Algorithmic Trading: Hidden Markov Models

Sebastian Jaimungal
University of Toronto

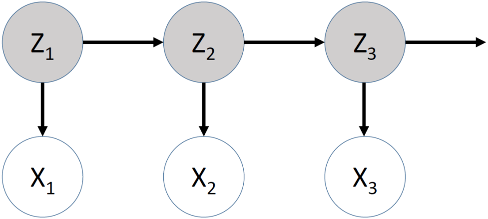
many thanks to

Álvaro Cartea, Oxford

Feb, 2017

Hidden Markov Models

HMM



- $Z = \{Z_t : t = 1, \dots, N\}$ is a K-state Markov chain, transition matrix A, modelling the **unobserved** or **latent** states
- $X = \{X_t : t = 1, \dots, N\}$ are the **observed data** with **emission probability**

$$\psi_i(x) := \mathbb{P}(X_t = x \mid Z_t = i)$$

[Note: it is independent of time]

HMM

- The **completed-data log-likelihood** for observations $\boldsymbol{x} = \{x_1, \dots, x_n\}$ is

$$\begin{aligned}\bar{\ell}(\Theta; \boldsymbol{Z}) = & \sum_{i=1}^{K} (\log \pi_i)\, \mathbb{1}_{\{Z_1 = i\}} \\ & + \sum_{t=1}^{n} \sum_{i=1}^{K} (\log \psi_i(x_t))\, \mathbb{1}_{\{Z_t = i\}} \\ & + \sum_{t=1}^{n-1} \sum_{i,j=1}^{K} (\log A_{ij})\, \mathbb{1}_{\{Z_t = i, Z_{t+1} = j\}}\end{aligned}$$

HMM

- So that from the **e-step**

$$
\begin{aligned}
\bar{\ell}(\Theta) = \quad & \sum_{i=1}^{K} (\log \pi_i)\, \gamma_1^i \\
& + \sum_{t=1}^{n} \sum_{i=1}^{K} (\log \psi_i(x_m))\, \gamma_t^i \\
& + \sum_{t=1}^{n-1} \sum_{i,j=1}^{K} (\log A_{ij})\, \xi_t^{ij}
\end{aligned}
$$

- where

smoother $\quad \gamma_t^i := \mathbb{P}(\; Z_t = i \mid X_{1:T} = x_{1:T} \;)$

two-slice marginal $\quad \xi_t^{ij} = \mathbb{P}(\; Z_t = i,\; Z_{t+1} = j \mid X_{1:T} = x_{1:T} \;)$

HMM

- From the **m-step** we have the update rule

$$\pi_i = \frac{\gamma_1^i}{\sum_i \gamma_1^i}$$

$$\psi_i(x) = \frac{\sum_t \gamma_t^i \, \mathbb{1}_{x_t = x}}{\sum_t \gamma_t^i}$$

$$A_{ij} = \frac{\sum_t \xi_t^{i,j}}{\sum_t \sum_j \xi_t^{i,j}}$$

HMM

The following quantities are useful for estimating the model:

forward filter $$\alpha_t^i := \mathbb{P}(\ Z_t = i \mid X_{1:t} = x_{1:t}\ )$$

$$\eta_t := \mathbb{P}(\ X_t = x_t \mid X_{1:t-1} = x_{1:t-1}\ )$$

backwards filter $$\beta_t^i := \frac{\mathbb{P}(\ X_{t+1:T} = x_{t+1:T} \mid Z_t = i\ )}{\mathbb{P}(\ X_{t+1:T} = x_{t+1:T} \mid X_{1:t} = x_{1:t}\ )}$$

HMM

The forward recursion:

$$\alpha_1^i = \psi_i(x_1)\, \pi^i \qquad \textbf{initialize}$$

$$\tilde{\alpha}_t^i = \psi_i(x_t) \sum_{j=1}^{K} A_{ij}^\intercal\, \alpha_{t-1}^j \qquad \textbf{update}$$

$$\eta_t = \sum_{i=1}^{K} \tilde{\alpha}_t^i \qquad \textbf{normalization factor}$$

$$\alpha_t^i := \frac{\tilde{\alpha}_t^i}{\eta_t} \qquad \textbf{normalize}$$

HMM

The backward recursion:

$$\beta_T^i = 1 \qquad \textbf{initialize}$$

$$\beta_t^i = \frac{1}{\eta_{t+1}} \sum_{j=1}^{K} A_{ij}\, \beta_{t+1}^j\, \psi_j(x_{t+1}) \qquad \textbf{update}$$

HMM

The smoother can be obtained as follows

$$\gamma_t^i = \frac{1}{Z}\,\alpha_t^i\,\beta_t^i$$

$$Z = \sum_{j=1}^{K} \alpha_t^j\,\beta_t^j \qquad \text{normalization}$$

HMM

The smoothed two-slice marginal can be obtained as follows

$$\xi_t^{ij} = \frac{1}{Z}\,\alpha_t^i\,A_{ij}\,\psi_j(x_{t+1})\,\beta_{t+1}^j$$

$$Z = \sum_{i,j=1}^{K} \alpha_t^i\,A_{ij}\,\psi_j(x_{t+1})\,\beta_{t+1}^j \qquad \text{normalization}$$

HMM

Using **signed price changes** each second

Regime	π	A		ψ		
	INTC					
1	0.01	0.99	0.01	0.02	0.95	0.02
2	0.99	0.01	0.99	0.08	0.85	0.07
	SMH					
	π	A		ψ		
1	0.00	0.97	0.03	0.04	0.91	0.04
2	1.00	0.04	0.96	0.18	0.65	0.18
	NTAP					
	π	A		ψ		
1	0.04	0.96	0.04	0.01	0.97	0.02
2	0.96	0.06	0.94	0.15	0.68	0.16

HMM

Using **volume imbalance** (sell,neutral,buy) each second

Regime	π	A		ψ		
	INTC					
1	0.22	0.92	0.08	0.99	0.01	0.00
2	0.78	0.04	0.96	0.00	0.62	0.38
	SMH					
	π	A		ψ		
1	0.95	0.95	0.05	0.53	0.02	0.44
2	0.05	0.14	0.86	0.01	0.99	0.01
	NTAP					
	π	A		ψ		
1	0.73	0.98	0.02	0.00	0.61	0.39
2	0.27	0.09	0.91	0.99	0.01	0.00

HMM

Using **buy, sell order-flow** (one,one,both) each second

	INTC						
Regime	π	A			ψ		
1	0.49	0.91	0.00	0.09	0.74	0.23	0.03
2	0.00	0.01	0.99	0.00	0.46	0.35	0.19
3	0.51	0.02	0.00	0.98	0.95	0.04	0.01
	SMH						
Regime	π	A			ψ		
1	0.38	1.00	0.00	0.00	0.99	0.01	0.00
2	0.56	0.00	1.00	0.00	0.97	0.03	0.00
3	0.05	0.16	0.00	0.84	0.88	0.11	0.01