de rééchantillonnage

extensions

- les contraintes: $X_k \in D_k$
- beaucoup de variantes: rao–blackwellisation
- modèles hybrides: espaces continus/discrets/finis
- on n'est pas obligé de simuler selon $p(X_{k+1}|X_k)$: si on simule selon $q(X_{k+1})$, il faudra alors modifier la fonction de vraisemblance par $q(Y_k|X_k)$ (\rightarrow échantillonnage d'importance). À quoi cela sert ?
 - simuler selon $p(X_{k+1}|X_k)$ est très coûteux
 - on ne sait pas simuler selon $p(X_{k+1}|X_k)$

les contraintes

supposons que l'on sache que $X_k \in D_k$ (pour tout k)

```
 $\xi^{1:N} \stackrel{\text{iid}}{\sim} p(X_0)$  % initialisation
pour  $k = 1, 2, 3 \dots$  faire
   $\xi^i \sim p(X_k | X_{k-1} = \xi^i)$  pour  $i = 1 : N$  % mutation
   $\omega^i \leftarrow \mathbf{1}_{D_k}(\xi^i) p(Y_k | X_k = \xi^i)$  pour  $i = 1 : N$  % vraisemblance
   $\omega^i \leftarrow \omega^i / \text{sum}(\omega^{1:N})$  pour  $i = 1 : N$  % normalisation
   $\xi^{1:N} \leftarrow \text{resample}(\omega^{1:N}, \xi^{1:N})$  % sélection
  sortie  $\xi^{1:N}$ 
fin pour
```

attention !

un seul problème peut survenir: à un instant k , toutes les vraisemblances sont nulles, i.e.

$$p(Y_k | X_k = \xi^i) = 0 \quad \forall i$$

i.e. les particules ne correspondent plus à l'observation

i.e. le filtre est perdu

idée

- réinitialiser le filtre
- affiner l'étape de prédiction (par exemple simuler en tenant compte de X_k et Y_{k+1})

attention ! (suite)

```
 $\xi^{1:N} \stackrel{\text{iid}}{\sim} p(X_0)$  % initialisation
pour  $k = 1, 2, 3 \dots$  faire
   $\xi^i \sim p(X_k | X_{k-1} = \xi^i)$  pour  $i = 1 : N$  % mutation
   $\omega^i \leftarrow p(Y_k | X_k = \xi^i)$  pour  $i = 1 : N$  % vraisemblance
  si  $\text{sum}(\omega^{1:N}) \neq 0$  alors
     $\omega^i \leftarrow \omega^i / \text{sum}(\omega^{1:N})$  pour  $i = 1 : N$  % normalisation
     $\xi^{1:N} \leftarrow \text{resample}(\omega^{1:N}, \xi^{1:N})$  % sélection
  sinon
     $\omega^i \leftarrow 1/N$  pour  $i = 1 : N$  % reset
  fin si
  sortie  $\xi^{1:N}$ 
fin pour
```

