

Filtrage non linéaire en poursuite. Une approche par mélanges gaussiens

Serge Prosperi

Département ES21
École Supérieure d'Ingénieurs de Luminy
Luminy case 925,
13288 Marseille cedex 09

RÉSUMÉ

Nous proposons de modéliser les bruits du modèle de filtrage par des mélanges gaussiens, ce qui a été étudié dans [6], où des approximations par mélanges gaussiens des lois de probabilité scalaires usuelles sont présentées. Dans le cas d'équations linéaires - ou linéarisables - on en déduit des variantes des filtres de poursuite (du type PDA) intégrant les composantes du mélange. Nous nous intéressons ensuite à la densité exacte du vecteur d'état dont on observe qu'elle est elle-même un mélange gaussien, avec un calcul récursif de ses composantes, et une réduction de leur nombre conduisant à un algorithme approché. Ces résultats sont appliqués au problème de croisement de sources en sonar passif. Nous abordons pour finir le cas d'équations non linéaires, où les hypothèses des modèles classiques ne sont pas respectées.

ABSTRACT

We propose here to use a gaussian mixture model for the plant and the measurement filter noise, which has already been considered in [6], where approximations with gaussian mixtures of the common scalar probability laws are presented. If the equations are linear, or can be linearized, one obtain modified counterparts of the PDA Filter which accounts the components parameters of the mixture. We then consider the exact conditional probability density of the state vector which is itself a gaussian mixture, with a recursive computation of its components, and a reduction of their number which leads to a rough estimate algorithm. These results are applied to the crossing of passive sonar sources. We end the paper with an insight to the nonlinear equations case, for which model the classical assumptions are not true.

1 Cadre du problème de poursuite

Dans la théorie du filtrage linéaire avec variables d'état, les équations du modèle sont (X désignant le vecteur d'état et Y le vecteur de mesures) :

$$X'(t) = A(t) X(t) + \Gamma(t) W(t) \quad (1.1)$$

$$Y(t) = C(t) X(t) + V(t) \quad (1.2)$$

W et V sont des bruits blancs gaussiens centrés, et les matrices A , Γ , C , ainsi que les covariances Q et R des processus W et V sont connues. La solution optimale (l'estimateur bayésien du vecteur d'état) est fournie par le filtre de Kalman. On peut le généraliser, par linéarisation des équations, au modèle non linéaire :

$$X'(t) = f [t, X(t)] + \Gamma(t) W(t) \quad (1.3)$$

$$Y(t) = h [t, X(t)] + V(t) \quad (1.4)$$

Pour obtenir le Filtre de Kalman Étendu. Les bruits sont toujours dans ce cas supposés gaussiens, d'où une loi conditionnelle gaussienne, caractérisée par ses deux premiers moments. En poursuite multipistes avec fausse alarme, On observe simultanément plusieurs processus, $X_1 \dots X_p$, et plusieurs mesures $Y_1 \dots Y_q$ où les valeurs Y_j peuvent être des mesures effectives effectuées sur les processus observés - on ignore a priori lesquels, ou des fausses alarmes. L'observation est un vecteur $Z(t) = (Y_1, \dots, Y_q)$ dont les composantes sont des candidats de mesures des processus en t . Du fait de la non connaissance de l'affectation des mesures aux processus

d'origine ce n'est plus un modèle de filtrage linéaire (c'est un problème exponentiel) :

$$\hat{X}_t = E (X_t / Z_u, u \leq t) = \sum_H P (H) E (X_t / Z) \quad (1.5)$$

où l'hypothèse H désigne une branche du graphe d'affectation. Les solutions (approchées) usuelles calculent l'état estimé de manière récursive - sans mémorisation des mesures, dans les contextes mono ou multicible, soit en associant le plus proche voisin [5], soit en estimant le vecteur d'état comme une moyenne pondérée des estimations faites à partir des différentes mesures (Probabilistic Data Association, [1]). Il s'agit dans tous les cas de méthodes sous optimales. Il faut en outre disposer au préalable dans un contexte multicible d'une procédure d'initialisation des algorithmes de filtrage (détection - estimation des sources), fondée sur un test de vraisemblance généralisée [3].

2 Cas de bruits non gaussiens

Dans le pistage de sources passives (sonar), un modèle mélange gaussien permet de modéliser le phénomène de croisement de deux sources, avec la présence de deux ou trois modes. C'est l'objet des essais présentés au paragraphe 5.

