

Comparação de Desempenho entre FKE e Filtro de Partículas usando Modelos Não-Lineares

Luiz Henrique Monteiro Ferreira

Instituto Tecnológico de Aeronáutica, Praça Marechal Eduardo Gomes, 50 – Vila das Acácias – CEP 12228-900 – São José dos Campos – SP – Brasil

Resumo — Cada vez mais é imprescindível inserir modelos não-lineares às aplicações afim de se obter resultados mais precisos e confiáveis. A medida que a necessidade de processamento em tempo real e armazenamento aumentam, os sistemas devem se adaptar rapidamente as novas características do sinal. Este artigo apresenta uma revisão dos métodos de estimação tradicionais, filtro de Kalman e FKE (Filtro de Kalman Extendido), e as variações dos filtros de partículas, SIS (Sequential Importance Sampling) e SIR (Sampling Importance Resampling), ao final, compara o desempenho dos filtros FKE e SIR aplicados a um sistema não-linear com estatísticas gaussianas.

Palavras-chaves — Estimação bayesiana, filtro de partículas, métodos de Monte Carlo.

I. INTRODUÇÃO

Muitos campos da ciência requerem a estimação de estado de sistemas dinâmicos, sem dúvida, o modelo de espaço de estado é o mais empregado, este artigo trata deste modelo em tempo discreto. A modelagem de espaço de estado concentra-se no vetor de estados, este, por sua vez, contém todas as informações relevantes para descrever o sistema desejado, no entanto, somente o vetor de estado não é suficiente para realizar-se uma boa estimação, então surge a figura o vetor de observação, que geralmente tem uma dimensão menor do que a do vetor de estados, porque nem todos os parâmetros do vetor de estados são observáveis, contudo, dele são tiradas as informações reais que podem ser medidas através dos sensores. Por exemplo, um foguete é lançado, e deseja-se determinar com precisão a sua trajetória, para tal é necessário o conhecimento das características cinemáticas do foguete. O vetor de estado é o mais indicado para conter tais informações, no entanto, para uma boa estimação da sua trajetória, devem entrar em cena sensores capazes de atuar como vetores de observação, tais como, um sistema Radar ou um telêmetro a Laser.

Portanto, afim de analisar e fazer inferências sobre sistemas dinâmicos são necessários no mínimo dois modelos, primeiramente, um modelo que descreve a evolução do sistema ao longo do tempo, segundo, um modelo relacionado com a medida do estado.

Para este artigo assume-se conhecer todos os modelos probabilístico associados à dinâmica do sistema, esta formulação é fundamental na abordagem bayesiana, na qual se baseiam o filtro de Kalman e de partículas alvos desse trabalho.

A formulação inicial do problema de estimação será apresentada na primeira seção, e nas seções subsequentes, os filtros de Kalman e FKE, bem como, os filtros de partículas SIS e SIR, e por fim, a comparação entre os filtros FKE e SIR empregando um sistema não-linear de estatísticas gaussianas.

II. FORMULAÇÃO DO PROBLEMA

Para se definir o problema de rastreamento não-linear [2], deve-se considerar a evolução da sequência de estados $\{x_k, k \in \mathbb{N}\}$, tal que

$$x_k = f_k(x_{k-1}, v_{k-1}) \quad (1)$$

onde $f_k : \mathfrak{R}^{n_x} \times \mathfrak{R}^{n_v} \rightarrow \mathfrak{R}^{n_x}$ é função do estado anterior x_{k-1} , a função f_k pode ser não-linear, $\{v_{k-1}, k \in \mathbb{N}\}$ é um processo ruidoso iid (independent identically distributed), n_x e n_v são as dimensões dos vetores de estado e do processo ruidoso, respectivamente, e \mathbb{N} representa o conjunto dos números naturais. O objetivo do rastreamento é estimar o estado x_k através do modelo de observação dado por

$$y_k = h_k(x_k, n_k) \quad (2)$$

onde $y_k : \mathfrak{R}^{n_x} \times \mathfrak{R}^{n_n} \rightarrow \mathfrak{R}^{n_z}$ pode ser uma função não-linear, $\{n_k, k \in \mathbb{N}\}$ é uma sequência iid, n_z e n_n são as dimensões dos vetores de observação e do ruído, respectivamente. Para se realizar uma melhor estimação de x_k , a abordagem bayesiana procura estimar recursivamente x_k , baseando-se no conjunto de todas as medidas $y_{1:k} = \{y_1, y_2, \dots, y_k\}$ disponíveis até o instante k .