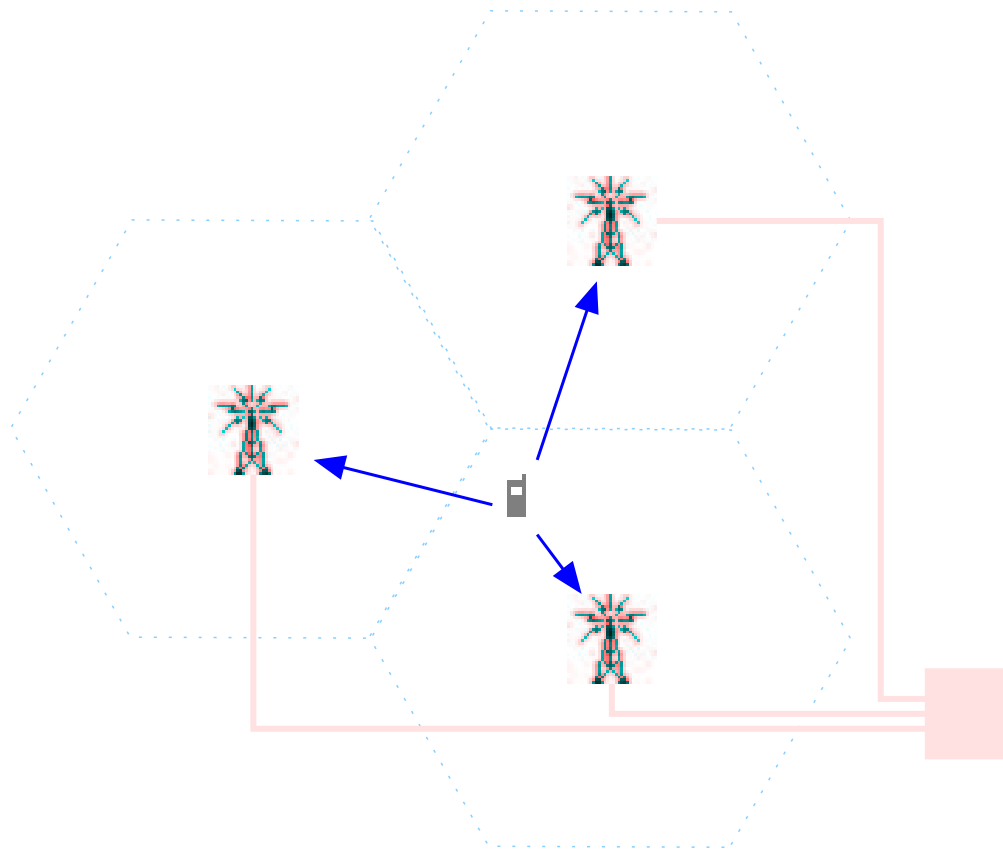
références (filtrage particulaire)

- Neil J. Gordon, David J. Salmond, Adrian F. M. Smith, Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings, Part F*, 140(2): 107–113, 1993.
- Arnaud Doucet, Nando de Freitas, Niel J. Gordon (eds), *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, 2001.
- S.M. Arulampalam, S. Maskell, N.J. Gordon, T.C. Clapp. A tutorial on particle filters for online nonlinear / non-Gaussian Bayesian tracking, *IEEE Transactions on Signal Processing SP-50*, 2 (special issue on Monte Carlo Methods for Statistical Signal Processing), 174-188, 2002.
- B. Ristic, S. Arulampalam, N. Gordon, *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House, 2004.

géolocalisation

introduction

- **but**: positionner un mobile (un piéton ou un automobiliste) dans un réseau cellulaire urbain
- **infrastructure**: zone urbaine – plusieurs stations de base – centre de calcul