De fait, toute loi continue peut être approchée avec une précision arbitraire par un mélange gaussien [2]. Cette propriété n'a d'intérêt que si le nombre de termes est restreint, ce qui est souvent le cas en pratique, rendant l'approximation

très intéressante par rapport aux autres approximations, issues de développements de Taylor (Développements de Edgeworth et Cram Charlier).

Si $W(t)$ suit une loi mélange gaussienne ayant pour densité $f_{W(t)}(x) = \sum_l p_{t,l} f_{t,l}(x)$, $f_{t,l}$ étant une densité gaussienne. On indexe (L) les applications $t \rightarrow l(t)$.

$$\hat{X}_t = E(X_t / Z_u, u \leq t) = \sum_{H,L} P_L P(H/L) E(X_t / Z, H, L) \tag{2.1}$$

3 Filtrage

Nous proposons deux approches du problème du filtrage, dans le cadre d'un modèle discret, l'une de type récursif, dérivée du modèle de filtrage linéaire classique, la seconde fondée sur le calcul direct de la densité de probabilité du vecteur d'état.

3.1 Filtrage récursif

On peut donner une version séquentielle de l'algorithme de poursuite (Analogie du PDA, que nous notons MPDA), sous la forme d'un filtre non linéaire qui consiste à approcher la loi a posteriori (loi mélange) du vecteur d'état par une loi gaussienne de même espérance et de même covariance. On peut alors proposer une écriture récursive :

$$\hat{X}(t) = \sum_{l,j} p_{t,l} \beta_{l,j} \hat{X}_{l,j}(t) \tag{3.1}$$

$\beta_{l,j} = P(H_j / l, Z)$, probabilité d'association de la mesure j dans l'hypothèse d'un bruit gaussien de densité $f_{t,l}$ est calculé sans difficulté par la règle de Bayes en supposant la probabilité de détection des sources connue et que la répartition des fausses alarmes suit une loi de poisson [1]. $\hat{X}_{l,j}(t)$ est l'estimateur obtenu par les équations de Kalman toujours dans l'hypothèse où $Y = Y_j$ et d'un bruit d'état gaussien de densité $f_{t,l}$.

On peut traiter de manière analogue le cas où le bruit de mesure est un mélange gaussien, de proportions $q_{t,r}$ et de densités $g_{t,r}$. L'estimateur a la même forme que précédemment, dans les expressions (2.1) et (3.1) :

$$\hat{X}(t) = \sum_{l,r,j} p_{t,l} p_{t,r} \beta_{l,r,j} \hat{X}_{l,r,j}(t) \tag{3.2}$$

La covariance du filtre MPDA peut être calculée immédiatement à partir de (3.2), dans le cas discret :

On pose $\gamma_{k,l,r,j} = p_{k,l} q_{k,r} \beta_{k,l,r,j}$. Le produit des probabilités des composantes pour les bruits d'état et de mesure et de la probabilité d'association.

$$P_{k/k-1} = A_k P_{k-1/k-1} A_k^t + \Gamma_k \text{cov}(V_k) \Gamma_k^t \tag{3.3}$$

$$Q_k = \text{cov}(V_k) = \sum_{k,l} p_{k,l} \text{cov}(V_{k,l}) \tag{3.4}$$

$$P_{k/k} = G_k \Lambda_k G_k^t + \sum_{l,r} \gamma_{k,l,r,0} P_{k/k-1} + (1 - \sum_{l,r} \gamma_{k,l,r,0}) (I - G_k C_k) P_{k/k-1} \tag{3.5}$$

$$G_k = P_{k/k-1} C_k^t S_k^{-1} \text{ est le gain du filtre, avec } S_k = C_k P_{k/k-1} C_k^t + \sum_r q_{k,r} R_{k,r} \tag{3.6}$$

$$\Lambda_k = \sum_{l,r,j} \gamma_{k,l,r,j} (v_{k,l,r,j}) (v_{k,l,r,j})^t - v_k v_k^t$$

où $v_{k,l,r,j}$ désignent les innovations et v_k les innovations pondérées.

3.2 Calcul direct de la loi de l'estimateur

Cette approche consiste à traiter directement la densité de probabilité conditionnelle :

$$p(X/Z) = \sum_{H,L} P_L P(H/L) p(X/Z, H, L) \tag{3.7}$$

Cette écriture peut là encore être faite de manière récursive dans le cas d'un modèle (linéaire) discret avec les équations du filtrage non linéaire.

