Hidden Markov Model

QING SHI, JINYI LUO, YUEER LI
APRIL 17, 2019

ANDREY MARKOV (1856-1922)

RUSSIAN MATHEMATICIAN

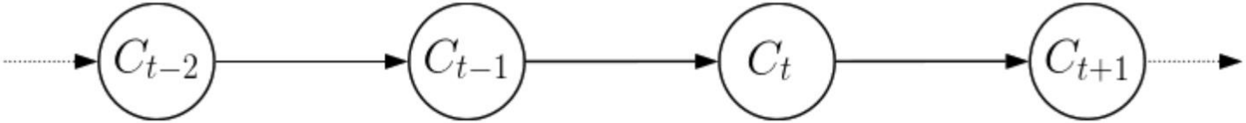
WORKED ON STOCHASTIC PROCESSES

PRIMARY ON MARKOV CHAINS & MARKOV PROCESS

Markov Chain Definition

- A sequence of discrete random variables $\{C_t : t \in \mathbb{N}\}$ is said to be a (discrete-time) **Markov chain** (MC) if, for all $t \in \mathbb{N}$, it satisfies the **Markov property**

$$\mathrm{P}(C_{t+1} \mid C_t, \ldots, C_1) = \mathrm{P}(C_{t+1} \mid C_t).$$



Transition Probability

- Let S be a finite set and $\gamma(i,j)$ a transition function for S. Let $\pi_0(i)$ be a probability distribution on S. Then the Markov chain corresponding to initial distribution π_0 and transition probabilities $\gamma(i,j)$ is the stochastic process C_n such that

$$P(C_0 = i) = \pi_0(i),$$

$$P(C_{n+1} = i_{n+1} | C_n = i_n, C_{n-1} = i_{n-1}, \dots, C_2 = i_2, C_1 = i_1) = \gamma(i_n, i_{n+1})$$

Which implies

$$P(C_{n+1} = i_{n+1} | C_n = i_n) = \gamma(i_n, i_{n+1})$$

Important Quantities

- Transition Probability Matrix (One step transition matrix)

$$\Gamma = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mm} \end{pmatrix}$$

- We can also consider the transition probabilities for longer times (m step).

$$\text{Let } \gamma^m(i,j) = P(C_{n+m} = j \,|C_n = i)$$

Important Quantities

- If these probabilities are homogenous, then they satisfy (Chapman-Kolmogorov equations)

$$\gamma^{t+u}(i,j) = \sum_k \gamma^t(i,k)\gamma^u(k,j)$$

- In Matrix form,

$$\Gamma(t+u) = \Gamma(t)\Gamma(u)$$

Important Quantities

- Unconditional Probabilities

$$u_i(t) = \mathrm{P}(C_t = i)$$

$$u(t) = (\mathrm{P}(C_t = 1), \dots, \mathrm{P}(C_t = m)), \qquad t \in \mathbb{N}.$$

$$u(t+1) = u(t)\Gamma.$$

Joint Probability

- $P(x_0 = i_0, \ldots, x_n = i_n)$

$$= P(x_0 = i_0)P(x_1 = i_1 \mid x_0 = i_0)P(x_2 = i_2 \mid x_1 = i_1, x_0 = i_0) \ldots$$

$$P(x_n = i_n \mid x_{n-1} = i_{n-1}, x_{n-2} = i_{n-2}, \ldots)$$

$$= P(x_0 = i_0)P(x_1 = i_1 \mid x_0 = i_0)P(x_2 = i_2 \mid x_1 = i_1) \ldots P(x_n = i_n \mid x_{n-1} = i_{n-1})$$

$$\boxed{= \pi_0 P(i_0, i_1)P(i_1, i_2) \ldots P(i_{n-1}, i_n)}$$

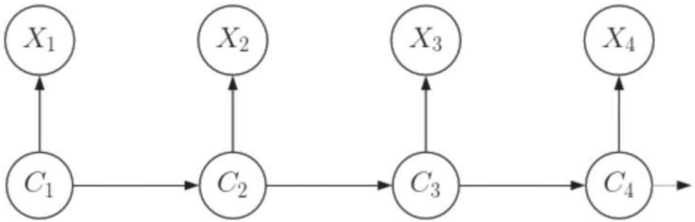
Stationary Distribution

A Markov chain with transition probability matrix Γ is said to have stationary distribution δ (a row vector with non-negative elements) if

$$\delta\Gamma = \delta \ and \ \delta 1 = 1$$

A Markov chain started from its stationary distribution will continue to have that distribution at all subsequent time points, and we shall refer to such a process as a stationary Markov chain .

