

Stock Price Prediction Through the Mixture of Gaussian Processes via the Precise Hard-cut EM Algorithm

(✉)

Jiwei Ma¹, Jie Chen², and Jun Wang³
¹100871, ²100871, ³100871
jwma@math.pku.edu.cn

Abstract.

Stock price prediction is a challenging task due to the high volatility and non-stationarity of stock prices. In this paper, we propose a novel method for stock price prediction based on the mixture of Gaussian processes (MGP) via the precise hard-cut EM algorithm. The MGP model is able to capture the complex and non-linear relationship between stock prices and time. The precise hard-cut EM algorithm is used to estimate the parameters of the MGP model. The experimental results show that the proposed method outperforms the traditional methods in terms of prediction accuracy and robustness.

Keywords:

Stock price prediction, Mixture of Gaussian processes, Precise hard-cut EM algorithm.

1 Introduction

Stock price prediction is a challenging task due to the high volatility and non-stationarity of stock prices. In this paper, we propose a novel method for stock price prediction based on the mixture of Gaussian processes (MGP) via the precise hard-cut EM algorithm. The MGP model is able to capture the complex and non-linear relationship between stock prices and time. The precise hard-cut EM algorithm is used to estimate the parameters of the MGP model. The experimental results show that the proposed method outperforms the traditional methods in terms of prediction accuracy and robustness.

1. Introduction

2. Related Work

3. Methodology

4. Experiments

5. Conclusion

6. Acknowledgments

7. References

$\hat{\mu}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}$, $\hat{\sigma}_i^2 = \frac{1}{n} \sum_{j=1}^n (x_{ij} - \hat{\mu}_i)^2$, $i = 1, 2, \dots, C$.

The EM algorithm for MGP is as follows:

- Initialize $\mu_i, \sigma_i^2, \pi_i, i = 1, 2, \dots, C$.
- E-step: Compute the posterior probabilities $\hat{\mu}_i, \hat{\sigma}_i^2, i = 1, 2, \dots, C$.
- M-step: Update the parameters $\mu_i, \sigma_i^2, \pi_i, i = 1, 2, \dots, C$.
- Repeat steps 2 and 3 until convergence.

2 The Precise Hard-cut EM Algorithm for MGPs

2.1 The MGP Model

Let $X = \{x_{ij}\}$ be a $n \times C$ matrix of observations, where x_{ij} is the value of the i -th variable for the j -th individual. Let $\theta = \{\pi, \mu, \sigma\}$ be the parameter vector, where $\pi = (\pi_1, \dots, \pi_C)$, $\mu = (\mu_1, \dots, \mu_C)$, and $\sigma = (\sigma_1, \dots, \sigma_C)$.

$$\pi_i \geq 0, \quad i = 1, 2, \dots, C, \quad \sum_{i=1}^C \pi_i = 1 \tag{1}$$

$$\mu_i \in \mathbb{R}, \quad i = 1, 2, \dots, C, \tag{2}$$

$$\sigma_i > 0, \quad i = 1, 2, \dots, C, \tag{3}$$

$$\begin{aligned}
 & \text{The likelihood function is } L(\theta) = \prod_{j=1}^n \sum_{i=1}^C \pi_i \prod_{k=1}^C \frac{1}{\sigma_k} \exp\left\{-\frac{1}{2} \sum_{k=1}^C \frac{(x_{kj} - \mu_k)^2}{\sigma_k^2}\right\} \\
 & \text{The log-likelihood function is } \ln L(\theta) = \sum_{j=1}^n \left\{ \sum_{i=1}^C \pi_i \sum_{k=1}^C \left[-\frac{1}{2} \frac{(x_{kj} - \mu_k)^2}{\sigma_k^2} - \ln \sigma_k \right] \right\}
 \end{aligned}$$

The EM algorithm for MGP is as follows:

- Initialize $\mu_i, \sigma_i^2, \pi_i, i = 1, 2, \dots, C$.
- E-step: Compute the posterior probabilities $\hat{\mu}_i, \hat{\sigma}_i^2, i = 1, 2, \dots, C$.
- M-step: Update the parameters $\mu_i, \sigma_i^2, \pi_i, i = 1, 2, \dots, C$.
- Repeat steps 2 and 3 until convergence.