Si on suppose que la densité conditionnelle de l'état à l'étape $k-1$ est un mélange gaussien :

$$p(x_{k-1} / Z^{k-1}) = \sum_L P_L \phi_L(x_{k-1}) \tag{3.8}$$

à l'étape suivante la densité conditionnelle $p(x_k / Z^k)$ s'écrit

$$\alpha \sum_L P_L \sum_{l,r,j} p_{k,l} q_{k,r} \beta_{k,l,r,j} g_{k,r} (y_{k,j} - C_k x_k) \int f_{k,l}(x_k - A_k x_{k-1}) \phi_L(x_{k-1}) dx_{k-1}$$

où α est une constante de normalisation. Il s'agit bien encore d'un mélange gaussien. On peut calculer les densités de probabilité qui apparaissent à l'intérieur de cette somme. Mais le résultat de ce calcul est connu, puisqu'il s'agit des densités gaussiennes dont les moments sont fournis par le filtre de Kalman.

Cette technique conduit donc ici encore à mettre en parallèle un banc de filtres. Comme dans tout problème de filtrage non linéaire, on est amené à estimer la densité de probabilité, c'est à dire une fonction multivariable, mais représentée ici par un nombre fini de paramètres. L'inconvénient est que ce nombre croît - rapidement - si on ne limite pas le nombre de termes du mélange. Cela paraît difficile à mettre en oeuvre en dimension élevée, ou avec des bruits nécessitant un grand nombre de composantes pour le mélange gaussien, mais peut être proposée en écoute passive (par exemple en sonar) avec mesures angulaires.

Dans tous les cas de figure, il faut néanmoins réduire en continu le nombre de composantes du mélange.

On peut également utiliser un estimateur Markovien à un ordre s quelconque, de manière analogue à ce que propose Singer dans [8], ce qui requiert le calcul de m^s filtres, si l'état estimé est un mélange comportant m composantes gaussiennes. Cela peut être vu soit comme une limitation de la propagation des équations (3.8) à un nombre fini s de pas, ou comme une écriture du filtre MPDA à l'ordre s . L'hypothèse faite ici est dans tous les cas (contrairement à [8]) que la loi a priori de l'état est une loi mélange.

4 Modélisation

4.1 Choix des bruits de modèle

Le choix du mélange (nombre de composantes et paramètres) peut être fait par apprentissage (identification de mélange gaussien) si on dispose d'un nombre suffisant de mesures pour chaque source, ce qui n'est pas toujours faisable pour des systèmes dynamiques non stationnaires. Si les bruits sont stationnaires, cette loi peut être évaluée paramétriquement, en fonction des caractéristiques des processus, dans la mesure où on peut les identifier (c'est à dire s'affranchir à cette étape du problème d'affectation). Cela ne posera pas de gros problèmes en présence de forts rapports signal sur bruit. Dans le cas contraire, ce problème est analogue au problème classique de normalisation.

La question de l'identification de lois mélange a reçu un grand intérêt depuis K. Pearson, et a instigué de nombreuses méthodes (méthode des moments, approximation stochastique ou encore techniques issues des problèmes de classification comme les nuées dynamiques, algorithme EM) car les équations de vraisemblance sont trop complexes pour être résolues directement. Il s'agit d'un problème difficile, en particulier pour l'estimation du nombre de composantes. Si la dimension des données est faible, et le nombre de modes connu (ou borné), le choix de la méthode n'est pas fondamental, et une approche de type vraisemblance est ici bien adaptée.

4.2 Réduction du nombre de composantes

Pour la détermination des composantes retenues dans la propagation du calcul de la densité du vecteur d'état, on peut chercher la meilleure approximation de cette loi sous la forme d'un nouveau mélange ayant un nombre N de composantes déterminé.

Mais ici on connaît une expression analytique de la loi (mélange gaussien) et on veut réduire le nombre de composantes. On peut pour cela proposer une méthode directe d'estimation des paramètres de la nouvelle loi, en minimisant une distance entre les deux fonctionnelles, comme l'écart quadratique - ce qui revient à la même complexité que le problème d'identification, ou utiliser une méthode ad hoc (de type classification) ce que nous avons mis en oeuvre.

1. On néglige les proportions inférieures à un seuil, fonction du nombre de modes retenu.
2. On établit des classes de composantes sur un critère de distance normalisée des moyennes, et on traite ces classes de façon indépendante.