III. MÉTODOS ÓTIMOS

O filtro de Kalman é um algoritmo que calcula recursivamente as variáveis de estado de um sistema representado por equações de estado lineares [3]. Ele apresenta condições especiais de aplicação, por exemplo, o sistema tem de ser perturbado por ruídos brancos e gaussianos, de forma que os estados possam ser tratados como variáveis aleatórias gaussianas. Devido a essas restrições algumas relações devem ser estabelecidas para o emprego do filtro de Kalman, são elas:

- v_{k-1} e n_k têm densidades de probabilidade gaussianas iid com parâmetros conhecidos (média e covariância);

- $f_k(x_{k-1}, v_{k-1})$ é conhecida e linear de parâmetros x_{k-1} e v_{k-1} ;
- $h_k(x_k, n_k)$ é linear com parâmetros x_k e n_k ;

Desta forma, pode-se escrever o conjunto de equações, que sintetizam o estimador, usando as condições citadas acima mais (1) e (2), como

$$x_k = F_k x_{k-1} + v_{k-1} \quad (3)$$

$$y_k = H_k x_k + n_k \quad (4)$$

onde F_k e H_k são matrizes conhecidas, definindo as funções lineares, v_{k-1} e n_k têm covariâncias iguais a Q_{k-1} e R_k , respectivamente. Com o propósito de se determinar a solução MMSE (Minimum Mean Square Estimation), pode-se considerar as médias de v_{k-1} e n_k nulas, então constrói-se um algoritmo para o cálculo recursivo de $\hat{x}_{k/k}$ ($E\{x_k/y_{1:k}\}$) dividido em passo de predição e atualização [3].

O passo de predição consiste em calcular $\hat{x}_{k/k-1}$ e $P_{k/k-1}$, a partir de $\hat{x}_{k-1/k-1}$ e $P_{k-1/k-1}$, pelas expressões

$$\hat{x}_{k/k-1} = F_k \hat{x}_{k-1/k-1} \quad (5)$$

$$P_{k/k-1} = Q_{k-1} + F_k P_{k-1/k-1} F_k^T \quad (6)$$

Para o passo de atualização deseja-se determinar $\hat{x}_{k/k}$ e $P_{k/k}$, a partir de $\hat{x}_{k/k-1}$ e $P_{k/k-1}$, processando a observação atual y_k ;

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k (y_k - H_k \hat{x}_{k/k-1}) \quad (7)$$

$$P_{k/k} = P_{k/k-1} - K_k H_k P_{k/k-1} \quad (8)$$

onde K_k é conhecido como ganho de Kalman, e pode ser calculado por

$$K_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1} \quad (9)$$

Para inicializar este algoritmo a média e a covariância de x_0 devem ser conhecidas, além de todas as funções densidade de probabilidade serem gaussianas.

III. MÉTODOS SUBÓTIMOS

Neste tópico serão apresentados os dois principais métodos hoje empregados na estimação não-linear, cuja comparação são o objeto deste artigo, são os seguintes:

- filtro de Kalman estendido (FKE);
- filtro de partículas (SIS e SIR).

