MTL-106
Probability and Stochastic Processes

Assignment - 2B

Deadline : 10th April 2024

In this assignment you'll learn about an application of DTMC in quantitative trading using Hidden Markov Models.

A consistent challenge for quantitative traders is the frequent behaviour modification of financial markets, often abruptly, due to changing periods of government policy, regulatory environment and other macroeconomic effects. Such periods are known colloquially as "market regimes" and detecting such changes is a common, albeit difficult process undertaken by quantitative market participants.

These various regimes lead to adjustments of asset returns via shifts in their means, variances/volatilities, serial correlation and covariances, which impact the effectiveness of time series methods that rely on stationarity.

This motivates a need to effectively detect and categorise these regimes in order to optimally select deployments of quantitative trading strategies and optimise the parameters within them. The modeling task then becomes an attempt to identify when a new regime has occurred and adjust strategy deployment, risk management and position sizing criteria accordingly.

Principal method for carrying out regime detection is to use a statistical time series technique known as a Hidden Markov Model. These models are well suited to the task as they involve inference on "hidden" generative processes via "noisy" indirect observations correlated to these processes. In this instance the hidden, or latent process is the underlying regime state, while the asset returns are the indirect noisy observations that are influenced by these states

Hidden Markov Model

A Hidden Markov model (HMM) is a Markov model in which the observations are dependent on a latent (or "hidden") Markov process (referred to as X). An HMM requires that there be an observable process Y whose outcomes depend on the outcomes of X in a known way. Since X cannot be observed directly, the goal is to learn about state of X by observing Y. By definition of being a Markov model, an HMM has an additional requirement that the outcome of Y at time $t = t_0$ must be "influenced" exclusively by the outcome of X at $t = t_0$ and that the outcomes of X and Y at $t < t_0$ must be conditionally independent of Y at $t = t_0$ given X at time $t = t_0$.

More formally, HMM is a discrete time stochastic process (X_n, Y_n) where

1. X_n is a DTMC (unobservable or hidden)
2. $P(Y_n \in A \mid X_n = x_n, \ldots, X_1 = x_1) = P(Y_n \in A \mid X_n = x_n)$, where A is some borel set.

The states of process X_n are called *hidden states* and $P(Y_n \in A|X_n = x)$ is called *emission probability*.

Building the mathematical model

We can think of hidden states in our markov model as "regimes" under which a market might be acting while the observations are the asset returns that are directly visible.

Suppose $Y_1, \ldots, Y_T$ are the asset returns at discrete times and $X_1, \ldots, X_T$ are the unobserved market regimes, then the joint density of $X_1, \ldots, X_T$ is given by,

$$
\begin{aligned}
P(X_{1:T}|Y_{1:T}) &= \frac{P(X_{1:T})P(Y_{1:T}|X_{1:T})}{P(Y_{1:T})} \\
&= \frac{P(X_1)\prod_{t=2}^{T} P(X_t|X_{t-1})\prod_{t=1}^{T} P(Y_t|X_t)}{P(Y_{1:T})}
\end{aligned}
$$

In above expression terms of form $P(X_t|X_{t-1})$ can be given by $K \times K$ state transition matrix. However, for the application considered here, namely observations of asset returns, the values are in fact continuous. This means the model choice for the emission probability is more complex. The common choice is to make use of a conditional multivariate Gaussian distribution

$$Y_t|X_t = x \sim \mathcal{N}(\mu_x, \sigma_x^2)$$

Depending upon the specified state and emission probabilities a Hidden Markov Model will tend to stay in a particular state and then suddenly jump to a new state and remain in that state for some time. This is precisely the behaviour that is desired from such a model when trying to apply it to market regimes. The regimes themselves are not expected to change too quickly (consider regulatory changes and other slow-moving macroeconomic effects). However, when they do change they are expected to persist for some time

Market Regimes

Applying Hidden Markov Models to regime detection is tricky since the problem is actually a form of *unsupervised learning*. That is, there is no "ground truth" or labelled data on which to "train" the model. In particular it is not clear how many regime states exist apriori. Are there two, three, four or more "true" hidden market regimes?