Hidden Markov Model



Definition

A hidden Markov model $\{X_t : t \in N\}$ is a particular kind of dependent mixture. With $X^{(t)}$ and $C^{(t)}$ representing the histories from time 1 to time t, one can summarize the simplest model of this kind by:

$$P\,(C_t \mid C^{(t-1)}) = P\,(C_t \mid C_{t-1}), \qquad t = 2, 3, \ldots$$
$$P\,(X_t \mid X^{(t-1)}, C^{(t)}) = P\,(X_t \mid C_t), \qquad t \in N$$

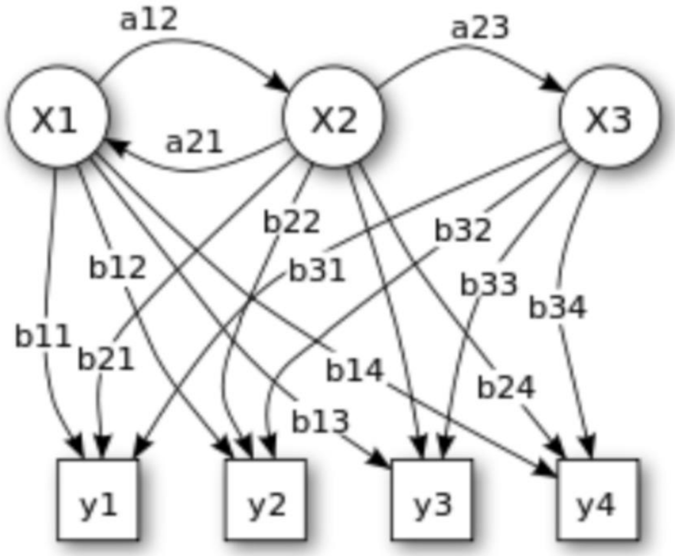
Notation

{Ct} unobserved process, satisfied Markov Property

{Xt} state-dependent process, Xt only depend on current state Ct.

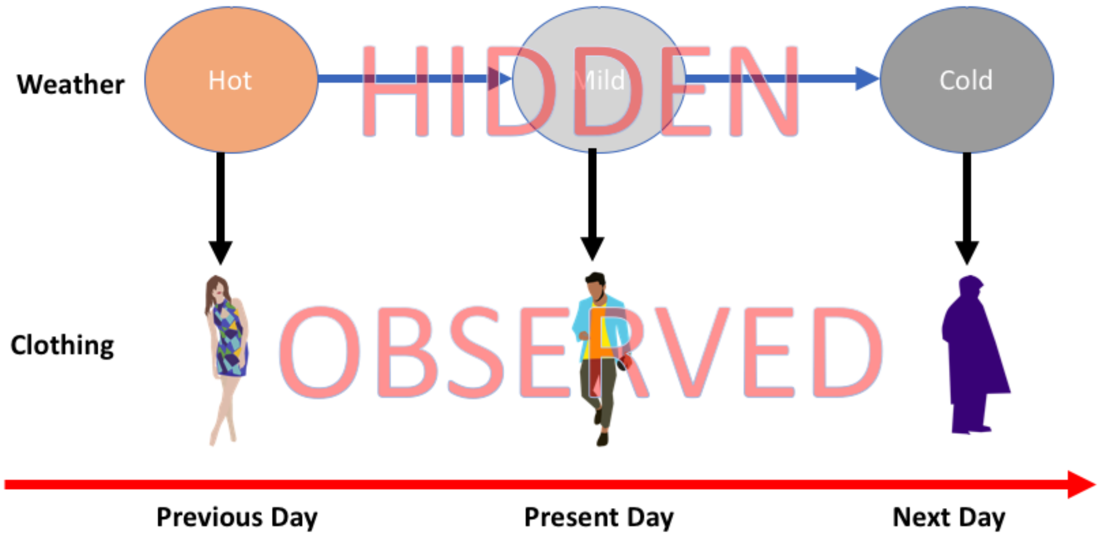
If {Ct} has m state, {Xt} is m-state HMM.