A. Filtro de Kalman Estendido (EKF)

Caso f_k e h_k , em (1) e (2), sejam funções não-lineares, uma linearização local das equações pode ser suficiente para descrever totalmente a não-linearidade. O filtro estendido de Kalman baseia-se na aproximação de primeira ordem das

funções f_k e h_k , de maneira que, os passos de predição e atualização são dados por

$$\hat{x}_{k/k-1} = f_k(\hat{x}_{k-1/k-1}) \quad (10)$$

$$P_{k/k-1} = Q_{k-1} + \hat{F}_k P_{k-1/k-1} \hat{F}_k^T \quad (11)$$

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + K_k [y_k - h_k(\hat{x}_{k/k-1})] \quad (12)$$

$$P_{k/k} = P_{k/k-1} - K_k \hat{H}_k P_{k/k-1} \quad (13)$$

onde

$$\hat{F}_k = \left. \frac{\partial f_k(x)}{\partial x} \right|_{x=\hat{x}_{k-1/k-1}} \quad (14)$$

$$\hat{H}_k = \left. \frac{\partial h_k(x)}{\partial x} \right|_{x=\hat{x}_{k/k-1}} \quad (15)$$

$$K_k = P_{k/k-1} \hat{H}_k^T (\hat{H}_k P_{k/k-1} \hat{H}_k^T + R_k)^{-1} \quad (16)$$

No entanto, deve-se ficar claro que para se aplicar o FKE as funções densidades de probabilidade devem ser gaussianas, inclusive as pdf condicionais, se não o filtro FKE não obterá bons resultados.

B. Filtro de Partículas

O algoritmo SIS (Sequential Importance Sampling) é um método MC (Monte Carlo) que forma a base para a maioria dos filtros MC, aquele é uma técnica para implementação de um filtro bayesiano recursivo através de simulações MC [1]-[4]. A idéia principal é representar a pdf (probability density function) a posteriori requerida por um conjunto de amostras aleatórias com os seus respectivos pesos associados, e calcular as estimativas baseadas nessas amostras e pesos. Quando o número de amostra torna-se muito grande, esta caracterização MC torna-se uma representação equivalente da pdf a posteriori desejada, então o filtro SIS aproxima-se da estimativa ótima bayesiana [1]-[7].

Afim de descrever o algoritmo, definem-se $\{x_{0:k}^i, w_k^i\}_{i=1}^{N_s}$ como medidas aleatórias que caracterizam a pdf a posteriori $p(x_{0:k}/y_{1:k})$, onde $\{x_{0:k}^i, i=0, \dots, N_s\}$ é um conjunto de pontos com pesos associados $\{w_k^i, i=0, \dots, N_s\}$, sendo aquele o conjunto de todos os estados associados até o tempo k . Os pesos são normalizados tal que $\sum_i w_k^i = 1$. Então, a função densidade a posteriori em k pode ser aproximada como

$$p(x_{0:k}/y_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_{0:k} - x_{0:k}^i) \quad (17)$$

Desta forma, tem-se uma aproximação discreta para a pdf de acordo com cada peso, onde os pesos são escolhidos através do princípio da amostragem por importância [4]. Suponha que $p(x) \propto \pi(x)$ é uma densidade de probabilidade da qual é difícil obter-se amostras, mas que para $\pi(x)$ isto pode ser feito sem maiores problemas, então amostra-se $x^i \sim q(x), i=1, \dots, N_s$ da função $q(x)$, chamada de densidade

de importância., e em seguida, aproxima-se a densidade de probabilidade $p(x)$ por

$$p(x) \approx \sum_{i=1}^{N_s} w_k^i \delta(x - x^i) \quad (18)$$

onde

$$w^i \propto \frac{\pi(x^i)}{q(x^i)} \quad (19)$$

é o peso normalizado da i -ésima partícula.