5 Croisement de sources (sonar passif)

5.1 Le problème

Le contexte est celui de deux sources donnant lieu à des mesures angulaires (azimut) qui correspondent à des pics d'énergie dans le domaine spatial, représentées dans le cas de mouvements lents par une équation d'état approchée. On suppose que l'accélération est nulle, ce qui est le cas en dehors du point de contact entre les sources de bruit et l'observateur. Lorsque les sources se rapprochent l'une de l'autre, les mesures sont entachées d'un biais, jusqu'à un point où elles ne sont plus résolues - on peut noter que ce problème de capture a déjà été abordé (JPDAM, [1]), mais dans le contexte d'échos radar, avec une discrétisation en cases distance. Le biais moyen est estimé par une phase d'apprentissage (il dépend des niveaux de puissance rayonnée par les sources). Ici, le bruit de mesure n'est plus gaussien. En outre, dans la phase de transition entre les zones résolue (avec présence de deux mesures) en non résolue (une seule mesure), chaque mesure peut être considérée comme suivant un mélange à deux composantes.

5.2 Résultats

Les essais ont été réalisés jusqu'ici uniquement sur signaux synthétiques, en générant des sorties de voies, avec un critère de détection par seuil qui fournit des mesures correspondant à une probabilité de détection P_D (estimée) ainsi que des fausses alarmes. Nous avons pris pour vecteur d'état un vecteur à 4 composantes qui comprend les positions et vitesses des deux sources, et identifié la loi du bruit en azimut dans la phase de croisement en fonction de la distance estimée des sources et de leur puissance relative, par un mélange gaussien à 4 composantes, puis filtré les pistes par un filtre de type MPDA, et comparé avec l'utilisation d'une technique NNA et d'un filtrage PDA conjoint des deux sources (JPDA). Les filtres sont initialisés avec les valeurs exactes.

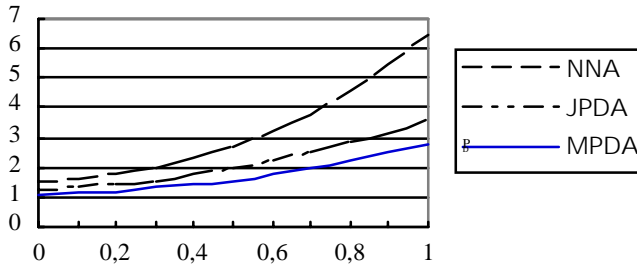
Le scénario retenu est celui de deux sources de même puissance, en l'absence de manoeuvre pendant le croisement. Contrairement à [3] nous ne tenons pas compte de l'énergie des sources, et le réglage du seuil fournit simplement le taux de fausse alarme (un rapport signal sur bruit de 10 dB assure une P_D supérieure à 0.9 dans tous les cas). L'écart - type des mesures hors croisement est de 1 degré.

Nous représentons à la figure 5.1 l'erreur d'estimation en fonction de la densité μ de fausses alarmes exprimée en nombre moyen d'alarmes par degré, observée sur toute la phase de croisement avant capture. le modèle exact de bruit est supposé connu.

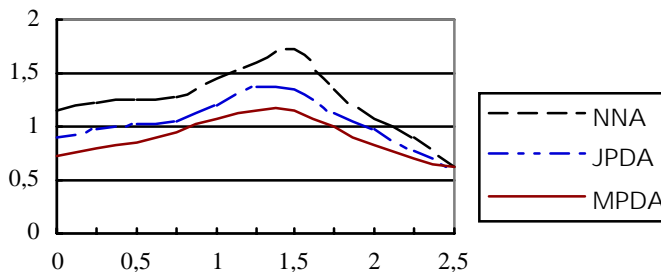
La figure 5.2 présente également l'écart type de l'erreur, mais en l'absence de fausses alarmes, en fonction de la distance d des deux pistes exprimée en degrés, ce qui permet de juger de manière intrinsèque la discrimination des sources. Le modèle pour le bruit de mesure a été déterminé une seule fois pour la génération des mesures et repris par la suite, et est ici estimé à chaque tirage.

Ces valeurs ont été obtenues en simulation de type Monte Carlo sur 100 tirages. Nous n'avons pas comparé les résultats

à la borne d'estimation en poursuite [4] car il était difficile de tenir compte dans la matrice de Fisher du modèle exact pour les lois mélange, mais cette comparaison est certainement souhaitable pour avoir un étalon absolu des performances de la poursuite.