3 Stock Price Prediction

3.1 The General Prediction Model

Let $\{x_t\}$ be a time series of stock prices. The general prediction model is given by $\hat{x}_t = f(x_{t-1}, x_{t-2}, \dots, x_{t-k})$, where k is the lag order. The prediction error is defined as $e_t = x_t - \hat{x}_t$.

1, 2, 3, 4, τ 1, 2, 3, 4, $f_{i,t}$, 700
 $\frac{1}{\sigma}$, σ , 700, 600
 100

3.2 Prediction Results and Comparisons

(1) $\hat{y}_t - \hat{y}_t \sigma$

- (1) $\hat{y}_t - \hat{y}_t \sigma$ 12 $f_{i,t}$
- (2) $\hat{y}_t - \hat{y}_t \sigma$ 13 $f_{i,t}$
- (3) $\hat{y}_t - \hat{y}_t \sigma$ $f_{i,t}$

$$\sqrt{\frac{1}{L} \sum_{t=1}^L (\hat{y}_t - \hat{y}_t)^2} \tag{10}$$

\hat{y}_t , 5, 16.00, 2014

1. $f_{i,t}$

3, $\tau = 1$, 14, 15.

Table 1.

τ									
1	1	21.1933	0.2224	21.2151	0.2218	21.2106	0.2194	21.5791	0.2204
1	2	31.8931	0.3174	32.5150	0.3258	31.5769	0.3134	33.4016	0.3188
1	3	38.8945	0.3713	40.4267	0.3864	39.1431	0.3654	41.3143	0.3784
1	4	43.9104	0.4144	45.9821	0.4207	43.8947	0.4030	47.9299	0.4159
2	1	21.2464	0.2192	21.4326	0.2209	21.2879	0.2195	21.8939	0.2216
2	2	32.0199	0.3133	32.8077	0.3428	31.8553	0.3146	33.5479	0.3232
2	3	39.3107	0.3737	41.6617	0.3918	38.0136	0.3556	47.1278	0.3621
2	4	43.7612	0.3940	45.8229	0.4418	43.7895	0.4016	53.1479	0.3975
3	1	21.0782	0.2183	21.4797	0.2203	21.0879	0.2206	22.9824	0.2272
3	2	32.0317	0.3192	33.1330	0.3547	31.5318	0.3159	36.3263	0.3126
3	3	39.1218	0.3448	42.0259	0.4089	38.5492	0.3545	43.5631	0.3579
3	4	44.9140	0.3901	52.6917	0.4585	43.7351	0.3726	57.8089	0.4096
4	1	21.1804	0.2225	21.8933	0.2212	21.1461	0.2219	22.9345	0.2341
4	2	33.3603	0.3125	33.3099	0.4030	32.1945	0.3104	38.2823	0.3131
4	3	39.5789	0.3695	43.0564	0.4681	39.1205	0.3586	54.2466	0.3705
4	4	45.9405	0.4067	58.1039	0.5645	45.3620	0.3960	72.0162	0.4536