Portanto, se as amostras $x_{0:k}^i$ forem retiradas de uma densidade de importância $q(x_{0:k}/y_{1:k})$, então os pesos em (17) são definidos por (19), tornando-se

$$w_k^i \propto \frac{p(x_{0:k}^i/z_{1:k})}{q(x_{0:k}^i/z_{1:k})}. \quad (20)$$

Para o caso sequencial, a cada iteração, poderia ter amostras constituindo uma aproximação de $p(x_{0:k-1}/y_{1:k-1})$ e querer aproximar $p(x_{0:k}/y_{1:k})$ com um conjunto de amostras. Se a densidade de importância é escolhida de forma fatorada como se segue

$$q(x_{0:k}/y_{1:k}) = q(x_k/x_{0:k-1}, y_{1:k})q(x_{0:k-1}/y_{1:k-1}) \quad (21)$$

então pode-se amostrar $x_{0:k}^i \sim q(x_{0:k}/y_{1:k})$ por adição de amostras existentes $x_{0:k-1}^i \sim q(x_{0:k-1}/y_{1:k-1})$ com o novo estado $x_k^i \sim q(x_k/x_{0:k-1}, y_{1:k})$. Para derivar o peso da equação atualizada, deve-se escrever $p(x_{0:k}/y_{1:k})$ em função de $p(x_{0:k-1}/y_{1:k-1})$, $p(y_k/x_k)$ e $p(x_k/x_{k-1})$ como segue

$$\begin{aligned} p(x_{0:k}/y_{1:k}) &= \frac{p(y_k/x_k)p(x_k/x_{k-1})}{p(y_k/y_{1:k-1})} \times p(x_{0:k-1}/y_{1:k-1}) \\ &\propto p(y_k/x_k)p(x_k/x_{k-1})p(x_{0:k-1}/y_{1:k-1}). \end{aligned} \quad (22)$$

Por substituição de (21) e (22) em (20), a equação do peso atualizado pode então ser mostrada como

$$\begin{aligned} w_k^i &\propto \frac{p(y_k/x_k^i)p(x_k^i/x_{k-1}^i)p(x_{0:k-1}^i/y_{1:k-1})}{q(x_k^i/x_{0:k-1}^i, y_{1:k})q(x_{0:k-1}^i/y_{1:k-1})} \\ &= w_{k-1}^i \frac{p(y_k/x_k^i)p(x_k^i/x_{k-1}^i)}{q(x_k^i/x_{0:k-1}^i, y_{1:k})}. \end{aligned} \quad (23)$$

Para o caso $q(x_{0:k}/x_{0:k-1}, y_{1:k}) = q(x_k/x_{k-1}, y_k)$ a densidade de importância torna-se dependente apenas de x_{k-1} e y_k . Este modo em particular é útil quando a estimativa de $p(x_k/y_{1:k})$ é requerida a cada passo. Assumindo este cenário, somente x_k^i precisa ser armazenada, conseqüentemente, a trajetória $x_{0:k-1}^i$ pode ser descartada e o histórico das observações $y_{1:k-1}$ também. O peso modificado torna-se

$$w_k^i = w_{k-1}^i \frac{p(y_k/x_k^i)p(x_k^i/x_{k-1}^i)}{q(x_k^i/x_{k-1}^i, y_k)} \quad (24)$$

e a pdf a posteriori $p(x_k/y_{1:k})$ é dada por

$$p(x_k/y_{1:k}) \approx \sum_{i=1}^{N_s} w_k^i \delta(x_k - x_k^i) \quad (25)$$

onde os pesos são definidos em (24). Pode-se mostrar que quando $N_s \rightarrow \infty$, a aproximação (25) tende para a verdadeira densidade a posteriori $p(x_k/y_{1:k})$ [6].

Uma desvantagem do filtro SIS é que a variância dos pesos de importância $\{w_k^i\}$ cresce à medida que k aumenta, gerando um efeito conhecido como degeneração ou “degenerescência” de partículas, segundo o qual após algumas iterações, apenas um pequeno número de amostras têm pesos normalizados significativos (próximos a 1), enquanto a maioria das amostras tem pesos negligenciados (próximos a zero) [5]. Esse fenômeno é o equivalente a redução do número efetivo de partículas na população amostral, exigindo que se use um número muito grande de amostras para garantir a convergência da aproximação da estimativa MMSE.