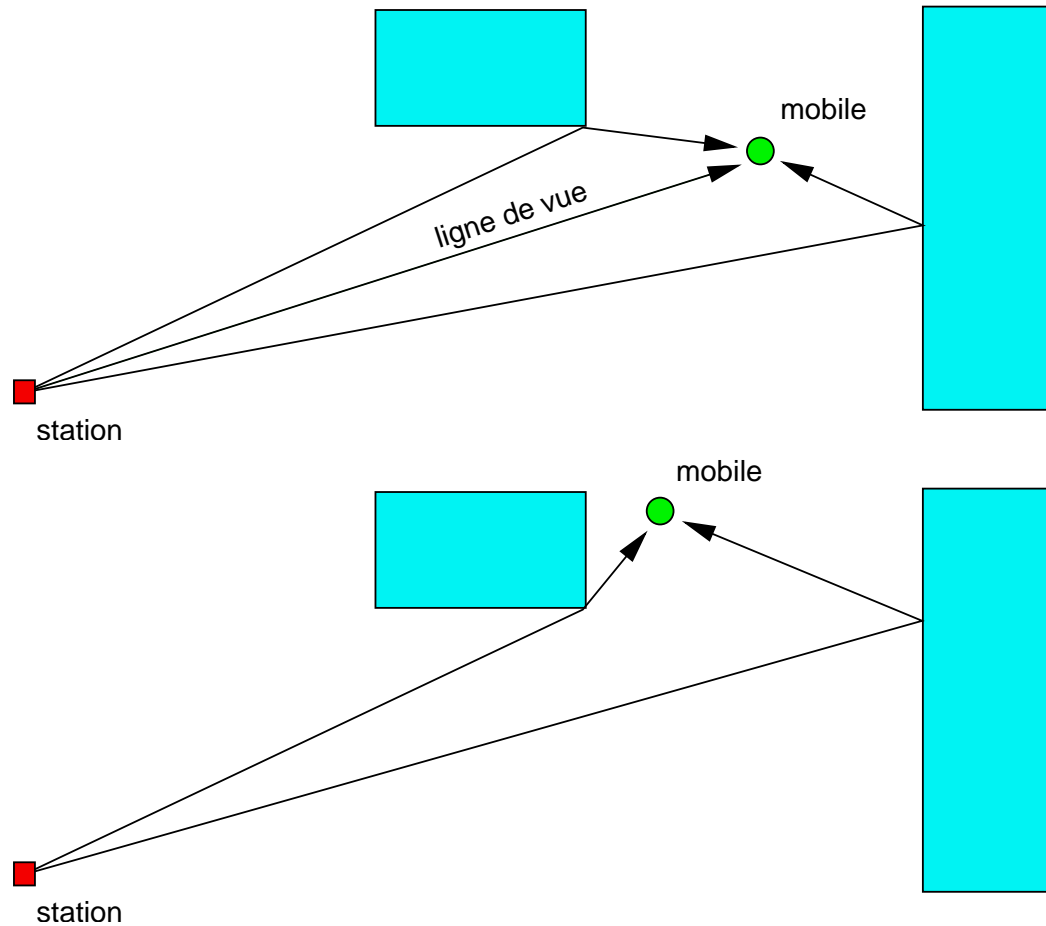
trois approches

- **Localisation orientée réseau (network-based)**: Une ou plusieurs stations de base font les mesures nécessaires et les transmettent à un centre où la position du mobile est calculée. Cette possibilité ne demande pas de changement de mobiles mais implique un fort investissement de la part des opérateurs.
- **Localisation orientée mobile (mobile-based)**: Les mesures et les calculs sont effectués par la station mobile. Cette solution est moins coûteuse pour les opérateurs mais demande une évolution technologique et logicielle des mobiles.
- **Approche hybride**: Dans ces solutions hybrides les mesures sont faites par le mobile qui les transmet aux stations de base où seront fait les calculs.

