SISEA : Filtrage de Kalman et modèles de Markov cachés

TP 2 : Modèles de Markov cachés

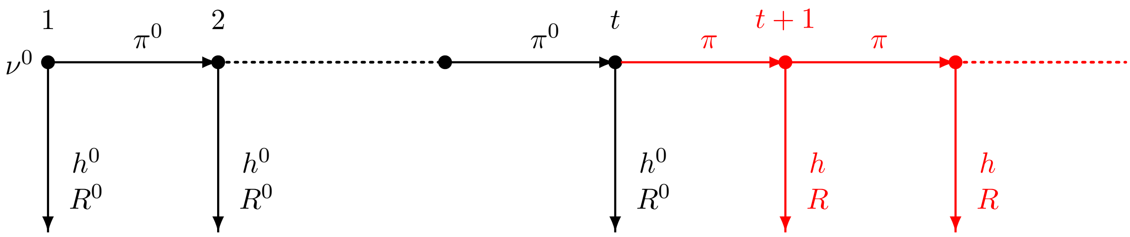
jeudi 15 décembre 2016

L'objectif de ce TP est de montrer comment une utilisation ingénieuse des variables forward et backward permet de détecter un éventuel changement de modèle, et le cas échéant d'estimer l'instant de changement.

On considère un modèle de Markov caché à 2 états et avec des observations uni–dimensionnelles conditionnellement gaussiennes, c'est–à–dire que les densités d'émission sont gaussiennes, avec des moyennes et des variances qui dépendent de l'état caché.

Au vu des observations $(Y_1, \cdots, Y_n)$, on veut détecter s'il existe ou non un instant t compris entre l'instant initial 1 et l'instant final n, tel que

- à l'instant 1, la loi initiale est $\nu^0 = (\nu_i^0)$,
- entre l'instant initial 1 et l'instant t, le modèle est nominal, c'est–à–dire que la matrice de transition est $\pi^0 = (\pi_{ij}^0)$ et les densités d'émission sont gaussiennes, de moyennes $h^0 = (h_i^0)$ et de variances $R^0 = (R_i^0)$,
- entre l'instant $(t+1)$ et l'instant final n, le modèle est modifié, c'est–à–dire que la matrice de transition est $\pi = (\pi_{ij})$ et les densités d'émission sont gaussiennes, de moyennes $h = (h_i)$ et de variances $R = (R_i)$.



On rappelle que la vraisemblance L_n^0 du modèle nominal $\boldsymbol{M}^0 = (\nu^0, \pi^0, h^0, R^0)$ au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer comme

$$L_n^0 = \sum_{i \in E} p_n^{0,i} = \sum_{i \in E} p_k^{0,i} \, v_k^{0,i} \;, \qquad \text{pour tout } k = 1, \cdots, n$$

au moyen des variables forward et backward associées au modèle nominal $\boldsymbol{M}^0$. De même, la vraisemblance L_n du modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ (où le changement est déjà intervenu avant même l'instant initial 1) au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer comme

$$L_n = \sum_{i \in E} p_n^i = \sum_{i \in E} p_k^i \, v_k^i \;, \qquad \text{pour tout } k = 1, \cdots, n$$

au moyen des variables forward et backward associées au modèle alternatif $\boldsymbol{M}$.

On admet le résultat suivant : la vraisemblance $L_n(t)$ du modèle avec un changement à l'instant t au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer comme

$$L_n(t) = \sum_{i \in E} p_t^{0,i} \; v_t^i \; ,$$

au moyen de la variable forward associée au modèle nominal $\boldsymbol{M}^0$ et de la variable backward associée au modèle alternatif $\boldsymbol{M}$.

On rappelle que la version normalisée des variables forward associées au modèle nominal $\boldsymbol{M}^0$ et au modèle modifié $\boldsymbol{M}$ sont définies pour tout instant $k = 1, \cdots, n$ par

$$\bar{p}_k^{0,i} = \frac{p_k^{0,i}}{\displaystyle\sum_{j \in E} p_k^{0,j}} \qquad \text{et par} \qquad \bar{p}_k^i = \frac{p_k^i}{\displaystyle\sum_{j \in E} p_k^j} \qquad \text{pour tout } i \in E,$$

respectivement, et que la version normalisée de la variable backward associée au modèle alternatif $\boldsymbol{M}$ est définie pour tout instant $k = 1, \cdots, n$ par

$$\bar{v}_k^i = \frac{v_k^i}{\displaystyle\sum_{j \in E} \bar{p}_k^j \; v_k^j} \qquad \text{pour tout } i \in E.$$

