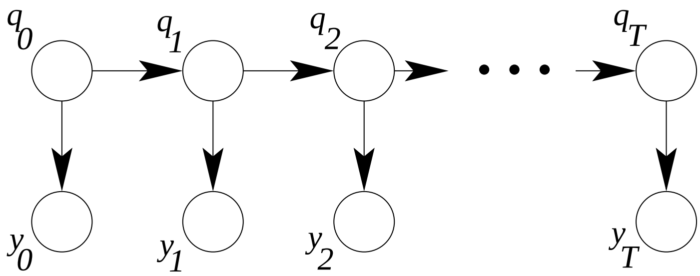
Mastere M2 MVA 2007 - Modèles graphiques

Exercices à rendre pour le 12 Décembre 2007.

Ces exercices peuvent s'effectuer par groupe de deux élèves.

1 Filtre de Kalman

On considère le modèle de Markov caché suivant :



avec $q_t \in \mathbb{R}^p$ et $y_t \in \mathbb{R}^q$ pour tout $t \in \{0, \ldots, T\}$, et les distributions suivantes :

$$q_0 \sim \mathcal{N}(0, \Sigma_0),\; q_t|q_{t-1} \sim \mathcal{N}(Aq_{t-1}, H),\; y_t|q_t \sim \mathcal{N}(Cq_t, R).$$

On suppose $y_0, \ldots, y_T$ observés. La "récursion α" vue en cours est définie par

$$\begin{aligned}
\alpha_0(q_0) &= p(y_0|q_0)p(q_0) \\
\alpha_{t+1}(q_{t+1}) &= \int \alpha_t(q_t)p(q_{t+1}|q_t)p(y_{t+1}|q_{t+1})dq_t
\end{aligned}$$

(a) Quelle est la forme fonctionnelle de $\alpha_t(q_t)$?

(b) Montrer que pour une représentation bien choisie de $\log(\alpha_t(q_t))$ (pour laquelle les termes constants sont laissés non spécifiés), la récursion α est équivalente à la récursion suivante (elle-même équivalente aux récursions (15.55-58) et aux récursions classiques du filtre de Kalman) :

$$\begin{aligned}
\xi_0 &= C^\top R^{-1} y_0 \\
\Omega_0 &= \Sigma_0^{-1} + C^\top R^{-1} C \\
\xi_{t+1} &= C^\top R^{-1} y_{t+1} + H^{-1}A(\Omega_t + A^\top H^{-1} A)^{-1}\xi_t \\
\Omega_{t+1} &= H^{-1} + C^\top R^{-1} C - H^{-1}A(\Omega_t + A^\top H^{-1}A)^{-1}A^\top H^{-1}
\end{aligned}$$

On utilisera le lemme suivant :

Lemme 1 *Si* $u(x,y) \propto \exp\left(-\frac{1}{2}\begin{pmatrix} x \\ y \end{pmatrix}^\top \begin{pmatrix} A & B \\ B^\top & C \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ c \end{pmatrix}^\top \begin{pmatrix} x \\ y \end{pmatrix}\right)$, *alors*

$$\int u(x,y)dx \propto \exp\left(-\frac{1}{2}y^\top(C - B^\top A^{-1}B)y + y^\top(c - B^\top A^{-1}a)\right)$$

2 Implémentation - mélange de Gaussiennes

Le fichier "EMGaussienne.dat" contient un ensemble de données (x_n, y_n) où $(x_n, y_n) \in \mathbb{R}^2$. Le but de cet exercice est d'implémenter l'algorithme EM pour certains mélanges de K Gaussiennes dans $\mathbb{R}^d$ (dans cet exercice, $d = 2$ et $K = 4$), avec des données IID. (NB : dans cet exercice, il n'est pas nécessaire de démontrer les formules utilisées).

Le langage de programmation est libre (MATLAB, Octave, Scilab ou R sont néanmoins recommandés). Le code source doit être remis avec les résultats. Le corrigé sera codé en MATLAB.

(a) Implémenter l'algorithme K-means (chapitre 10 du polycopié). Représenter graphiquement les données d'apprentissage, les centres obtenus et les différents groupes ("clusters"). Essayer plusieurs initialisations et comparer les résultats (centres et mesures de distortions).

(b) Implémenter l'algorithme EM pour un mélange de Gaussiennes avec des matrices de covariance proportionnelles à l'identité (initialiser l'algorithme EM avec les moyennes trouvées par K-means).

Représenter graphiquement les données d'apprentissage, les centres et les covariances obtenus (une manière élégante de représenter graphiquement est de représenter l'ellipse contenant un certain pourcentage (e.g., 90%) de la masse de la Gausienne). Estimer et représenter la variable latente pour chaque point (pour le jeu de paramètres appris par EM).

(c) Implémenter l'algorithme EM pour un mélange de Gaussiennes avec des matrices de covariance générales. Représenter graphiquement les données d'apprentissage, les centres et les covariances obtenus. Estimer et représenter la variable latente pour chaque point (pour le jeu de paramètres appris par EM).

(d) Commenter les différents résultats obtenus. En particulier, comparer les log-vraisemblances des deux modèles de mélanges, sur les données d'apprentissage, ainsi que sur les données de test (dans "EMGaussienne.test").