Answers to these questions depend heavily on the asset class being modelled, the choice of time frame and the nature of data utilised. For instance, daily returns data in equities markets often exhibits periods of calm lower volatility, even over a number of years, with exceptional periods of high volatility in moments of "panic" or "correction". Is it natural then to consider modelling equity indices with two states? Might there be a third intermediate state representing more vol than usual but not outright panic?

Simulated Data

For this assignment simulated returns data will be generated from separate Gaussian distributions, each of which represents a "bullish" or "bearish" market regime. The bullish returns draw from a Gaussian distribution with positive mean and low variance, while the bearish returns draw from a Gaussian distribution with slight negative mean but higher variance.

This is an example of five separate simulated market regimes, stream of returns is then utilised by a Hidden Markov Model in order to infer posterior probabilities of the regime states, given the sequence of observations.

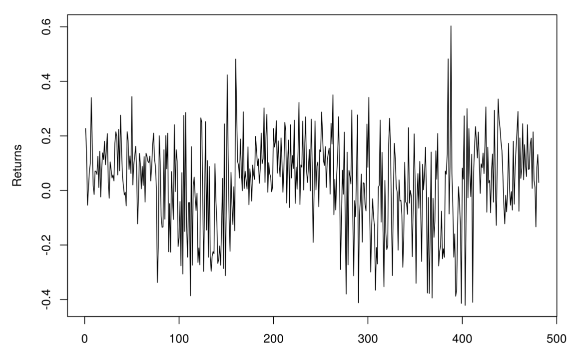

Figure 1: Asset Returns

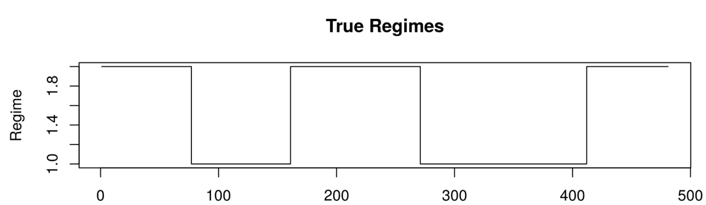

Figure 2: Market Regimes

Fitting HMM parameters

Given the observations Y, how well can we fit the transition probabilities and emission probabilities ? We'll use Baum-Welch algorithm (a special case of Expectation-Maximization algorithm)

The hidden states have a transition matrix A, observations are gaussian distributed with μ_x, σ_x^2 parameters for each hidden state x, initial distribution of hidden states is π. The BW algorithm finds local maximum for $\theta^* = \arg\max_\theta P(Y;\theta)$ (i.e. the HMM parameters θ that maximize the probability of the observation)

ALGORITHM

1. Set $\theta = (A, \mu, \sigma^2, \pi)$ with random initial conditions.
2. Let $\alpha_i(t) = P(Y_1 = y_1, \ldots, Y_t = y_t, X_t = i; \theta)$, the probability of seeing the observations $y_1, y_2, \ldots, y_t$ and being in state i at time t. This is found recursively as :
   (a) $\alpha_i(1) = \pi_i f_i(y_1)$,
   (b) $\alpha_i(t+1) = f_i(y_{t+1}) \sum_{j=1}^{K} \alpha_j(t) a_{ji}$.