(i) Montrer que la vraisemblance $L_n(t)$ du modèle avec un changement à l'instant t au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer aussi comme

$$L_n(t) = [\sum_{i \in E} p_t^{0,i}] \; [\sum_{i \in E} \bar{p}_t^{0,i} \; \bar{v}_t^i] \; [\sum_{i \in E} \bar{p}_t^i \; v_t^i] \; .$$

(ii) Montrer que la vraisemblance L_n du modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer aussi comme

$$L_n = \sum_{i \in E} p_t^i \; v_t^i = [\sum_{i \in E} p_t^i] \; [\sum_{i \in E} \bar{p}_t^i \; v_t^i] \; .$$

Si on introduit pour chaque instant de changement t possible, la vraisemblance L_t^0 du modèle nominal $\boldsymbol{M}^0$ au vu des observations $(Y_1, \cdots, Y_t)$ et la vraisemblance L_t du modèle alternatif $\boldsymbol{M}$ au vu des observations $(Y_1, \cdots, Y_t)$, alors on remarque que la fonction de vraisemblance pour l'estimation d'un éventuel instant de changement au vu des observations $(Y_1, \cdots, Y_n)$ peut s'exprimer comme

$$t \mapsto \frac{L_n(t)}{L_n} = \frac{L_t^0}{L_t} \; [\sum_{i \in E} \bar{p}_t^{0,i} \; \bar{v}_t^i] \; ,$$

et la fonction de log–vraisemblance comme

$$t \mapsto \log \frac{L_n(t)}{L_n} = \log L_t^0 - \log L_t + \log[\sum_{i \in E} \bar{p}_t^{0,i} \; \bar{v}_t^i] \; . \qquad (\star)$$

Pour chacun des scénarios étudiés, on disposera de la réalisation particulière de la suite des états cachés (inconnue en principe) et de la suite des observations uni–dimensionnelles, et il s'agira d'estimer l'instant t d'un éventuel changement, au vu de ces observations.

Le modèle nominal $\boldsymbol{M}^0 = (\nu^0, \pi^0, h^0, R^0)$ à $K = 2$ états, est défini par les caractéristiques suivantes (on rappelle que la loi initiale est un vecteur–ligne $1 \times K$, que la matrice de transition est une matrice $K \times K$, et que les K densités d'émission sont gaussiennes, dont les K moyennes respectives sont vues comme une matrice $1 \times K$ et dont les K variances respectives sont vues comme une matrice $1 \times K$) :

$$\nu^0 = \begin{bmatrix} 0.2 & 0.8 \end{bmatrix} \qquad \pi^0 = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix} \qquad h^0 = \begin{bmatrix} -2 & +2 \end{bmatrix} \qquad \text{et} \qquad R^0 = \begin{bmatrix} 5 & 5 \end{bmatrix} .$$

Changement dans la moyenne

Dans ce scénario, le modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ est défini par les caractéristiques suivantes :

$$\pi = \pi^0 \qquad h = \begin{bmatrix} -10 & +10 \end{bmatrix} \qquad \text{et} \qquad R = R^0 .$$

(iii) Lire la suite des états cachés $(X_1, \cdots, X_n)$ dans le fichier `etat_cache_1.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes, ainsi que le vrai instant de changement. Lire la suite d'observations $(Y_1, \cdots, Y_n)$ dans le fichier `observation_1.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes. Pour récupérer ces données

```
load('etat_cache_1.mat','x','tc');
n = size(x,2);
fprintf('instant de changement vrai %i\n',tc);
load('observation_1.mat','y');
```

et pour la visualisation

```
for i=1:K
   axe = subplot(K+1,1,i);
   plot((x==i),'+r');
   axis([1 n -0.5 1.5]);
   set(axe,'YTick',[0 1]);
   title(['etat cache # ',num2str(i)]);
end
subplot(K+1,1,K+1);
plot(y);
title('observation')
```

L'instant de changement peut–il déjà se voir sur la représentation graphique ?