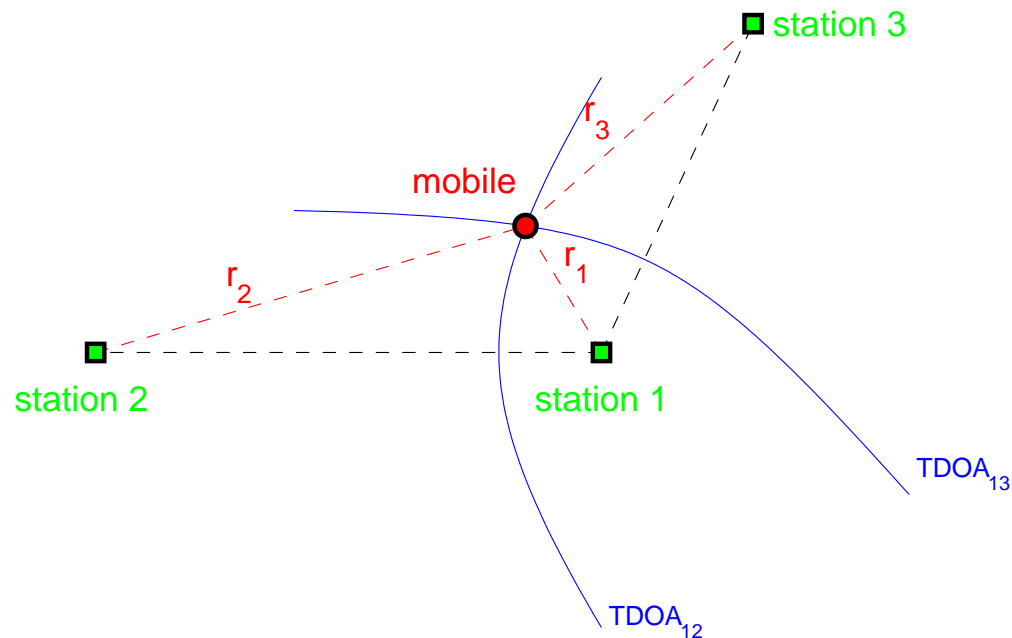
différents types de mesure

- **angles**: Le mobile émet un signal en direction de différentes stations. Les angles d'arrivée (AOA/angle of arrival) de ce signal en plusieurs stations permet de calculer la position du mobile par triangulation.
- **distances**: Le temps de propagation du signal (du mobile à la station ou de la station au mobile) est souvent utilisé en géolocalisation.
 - TOA (time of arrival): un signal émis par la station et renvoyé par le mobile à cette même station permet de calculer le temps de propagation → mesures médiocres.
 - TDOA (time difference of arrival): ici on mesure les temps d'arrivée du mobile à trois stations différentes (ou vice-versa). Cela demande une synchronisation des stations. La position est alors donnée à l'intersection de deux hyperboles: points où la différence des distances (c'est-à-dire des temps de propagation) de deux points est constante, ces deux points fixes étant les foyers de l'hyperbole.
- **puissance du signal reçu**: Utiliser la puissance du signal reçu par le mobile nécessite soit un bon modèle d'atténuation ou bien une carte empirique échantillonnée en toutes les positions intéressantes de la zone considérée.

LOS / no LOS (lign of sight)

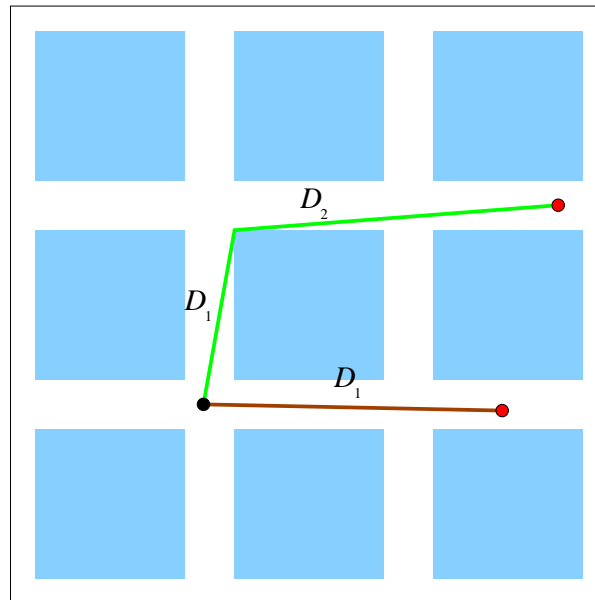


TDOA



Le TDOA est mesuré: i.e. $k = r_2 - r_1 > 0$ est connu, on sait donc que le mobile se trouve sur la branche de l'hyperbole dont les foyers sont les stations.
Pour localiser le mobile à partir des mesures TDOA il est nécessaire que le mobile communique avec 3 stations: les stations 1 et 2 donnent une première hyperbole, les stations 1 et 3 en donnent une deuxième, le mobile est localisé à l'intersection.

carte d'atténuation simplifiée



$$P_{\text{reçu}} = P_{\text{transmis}} - 10 \alpha \log \left(\frac{d_n}{K} \right) + N(0, \sigma^2)$$