Uma tentativa de controlar a variância dos pesos de importância $\{w_k^i\}$ é introduzir um passo de seleção de partículas, tal que, a cada instante k , cada partícula i gera um número $N^{(i)}$ de cópias de si mesma tal que $E\{N^{(i)}\} = N_p w_k^i$, onde N_p é o número total de partículas. Em outras palavras, o passo de seleção equivale a substituir o conjunto original de partículas por um novo conjunto formado por cópias das anteriores, onde o número de cópias por partículas é proporcional ao seu peso de importância de modo que partículas de alto peso são multiplicadas e partículas de baixo peso são descartadas [5]. Na prática, este processo de reamostragem não é repetido para todas as iterações do algoritmo, ao invés disso, calcula-se um indicador de degeneração, introduzido por [8],

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N_s} [w_k^i]^2}. \quad (26)$$

E se $\hat{N}_{eff} \leq \alpha N_p$, normalmente α é igual a 0,6 ou 0,7, então efetua-se um passo de reamostragem e modificam-se os pesos para $w_k^i = \frac{1}{N_p}$, do contrário, não se amostram as partículas e são mantidos os pesos originais.

III. TESTES E RESULTADOS

Com o objetivo de comparar os desempenhos dos filtros FKE e SIR, modelou-se um sistema fictício pelas equações a seguir:

$$x_k = \frac{1}{2}x_{k-1} + \frac{25x_{k-1}}{1+x_{k-1}^2} + 8\cos(1.2k) + v_{k-1} \quad (27)$$

$$y_k = \frac{x_k^2}{20} + n_k \quad (28)$$

onde v_{k-1} e n_k são variáveis aleatórias gaussianas de média nula e variâncias 10 e 1, respectivamente, e x_0 gaussiana de média nula e variância 10. Afim de quantificar as performances dos filtros FKE e SIR escolheu-se como critério básico o desvio padrão dos resultados em relação aos valores reais.

Os valores reais e observados são apresentados pela Fig. 1, correspondendo a um total de cem amostras no tempo, no entanto, diferentemente do vetor de estado, o vetor de observação apresentou, na maioria das suas amostras, valores não-nulos devido a (28) e a pequena variância de n_k .

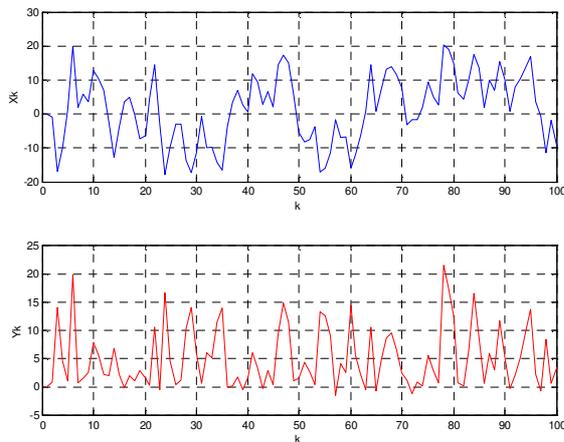


Fig. 1. Figura dos valores reais de x_k e observados y_k .

A linearização local realizada pelo filtro FKE não apresentou bom resultado para o modelo descrito por (27) e (28), isto pode ser constatado pela Fig. 2, que em alguns pontos obtidos pelo filtro referido, mostrou resultados bastante distorcidos, na forma de picos, isso ocorreu devido a características altamente não-lineares do sistema. No entanto, para o caso do filtro SIR foram empregados 1000 partículas, que apresentaram resultados melhores quando comparados ao FKE como pode ser visto na Fig. 3.