600

2, 3

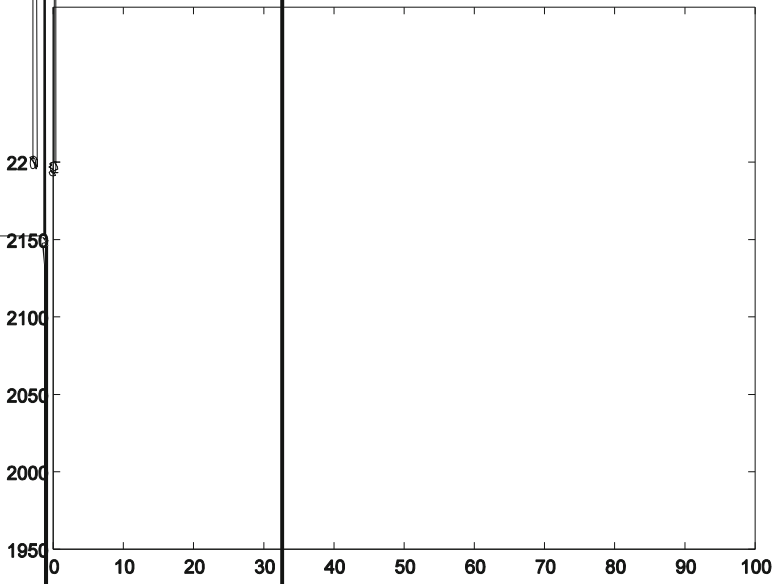
$\tau = 1$, 100

21.0782 0.2183

± 0.2 ± 0.4 2, 3.

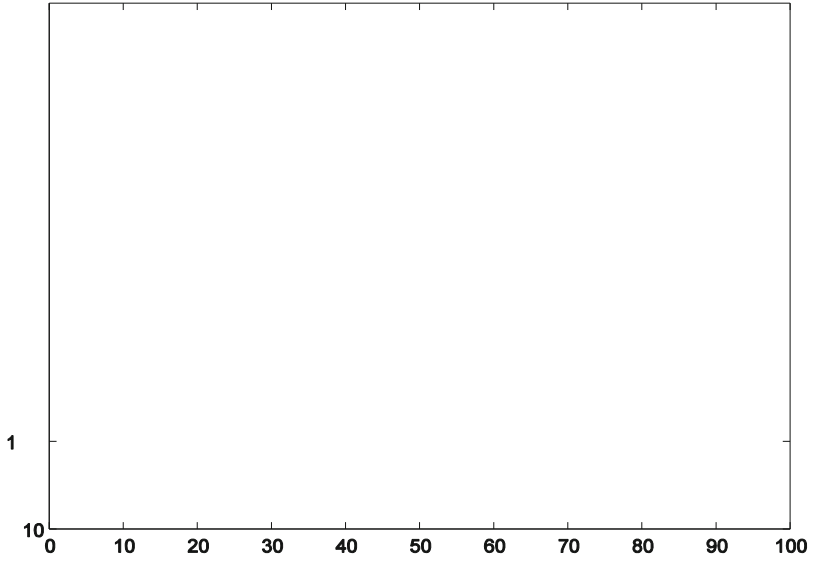
2.

6



(a)

(b)



(a)

(b)

(a)

(

4 Conclusion

Acknowledgement.

61171138.

References

1. *Journal of Applied Mathematics*, **22**, 130 (2010)
2. *Journal of Applied Mathematics*, **23**(4), 14–15, 38 (2004)
3. *Journal of Applied Mathematics*, **41**(4), 14–19 (2011)
4. *Journal of Applied Mathematics*, (2006)
5. *Journal of Applied Mathematics*, **13**, 654–660 (2000)
6. *Journal of Applied Mathematics*, **18**, 883–890 (2006)
7. *Journal of Applied Mathematics*, (2014), 858, 68–75. (2014)
8. *Journal of Applied Mathematics*, **31**(4), 383–400 (2011)
9. *Journal of Applied Mathematics*, **3**, 1011–1016 (2013)
10. *Journal of Applied Mathematics*, 1126–1133 (2012)
11. *Journal of Applied Mathematics*, **24**(1), 42–45 (2007)
12. *Journal of Applied Mathematics*, (2011), 2011, 6676, 165–174. (2011)
13. *Journal of Applied Mathematics*, **31**, 145–153 (2014)
14. *Journal of Applied Mathematics*, **7**(13), 1296–1306 (2014)
15. *Journal of Applied Mathematics*, (2014), 2014, 8866, 576–583. (2014)