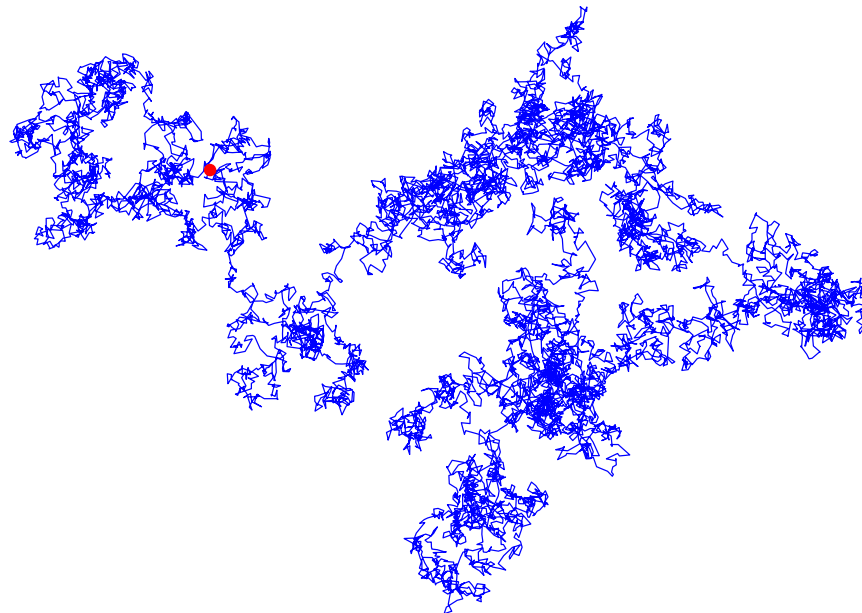
avec

$$d_\ell = d_{\ell-1} + C_\ell + \beta_\ell D_\ell, \quad d_0 = 0 \quad D_\ell = \text{longueur du segment } \ell$$

système d'état: un piéton

$$X_k = \begin{pmatrix} X_k^1 \\ X_k^2 \end{pmatrix} \quad \text{position du piéton à l'instant } t_k = \Delta k$$

à chaque instant il se déplace dans une direction choisit au hasard, indépendamment de ce qu'il a fait dans le passé $\rightarrow k \mapsto X_k$ est un mouvement brownien (une approximation)



système d'état: un piéton

$$X_{k+1}^1 = X_k^1 + \sigma \sqrt{\Delta} W_k^1$$

$$X_{k+1}^2 = X_k^2 + \sigma \sqrt{\Delta} W_k^2$$

où W_k^1 et W_k^2 sont i.i.d. $N(0, 1)$

$$p(X_{k+1}|X_k) = N(X_k, \sigma^2 \Delta I)$$

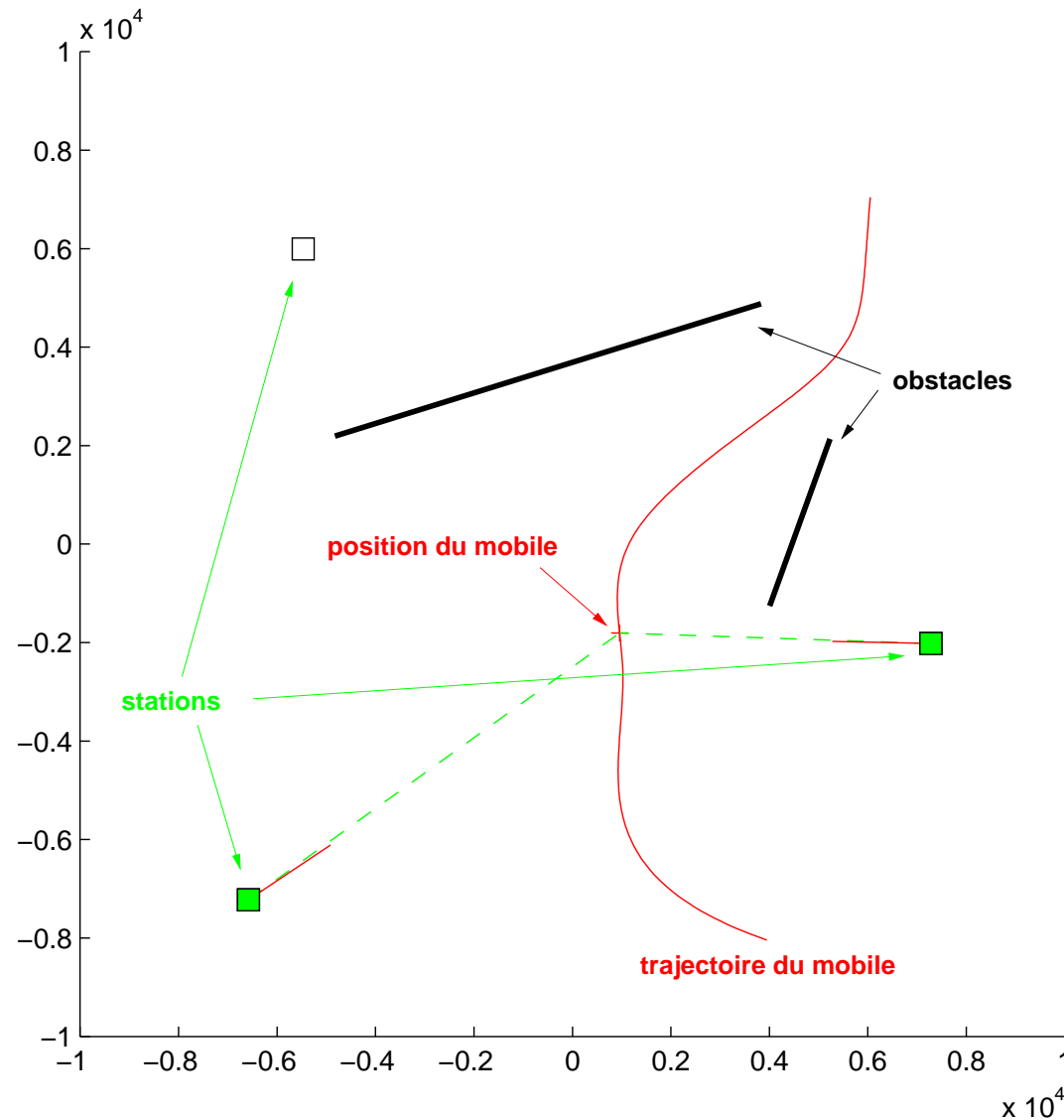
- modèle \neq réalité
- que se passe-t-il si le piéton prend une voiture ? (peut-on suivre une voiture ?)