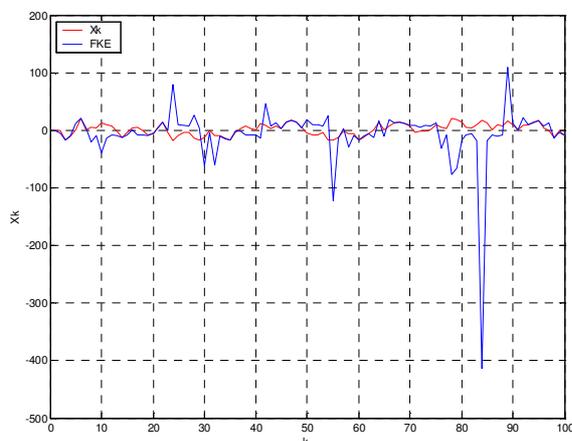


Fig. 2. Valores reais de x_k e estimados pelo FKE.

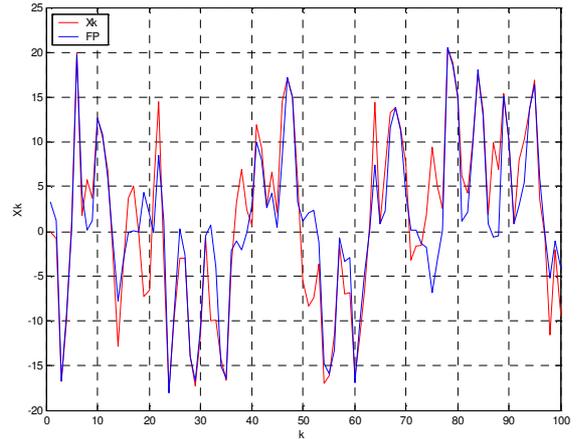


Fig. 3. Resultados reais e valores obtidos pelo filtro SIR.

Para descrever melhor quantitativamente a relação entre os desempenhos dos filtros foi construída a Tabela I, onde são apresentados resultados para os filtros FKE e SIR, no caso deste último, empregando-o com 1000 e 3000 partículas.

TABELA I. DESVIOS PADRÕES

Filtro	Nº de Partículas	Desvio Padrão
FKE	-	34.0025
SIR	1000	5.8196
SIR	3000	4.5902

IV. OBSERVAÇÕES FINAIS

Para a utilização do FKE para modelos não-lineares é necessário se determinar previamente sua região de atuação, do contrário ocorrerá pontos onde a aproximação da primeira derivada se tornará inconsistente. A aproximação dos filtros de partículas mostrou resultados melhores comparados com aqueles obtidos pelo FKE, porém, os filtros de partículas são sensíveis quanto ao número de partículas, e apresentam dificuldade extra de projeto, pois, para cada aplicação específica deve-se escolher uma função de importância adequada.

REFERÊNCIAS

- [1] A. Doucet, N. de Freitas, N. Gordon, "Sequential Monte Carlo in Practice", Springer-Verlag, 2001.
- [2] S. Arulampalam, S. Maskell, N. J. Gordon, and T. Clapp, "A Tutorial on Particle Filters for On-line Non-linear/Non-Gaussian Bayesian Tracking", IEEE Transactions of Signal Processing, Vol. 50, pages 174-188, February 2002.
- [3] A. H. Jazwinski, "Stochastic Processes and Filtering Theory". New York: Academic, 1970.
- [4] A. Doucet, "On sequential Monte Carlo methods for Bayesian filtering", Dept. Eng., Univ. Cambridge, UK, Tech. Rep., 1998.
- [5] J. Carpenter, P. Clifford, P. Fearnhead, "Improved particle filter for nonlinear problems", Proc. Inst. Elect. Eng., Radar, Sonar, Navig., 1999.
- [6] P. Del Moral, "Measure valued processes and interacting particle systems. Application to nonlinear filtering problems", Ann. Appl. Probab., vol. 8, no. 2, pp. 438-495, 1998.

- [7] A. Doucet, S. Godsill, C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering", *Statist. Comput.*, vol. 10, no. 3, pp. 197–208.
- [8] J. S. Liu, R. Chen, "Sequential Monte Carlo methods for dynamical systems", *J. Amer. Statist. Assoc.*, vol. 93, pp. 1032–1044, 1998.