Hidden Markov Model



X: hidden states
Y: possible observations
a: transition probability
b: emission probability

Hidden Markov Model



Hidden variable:
hot, mild or cold
Observed variable:
type of clothing worn
Arrow:
transitions from a hidden state to another hidden state or from a hidden state to an observed variable.

Hidden Markov Model

- Initial state probability: π_1
- Transition probability:

	Hot	Mild	Cold
Hot	0.6	0.3	0.1
Mild	0.4	0.3	0.2
Cold	0.1	0.4	0.3

What's $P(X_1 = winter, X_2 = casual, C_1 = mild, C_2 = hot)$?

$$P(X_1 = winter, X_2 = casual, C_1 = mild, C_2 = hot)$$
$$= P(C_1 = mild)P(X_1 = winter|C_1 = mild)$$
$$P(C_2 = hot|C_1 = mild)P(X_2 = casual|C_2 = hot)$$
$$= 0.3 * 0.2 * 0.3 * 0.8 = 0.0144$$

- Emission probability:

	Hot	Mild	Cold
Casual	0.8	0.19	0.01
Semi-Casual	0.5	0.4	0.1
Winter	0.01	0.2	0.79

$$i.e. P(X_1 = winter|C_1 = mild)$$
$$= 0.2$$
$$P(X_2 = casual|C_2 = hot)$$
$$= 0.8$$

Hidden Markov Model

- Joint Distribution

$$P(X_1, X_2, \dots, X_t, C_1, C_2, \dots, C_t)$$
$$= P(C_1)P(X_1|C_1)P(X_2, \dots, X_t, C_2, \dots, C_t)$$
$$= P(C_1)P(X_1|C_1)P(C_2|C_1)P(X_2|C_2) \dots P(C_t|C_{t-1})P(X_t|C_t)$$
$$= P(C_1)P(X_1|C_1) \prod_{t=2}^{t} P(C_t|C_{t-1})P(X_t|C_t)$$

Hidden Markov Model

- Marginal distribution of X_t - Univariate Distributions

$$
\begin{aligned}
&P(X_t = x) \\
&= \sum_{i=1}^{m} P(X_t = x, C_t = i) \\
&= \sum_{i=1}^{m} P(C_t = i) P(X_t = x | C_t = i) \\
&= \sum_{i=1}^{m} u_i(t) P_i(x) \\
&= \begin{pmatrix} u_1(t) & \dots & u_m(t) \end{pmatrix} \begin{pmatrix} P_1(t) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & P_m(t) \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \\
&= u(t) P(x) 1' \\
&= u(1) \Gamma^{t-1} P(x) 1' \\
&= \delta P(x) 1'
\end{aligned}
$$

Hidden Markov Model

- Marginal distribution of X_t - Bivariate Distributions

$$
\begin{aligned}
&P(X_t = v, X_{t+k} = w) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} P(\, X_t = v, X_{t+k} = w, C_t = i, C_{t+k} = j) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} P(C_t = i) P(X_t = v | C_t = i)\ P(C_{t+k} = j | C_t = i) P(X_{t+k} = w | C_{t+k} = j) \\
&= \sum_{i=1}^{m} \sum_{j=1}^{m} u_i(t) P_i(v) \gamma_{ij}(k) P_j(w) \\
&= u(t) P(v) \Gamma^k P(w) 1' \\
&= \delta P(v) \Gamma^k P(w) 1'
\end{aligned}
$$

Hidden Markov Model

Likelihood:

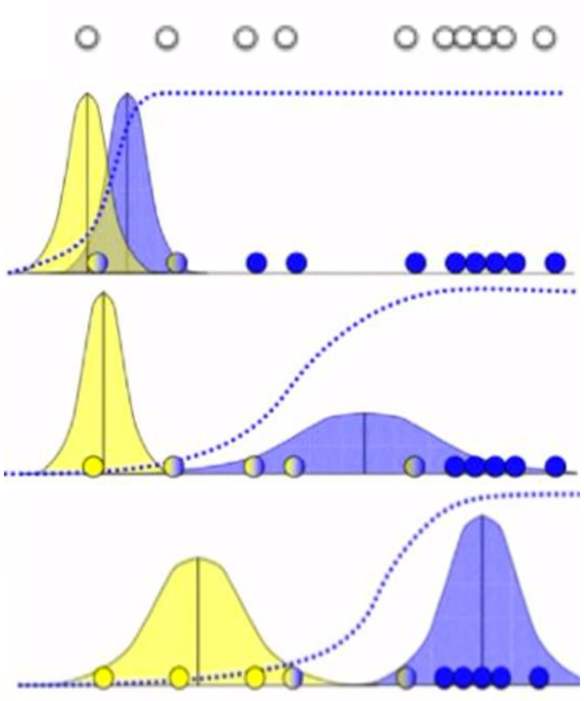
The likelihood is given by

$$L_T = \delta P(x_1)\Gamma P(x_2) \ldots \Gamma P(x_T) 1'$$

If δ, the distribution of C_1, is the stationary distribution of the Markov Chain, then in addition

$$L_T = \delta \Gamma P(x_1)\Gamma P(x_2) \ldots \Gamma P(x_T) 1'$$

Expectation Maximization (EM) Algorithm



$$P(X_i|b) = \frac{1}{\sqrt{2\pi\sigma_b}} \exp(-\frac{(X_i - \mu_b)^2}{2{\sigma_b}^2})$$

$$b_i = P(b|X_i) = \frac{P(X_i|b)P(b)}{P(X_i|b)P(b) + P(X_i|a)P(a)}$$

$$a_i = P(a|X_i) = 1 - b_i$$

$$\mu_b = \frac{b_1X_1 + b_2X_2 + \cdots + b_nX_n}{b_1 + b_2 + \cdots + b_n}$$

$${\sigma_b}^2 = \frac{b_1(X_1 - \mu_1)^2 + \cdots + b_n(X_n - \mu_n)^2}{b_1 + b_2 + \cdots + b_n}$$

$$\mu_a = \frac{a_1X_1 + a_2X_2 + \cdots + a_nX_n}{a_1 + a_2 + \cdots + a_n}$$

$${\sigma_a}^2 = \frac{a_1(X_1 - \mu_1)^2 + \cdots + a_n(X_n - \mu_n)^2}{a_1 + a_2 + \cdots + a_n}$$

Central Issues in HMM

1. Evaluation Problem: $P(X_t|C_t)$
2. Decoding Problem: $P(C_t|X_t)$ Posterior Probability
3. Learning Problem: Expectation Maximization / Baum-Welch Algorithm
   - Transition probability
   - Emission probability
   - Distribution of observation process

Hidden Markov Models for Regime Detection

Packages Used:

depmixS4

quantmod

TTR

forecast

Simulated Data

k=5 (Each of the k regimes will be bullish or bearish)

N_k(days of returns) $\in [50,150]$

Bull market is distributed as $N(0.1,0.1)$

Bear market is distributed as $N(-0.05,0.2)$

Reference

Lavrenko, V. (2014, January 19). Expectation Maximization: How it works. Retrieved from https://www.youtube.com/watch?v=iQoXFmbXRJA

(n.d.). Hidden Markov Models for Regime Detection using R. Retrieved from https://www.quantstart.com/articles/hidden-markov-models-for-regime-detection-using-r?fbclid=IwAR2-IcYOtnBwZvxBY164wzOzXcWoPcxML8NiFowCQD693I5xQURr0w6J_p8