références (géolocalisation)

- C. Åkerblom, *Tracking Mobile Phones in Urban Areas*. Thesis, Chalmers University of Technology, Göteborg, 2000.
<http://www.md.chalmers.se/Math/Research/Preprints/2000/57.ps.gz>
- C. Drane, M. Macnaughtan, C. Scott, Positioning GSM telephones. *IEEE Communications Magazine*, 36(4):46–54, 1998.
- F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P-J. Nordlund, Particle filters for positioning, navigation and tracking. *IEEE Transactions on Signal Processing*, 50(2), 2002.
<http://www.control.isy.liu.se/~fredrik/reports/01SPpf4pos.pdf>
- Masaharu Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, VT-29(3):317–325, 1980.
- Masaharu Hata, Takayoshi Nagatsu. Mobile location using signal strength measurements in a cellular system. *IEEE Transactions on Vehicular Technology*, VT-29(2):245–252, May 1980.
- F. Gunnarsson P-J. Nordlund, F. Gustafsson. Particle filters for positioning in wireless networks. In *Proc. of EUSIPCO*, Toulouse, 2002.
<http://www.control.isy.liu.se/~fredrik/reports/02eusipco.pdf>.

poursuite avec obstacles

exemple dû à Simon Maskell



modèle

- équation d'état

$$X_k = \begin{pmatrix} X_k^1 \\ X_k^2 \end{pmatrix} = \begin{cases} X_{k+1}^1 = X_k^1 + \sigma_W W_k^1 \\ X_{k+1}^2 = X_k^2 + \sigma_W W_k^2 \end{cases}$$

- initialisation

$$X_0 \sim N(\bar{x}_0, \sigma_0^2 I)$$

- équation d'observation: pour chacune des stations $s = 1 \dots S$
 - si la station s **ne voit pas le mobile** → pas de mesure
 - si la station s **voit le mobile** la mesure est :

$$Y_k^s = h_s(X_k) + \sigma_v V_k \quad \text{avec} \quad h_s(X) = \text{atan2}(X^1 - X^{1,s}, X^2 - X^{1,s})$$