   here f_i is density function of $\mathcal{N}(\mu_i, \sigma_i^2)$
3. Let $\beta_i(t) = P(Y_{t+1} = y_{t+1}, \ldots, Y_T = y_T \mid X_t = i; \theta)$ that is the probability of the ending partial sequence $y_{t+1}, \ldots, y_T$ given starting state i at time t. We calculate $\beta_i(t)$ as :
   (a) $\beta_i(T) = 1$
   (b) $\beta_i(t) = \sum_{j=1}^{K} \beta_j(t+1) a_{ij} f_j(y_{t+1})$.
4. $\gamma_i(t) = P(X_t = i \mid Y; \theta) = \dfrac{P(X_t = i, Y; \theta)}{P(Y; \theta)} = \dfrac{\alpha_i(t)\beta_i(t)}{\sum_{j=1}^{K} \alpha_j(t)\beta_j(t)}$

   which is the probability of being in state i at time t given the observed sequence Y and the parameters θ
5. $\xi_{ij}(t) = P(X_t = i, X_{t+1} = j \mid Y; \theta) = \dfrac{P(X_t = i, X_{t+1} = j, Y; \theta)}{P(Y; \theta)} = \dfrac{\alpha_i(t) a_{ij} \beta_j(t+1) b_j(y_{t+1})}{\sum_{k=1}^{K} \sum_{w=1}^{K} \alpha_k(t) a_{kw} \beta_w(t+1) b_w(y_{t+1})}$

   which is the probability of being in state i and j at times t and $t+1$ respectively given the observed sequence Y and parameters θ.
6. $\pi_i^* = \gamma_i(1)$, [expected frequency spent in state i at time 1.]
7. $a_{ij}^* = \dfrac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} \gamma_i(t)}$,
8. $\mu_i^* = \dfrac{\sum_{t=1}^{T} y_t \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}$,
9. $(\sigma_i^2)^* = \dfrac{\sum_{t=1}^{T} (y_t - \mu_i)^2 \gamma_i(t)}{\sum_{t=1}^{T} \gamma_i(t)}$,

   These steps are now repeated iteratively until a desired level of convergence.

Using this algorithm, if we fit the HMM and compute the regimes posterior probabilities on above simulated data we get the following plot of probabilities.

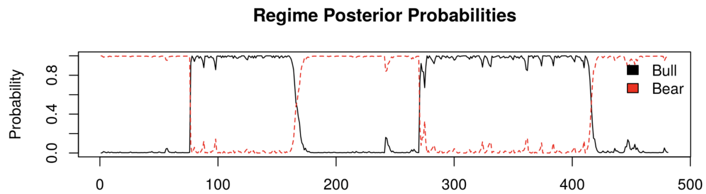

Figure 3: Posterior probabilities

Computing posterior probabilities

Once HMM model parameters are learned computing posterior probabilities is straight forward,

$$\begin{aligned} P(X_t = i | Y_t = y) &= P(X_t = i) P(Y_t = y | X_t = i) \\ &= \sum_{j=1}^{K} P(X_t = i \mid X_1 = j) \pi_j f(y; \mu_i, \sigma_i^2) \end{aligned}$$

denominator is omitted as values can be normalized later.

Coding part

Your task is to find the market regimes given the returns data as computed by HMM trained using BW-algorithm. Initialize $\mu_0 = \mu_1 = 0, \sigma_0^2 = \sigma_1^2 = 1$

Input

First line of the input is initialization of A as $a_{00}, a_{01}, a_{10}, a_{11}$. (here 0=bull, 1=bear).
Second line of the input is π_0, π_1
Third line of the input contains T, i.e. the number of time steps of data.
Subsequent T lines contains a value Y_i which is return on i^{th} day for $1 \leq i \leq T$.

Output

i^{th} line is market regime on that day which either "Bullish" or "Bearish" (without the quotes) for $1 \leq i \leq T$

Constraints and Time-Limit

$T \leq 10{,}000$ Time-Limit = 3sec

Sample Inputs and Outputs are provided separately (see MSTeams A2 channel), but your code will be tested on a much larger and exhaustive dataset so make sure to handle all the edge cases.

You are expected to write your own code from scratch using only the standard library of C++ or python. Any instances of plagiarism (either among students or copy from the internet) will be awarded 0.