(iv) Résoudre les équations forward et backward pour le modèle nominal $\boldsymbol{M}^0 = (\nu^0, \pi^0, h^0, R^0)$, c'est–à–dire calculer la fonction de log–vraisemblance $\log L_k^0$ et les versions normalisées $\bar{p}_k^0$ et $\bar{v}_k^0$ des variables forward et backward, pour tout instant $k = 1, \cdots, n$. Résoudre les équations forward et backward pour le modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$, c'est–à–dire calculer la fonction de log–vraisemblance $\log L_k$ et les versions normalisées $\bar{p}_k$ et $\bar{v}_k$ des variables forward et backward, pour tout instant $k = 1, \cdots, n$. Tracer la fonction de log–vraisemblance

$$t \mapsto \log \frac{L_n(t)}{L_n} = \log L_t^0 - \log L_t + \log[\sum_{i \in E} \bar{p}_t^{0,i} \; \bar{v}_t^i \,] \; .$$

définie en ($\star$) pour l'estimation d'un éventuel instant de changement au vu des observations $(Y_1, \cdots, Y_n)$, et calculer l'estimateur du maximum de vraisemblance de l'instant de changement.

On conseille d'écrire une fonction MATLAB qui prend en entrée

- une suite d'observations,
- les caractéristiques (loi initiale, matrice de transition, moyennes et variances des densités d'émission gaussiennes) d'un modèle de Markov caché,

et qui rend en sortie

- la suite des log–vraisemblances,
- la suite des variables forward normalisées,
- la suite des variables backward normalisées.

Changement dans la variance

Dans ce scénario, le modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ est défini par les caractéristiques suivantes :

$$\pi = \pi^0 \qquad h = h^0 \qquad R = \left[\, 0.5 \quad 0.5 \,\right] \; .$$

(v) Lire la suite des états cachés $(X_1, \cdots, X_n)$ dans le fichier `etat_cache_2.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes, ainsi que le vrai instant de changement. Lire la suite d'observations $(Y_1, \cdots, Y_n)$ dans le fichier `observation_2.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes. Pour récupérer ces données

```
load('etat_cache_2.mat','x','tc');
n = size(x,2);
fprintf('instant de changement vrai %i\n',tc);
load('observation_2.mat','y');
```

L'instant de changement peut–il déjà se voir sur la représentation graphique ?

(vi) Procéder comme à la question (iv).

Changement simultanné dans la moyenne et dans la variance

Dans ce scénario, le modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ est défini par les caractéristiques suivantes :

$$\pi = \pi^0 \qquad h = \begin{bmatrix} -1 & +1 \end{bmatrix} \qquad R = \begin{bmatrix} 3 & 3 \end{bmatrix} \ .$$

(vii) Lire la suite des états cachés $(X_1, \cdots, X_n)$ dans le fichier `etat_cache_3.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes, ainsi que le vrai instant de changement. Lire la suite d'observations $(Y_1, \cdots, Y_n)$ dans le fichier `observation_3.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes. Pour récupérer ces données

```
load('etat_cache_3.mat','x','tc');
n = size(x,2);
fprintf('instant de changement vrai %i\n',tc);
load('observation_3.mat','y');
```

L'instant de changement peut–il déjà se voir sur la représentation graphique ?

(viii) Procéder comme à la question (iv).

Changement dans la matrice de transition

Dans ce scénario, le modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, h, R)$ est défini par les caractéristiques suivantes :

$$\pi = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix} \qquad h = h^0 \qquad R = R^0 \ .$$

(ix) Lire la suite des états cachés $(X_1, \cdots, X_n)$ dans le fichier `etat_cache_4.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes, ainsi que le vrai instant de changement. Lire la suite d'observations $(Y_1, \cdots, Y_n)$ dans le fichier `observation_4.mat` où elle est stockée sous la forme d'une matrice à 1 lignes et n colonnes. Pour récupérer ces données

```
load('etat_cache_4.mat','x','tc');
n = size(x,2);
fprintf('instant de changement vrai %i\n',tc);
load('observation_4.mat','y');
```

L'instant de changement peut–il déjà se voir sur la représentation graphique ?

(x) Procéder comme à la question (iv).

Fournir

- un (court) rapport écrit présentant et discutant les résultats obtenus : observation de l'instant de changement directement sur la représentation graphique, apparence de la fonction de log–vraisemblance, pour chacun des scénarios étudiés.
- le code source MATLAB utilisé pour obtenir ces résultats.