• Fig. 5.1 Ecart - type $\sigma = f(\mu)$



• Fig. 5.2 Ecart - type $\sigma = f(d)$

6 Modèles non linéaires

Dans le modèle général de filtrage non linéaire, les équations de transition d'état et de mesure peuvent s'écrire:

$$X'(t) = \Phi [X(t)] \quad (6.1)$$

$$Y(t) = \psi [X(t)] \quad (6.2)$$

Φ et ψ étant des fonctions aléatoires. L'équation (6.1) n'a de sens que si on peut intégrer le second membre (intégrale stochastique), par exemple si $X_{k+1} = \Phi [k, X_k, W_{k+1}]$, où Φ est une fonction inversible par rapport à sa troisième variable, ce qui permet la récursion de la loi $p(X_k / Y_1, \dots, Y_k)$.

On peut également formuler le problème général de poursuite avec un modèle non linéaire :

$$X_j'(t) = \Phi [X_j(t)] \quad (6.3)$$

$$Y_j(t) = \psi [X_j(t)] \quad (6.4)$$

$$Z(t) = \zeta [Y_1(t), \dots, Y_p(t)] \quad (6.5)$$

La fonction aléatoire $\zeta = \gamma_r \circ \chi_s \circ \sigma$ est la composée d'une permutation, d'une troncation d'ordre s , et d'une adjonction de r valeurs aléatoires. Si les opérateurs Φ et ψ sont linéaires, le problème est linéaire (mais il ne s'agit pas ici d'un problème de filtrage linéaire).

Dans le cadre général du filtrage non linéaire on n'a pas d'autre choix que d'estimer directement la densité de probabilité conditionnelle de l'état.

On peut encore proposer de modéliser les bruits par des lois mélange, mais ce caractère n'est plus respecté pour la loi a

posteriori - comme le caractère gaussien est perdu par transformation non linéaire. Il faut dans ce cas pouvoir linéariser les équations du filtrage. De fait, on peut donner une version "étendue" de l'estimation de la densité de probabilité du vecteur d'état conditionnellement aux mesures dans le cas d'un modèle linéarisable, avec des versions à l'ordre 1 ou 2 du filtre de Kalman Etendu, aussi bien pour le filtrage classique que pour le problème de poursuite. Dans le cas le plus général (équations 6.1 à 6.5) on n'a pas d'autres ressources, sauf dans l'hypothèse de petites perturbations, que de discrétiser l'espace des états, ce qui n'est possible que si le vecteur d'état est de faible dimension. Nous reviendrons ultérieurement sur ces deux points, qui font l'objet d'une étude spécifique.

7 Conclusion

Lors du croisement de sources en sonar passif, l'hypothèse de bruits gaussiens n'est pas respectée, et nous proposons ici de modéliser les bruits du modèle de filtrage par des mélanges gaussiens. Dans le cas d'un modèle linéaire, on peut écrire de manière analytique les équations du filtre optimal, c'est à dire la densité de probabilité conditionnelle de l'état, qui est un mélange gaussien, et utiliser un estimateur markovien d'ordre r (pour $r = 1$, on retrouve les équations du PDA modifiées), ou un calcul récursif de la densité en limitant le nombre de termes du mélange. Nous avons testé l'utilisation du filtre d'ordre 1 (MPDA) sur des signaux synthétiques pour un croisement de sources, et vérifié l'intérêt de cette approche, en comparant avec l'utilisation des filtres classiques en poursuite. Les performances ne semblent pas être très sensibles à la qualité de l'estimation des paramètres du mélange. Compte tenu de l'évolution des puissances de calcul depuis 25 ans, l'approche par modèles de lois mélanges déjà proposée dans [6] peut retrouver aujourd'hui un regain d'intérêt.

8 Bibliographie

- [1] Y. Bar Shalom, T. Fortmann : "Tracking and Data Association" Academic Press, 1988.
- [2] W. Feller : "An Introduction to Probability Theory and its Applications", T2, John Wiley, 1971.
- [3] S. Prosperi : "New Passive Tracking Algorithms using Power Information with application to Passive Sonar", ICASSP, Mai 1991, Toronto.
- [4] S. Prosperi : "Bornes d'estimation en poursuite de pistes", GRETSI, Septembre 1993, Juan les pins.
- [5] R. A. Singer, J. J. St