Module 4: Lesson 4

Regime changes and the EM algorithm: An application to VIX

WORLDQUANT UNIVERSITY

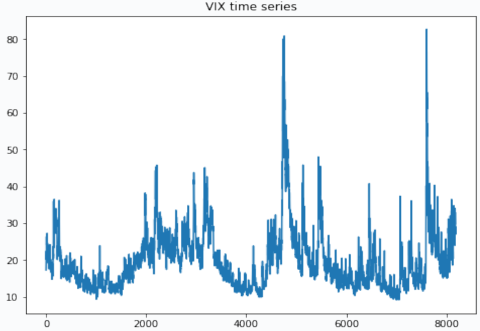
Outline

- An EM algorithm to estimate Hidden Markov Models
- An application to VIX.

A regime-switching model for VIX

In this lesson, we will develop a specific application of HMM estimation.

The aim is to estimate a regime-switching model for the daily realizations of the CBOE's VIX index, a popular measure of the stock market's volatility based on option volatilities for the S&P 500 index.



We assume that the observable VIX process is $y_t = \mu_t + \varepsilon_t$, where ε_t is drawn from a normal distribution with mean zero and variance σ_t^2.

The conditional mean and variance of the process at each time both depend on the outcome of a Markov process with N possible realizations, i.e., we have a tuple $\{(\mu_i, \sigma_i)_{i=1}^N\}$ and an associated transition matrix P.

Setting up the estimation problem

As in lesson 3, we are going to follow the Hamilton filter process. Further, we are going to apply a backward-smoothing method:

$$\xi_{t|T}(j) = \xi_{t|t}(j) \sum_{i=1}^{N} p_{ji} \frac{\xi_{t+1|T}(i)}{\xi_{t+1|t}(i)} \tag{1}$$

These smoothed probabilities are computed backwards, starting with $\xi_{T|T}(j)$ as it arises from the forward recursion of the Hamilton filter. This procedure is called the Kim smoother, due to Kim (1993).

Thus, the log-likelihood function can be expressed as:

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} log \ f(y_t|\mathcal{Y}_{t-1};\theta) = \sum_{t=1}^{T} log \ \sum_{i=1}^{N} \xi_{t|T}(i) \times \phi_t(i) \tag{2}$$

We are now developing a procedure that reduces the computational burden of this type of problem. This procedure is a version of the expectation-maximization (EM) algorithm.

EM algorithm: Expectation

In the first step, Expectation (E), we assume that we know the actual path of the unobservable Markov process $\mathcal{S}_{t-1}$ and the vector of parameters $\theta = \{(\mu_i, \sigma_i)_{i=1}^N, (p_{ij})_{i,j=1}^N\}$. In that case, we would have that:

$$\mathbb{P}(y_t, s_t = i | \mathcal{Y}_{t-1}, \mathcal{S}_{t-1}; \theta) = \phi_t(i) \prod_{j=1}^{N} p_{ji}^{\mathbb{I}(s_{t-1}=j;\theta)} \tag{3}$$

where $\mathbb{I}(s_{t-1} = j; \theta)$ is an indicator variable that denotes if state j was realized at $t-1$.

The "complete" log-likelihood function considers all the observations and potential state realizations:

$$\mathcal{L}(\theta; \mathcal{Y}_T, \mathcal{S}_T) = \sum_{t=1}^{T} \sum_{i=1}^{N} \mathbb{I}(s_t = i; \theta) log\ \phi_t(i) + \sum_{t=2}^{T} \sum_{i=1,j=1}^{N} \mathbb{I}(s_t = i; \theta) \mathbb{I}(s_{t-1} = j; \theta) log\ p_{ji} + \sum_{i=1}^{N} \mathbb{I}(s_1 = i; \theta) log\ \pi_{ij}$$

We compute the "expectation" by replacing the indicators $\mathbb{I}(s_t = j; \theta)$ with the estimated probabilities, $\xi_{t|T}(j)$.

EM algorithm: Maximization

In the Maximization step, instead of trying to optimize in a single step as in Lesson 3, we generate a series of parameter estimates $\theta^{(k)}$, such that:

$$\theta^{(k)} = \arg\max_{\theta} E\left[\mathcal{L}(\theta; \mathcal{Y}_T, \mathcal{S}_T)|\mathcal{Y}_T; \theta^{(k-1)}\right] \tag{4}$$

The normality assumption yields a sequence of estimates with analytical expressions:

$$\sigma_i^{(k)} = \sqrt{\frac{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)\left(y_t - \mu^{(k)}\right)^2}{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)}} \tag{5}$$

$$\mu_i^{(k)} = \frac{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i) y_t}{\sum_{t=1}^{T} \xi_{t|T}^{(k-1)}(i)} \tag{6}$$

The estimates are the standard deviations and sample averages of the data, weighted by the estimated probability that the process is in state i at each date t.

EM algorithm: Estimation of state probabilities

The k-th step estimates for the transition matrix and the initial distribution are given by:

$$p_{ij}^{(k)} = \frac{\sum_{t=2}^{T} p_{ij}^{(k-1)} \xi_{t-1|t-1}(i)\xi_{t|T}(j)/\xi_{t|t}(j)}{\sum_{t=2}^{T} \xi_{t-1|T}^{(k-1)}(i)} \tag{7}$$

$$\pi_i^{(k)} = \xi_{1|T}^{(k-1)}(i) \tag{8}$$

Notice the resemblance of the solution $p_{ij}^{(k)}$ with the MLE estimate for the transition probabilities we learned in Lesson 1 of this module.

With all these ingredients, we are ready to code the optimization algorithm and estimate the parameters of the model.

Summary of Lesson 4

In Lesson 4, we have looked at:

- An EM algorithm to estimate the underlying parameters of a HMM model

⇒ **References for this lesson:**

Hamilton, James D. *Time Series Analysis.* Princeton University Press, 1994. (see Chapter 22: "Modelling time series with changes in regime")

Kim, C-J. "Dynamic Linear Models with Markov-Switching." *Journal of Econometrics.* vol. 60, 1994, pp. 1–22.

Kole, E. "Markov Switching Models: An Example for a Stock Market Index." 2019. Available at SSRN: https://ssrn.com/abstract=3398954

TO DO NEXT: Now, please go to the associated Jupyter notebook for this lesson to study the implementation of the EM algorithm for the VIX.

With this lesson, we finalize Module 4 on Markov Processes as an introduction to Reinforcement Learning. Module 5 is dedicated to Dynamic Programming.