Ḇ003-0530˝ 2016˛ 08-0960-13

Ḇ

1 2 1 1
1. 100871
2. 300401

Ḇ GP
GP MGP GP MGP
GP
GP MGP MGP

Ḇ
Ḇ TP18 ḆA **DOI** Ḇ10. 16798/j. issn. 1003-0530. 2016. 08. 11

From Gaussian Processes to the Mixture of Gaussian Processes A Survey

ZHOU Ya-tong[1 2] CHEN Zi-yi[1] MA Jin-wen[1]
1. School of Mathematical Science and LMAM Peking University Beijing 100871 China
2. School of Electronic and Information Engineering Hebei University of Technology Tianjin 300401 China

Abstract Gaussian process GP model is a paradigmatic machine learning model that combines the advantages of both kernel learning method and Bayesian inference mechanism and thus has become a very popular area in machine learning in recent years. As an extension of the GP model the Mixture of Gaussian Processes MGP fits datasets more effectively and thus it has a better ability of learning and generalization. However

2015-09-30 2016-02-22
61171138 15YJA630108 2014M561053

[6] GP

GP

Tresp [7] 2000

Mixture of Gaussian Processes MGP

MGP

GP

1996 SVM 1992

Rasmussen [8] GP

20 GP

GP

[9] [10]

[9,11]

GP [12] GP MGP

GP

GP

MGP MGP

GP MGP

2

1 2 6 10

1 GP

N $\boldsymbol{x}_1$ $\boldsymbol{x}_2$ … $\boldsymbol{x}_N$

t_1 t_2 … t_N $f(\boldsymbol{x})$

$\boldsymbol{x}_{N+1}$ $f(\boldsymbol{x}_{N+1})$ GP

$f(\boldsymbol{x})$ $\boldsymbol{t}_N =$

$(t_1\ t_2\ \dots\ t_N)^{\mathrm{T}}$ $t(\boldsymbol{x})$ $\boldsymbol{X}_N =$

$(\boldsymbol{x}_1\ \boldsymbol{x}_2\ \dots\ \boldsymbol{x}_N)^{\mathrm{T}}$ $f(\boldsymbol{x})$ $t(\boldsymbol{x})$

$t(\boldsymbol{x}) = f(\boldsymbol{x}) + \mathrm{e}(\boldsymbol{x})$ $\mathrm{e}(\boldsymbol{x})$

σ_ν^2 $f(\boldsymbol{x})$ H

$$f(\boldsymbol{x}) = \sum_{h=1}^{H} w_h \Phi_h(\boldsymbol{x})$$

$\boldsymbol{w} = (w_1\ w_2\ \dots\ w_H)^{\mathrm{T}}$ $\boldsymbol{w} \sim N(0, \sigma_w^2 \boldsymbol{I}_H)$ $f(\boldsymbol{x})$

$t(\boldsymbol{x})$ $\boldsymbol{t}_N \sim N(0, \boldsymbol{C}_N)$

$\boldsymbol{C}_N$

$$C(\boldsymbol{x}_n, \boldsymbol{x}_m) = \sigma_w^2 \sum_{h=1}^{H} \Phi_h(\boldsymbol{x}_n) \Phi_h(\boldsymbol{x}_m) + \sigma_\nu^2 \delta_{nm} \quad (1)$$

$n = m$ $\delta_{nm} = 1$ $\delta_{nm} = 0$

$\boldsymbol{x}_{N+1}$ t_{N+1}

$$P(t_{N+1} | \boldsymbol{t}_N) = P(t_{N+1}, \boldsymbol{t}_N) / P(\boldsymbol{t}_N) \quad (2)$$

GP

$$P(t_{N+1}, \boldsymbol{t}_N) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{t}_N^{\mathrm{T}}\ t_{N+1}) \boldsymbol{C}_{N+1}^{-1} \begin{bmatrix} \boldsymbol{t}_N \\ t_{N+1} \end{bmatrix}\right\} \quad (3)$$

$\boldsymbol{C}_{N+1}$ $\boldsymbol{k}$ $\boldsymbol{k}^{\mathrm{T}}$ κ

t_{N+1}

$$P(t_{N+1} | \boldsymbol{t}_N) \sim N(\hat{t}_{N+1}, \sigma_{\hat{t}_{N+1}}^2) \quad (4)$$