$(X^{1,s}, X^{1,s})$ sont les coordonnées de la station s

filtre particulaire

initialisation

$$\xi_0^{1,i} \leftarrow \bar{x}_0^1 + \sigma_0 N(0, 1)$$

$$\xi_0^{2,i} \leftarrow \bar{x}_0^2 + \sigma_0 N(0, 1)$$

mutation

$$\xi_{k-}^{1,i} = \xi_{k-1}^{1,i} + \sigma_W W_k^1$$

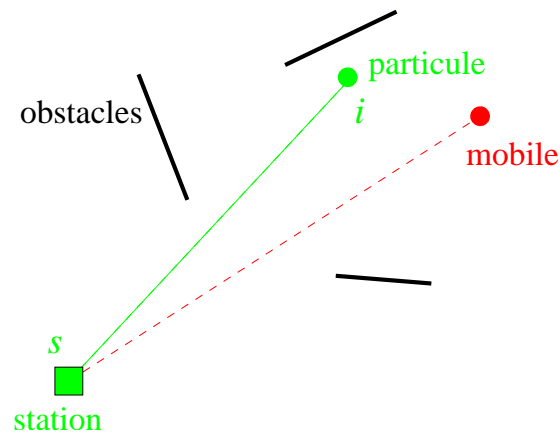
$$\xi_{k-}^{2,i} = \xi_{k-1}^{2,i} + \sigma_W W_k^2$$

filtre particulaire (suite)

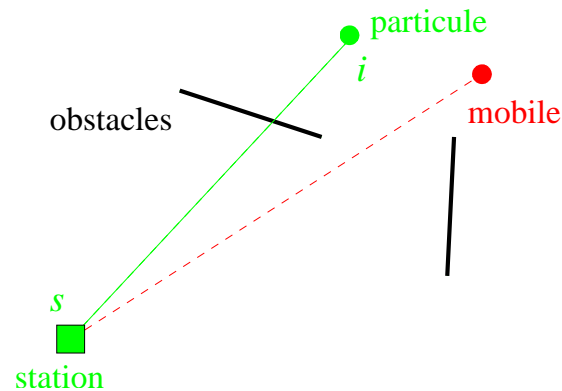
vraisemblance

l'instant k est fixé, on calcule la vraisemblance $\omega_k^{i,s}$ de la particule i pour la station s

cas 1 : la station voit le mobile



la station voit la particule

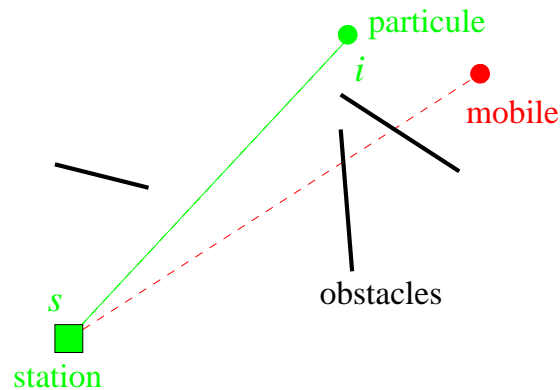


la station ne voit pas la particule

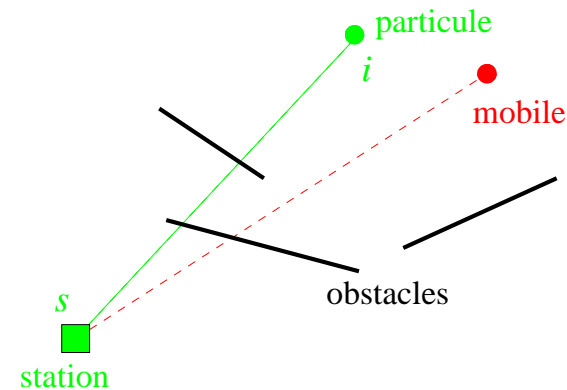
$$\omega_k^{i,s} = \begin{cases} \exp \left(-\frac{|Y_k^s - h_s(\xi_{k-}^i)|^2}{2\sigma_V^2} \right) & \text{si } s \text{ voit la particule } \xi_{k-}^i \\ 0 & \text{si } s \text{ ne voit pas la particule } \xi_{k-}^i \end{cases} \quad i = 1 \dots N$$

filtre particulaire (suite)

cas 2 : la station ne voit pas le mobile



la station voit la particule



la station ne voit pas la particule

$$\omega_k^{i,s} = \begin{cases} 0 & \text{si } s \text{ voit la particule } \xi_{k-}^i \\ 1 & \text{si } s \text{ ne voit pas la particule } \xi_{k-}^i \end{cases} \quad i = 1 \dots N$$

la vraisemblance de la particule i à l'instant k : $\omega_k^i = \prod_{s=1}^S \omega_k^{i,s}$