$\hat{t}_{N+1} = \boldsymbol{k}^{\mathrm{T}} \boldsymbol{C}_N^{-1} \boldsymbol{t}_N$ $\sigma_{\hat{t}_{N+1}}^2 = \kappa - \boldsymbol{k}^{\mathrm{T}} \boldsymbol{C}_N^{-1} \boldsymbol{k}$ $f(\boldsymbol{x}_{N+1})$ $= \hat{t}_{N+1}$ GP $\boldsymbol{x}_{N+1}$

2 GP

GP 2

$P(t_{N+1} | \boldsymbol{t}_N)$ $P(t_{N+1} |$ $\boldsymbol{t}_N)$ $P(t_{N+1} | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N)$

$$P(t_{N+1} = 1 | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N) = \int P(t_{N+1} = 1 | f(\boldsymbol{x}_{N+1})) P(f(\boldsymbol{x}_{N+1}) | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N) \mathrm{d}f(\boldsymbol{x}_{N+1}) \quad (5)$$

$$P(t_{N+1} = 1 | f(\boldsymbol{x}_{N+1})) = \frac{1}{1 + \mathrm{e}^{-f(\boldsymbol{x}_{N+1})}} \quad (6)$$

$$P(f(\boldsymbol{x}_{N+1}) | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N) = \int P(f(\boldsymbol{x}_{N+1}), \boldsymbol{f}_N | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N) \mathrm{d}\boldsymbol{f}_N = \int P(\boldsymbol{f}_{N+1} | \boldsymbol{X}_{N+1}, \boldsymbol{t}_N) \mathrm{d}\boldsymbol{f}_N \quad (7)$$

$\boldsymbol{f}_N = (f(\boldsymbol{x}_1)\ f(\boldsymbol{x}_2)\ \dots\ f(\boldsymbol{x}_N))^{\mathrm{T}}$ $P(\boldsymbol{f}_{N+1} | \boldsymbol{X}_{N+1}$

$$t_N \propto P \; f_{N+1} \mid X_{N+1} \; \prod_{n=1}^{N+1} P \; t_n \mid f \; x_n \qquad P \; f_{N+1} \mid X_{N+1} \propto \exp\left(-\frac{1}{2} f_{N+1}^{\mathrm{T}} K_{N+1}^{-1} f_{N+1}\right)$$

6 6 7

5 5 GP

[4 13] [14]

[15] 5

3

GP

1 GP

GP

GP

GP

[16]

GP

GP

[10]

Seeger [17]

Schwaighofer [18]

GP

Sundararajan [19] Geisser

2 GP

GP

Chatzis [20] GP

Soh GP

[21] Snelson [22] GP

Boyle [23]

GP

Law-rence [24] [25] GP

Pillonetto [26] GP

Gilboa E. GP

[27] GP

Zhao [28]

Dallaire [29]

GP

3 GP

GP

Williams [30] GP

Opper [31]

Opper-Vivarelli OU

Sollich [32]

Malzahn [33]

Kakade [34] GP

4 GP

GP

GP

Mackay [6]

[35]

GP KD [36]

Nyström [37] GP

GP

[4,13]

[14]

[15] Chalupka

[38] GP

GP

[39]

[39] Snelson [40]

Quiñonero-Candela [39] GP

5 GP

GP

GP

GP

RBF SVM

Rasmussen [10] 6 GP

Kriging

[6]

Mackay [6] GP RBF

GP

RBF Sollich [41] GP

SVM SVM

GP

Gestel [42]

GP SVM

6 GP

Rasmussen [43] GP

[12]

GP

Engel [44]

Ko [45]

GP

GP

Wang [46] GPDM

[47]

GPDM Deisenroth [48]

GPDP

Amoto [49]

GP

7

GPLVM GP

[50]

GPLVM PPCA

Lawrence [51] PPCA

GPLVM GPLVM

GPLVM

Lawrence [52]

GPLVM

GPLVM

Urtasun [53] GPLVM

8 GP

GP

Cheng [54] JAFFE

GP

GP

[55] GP

[56]

[57] [58] [59]

9 GP

GP

GP

GP Ľ

GP

GP

GP

"

$\boldsymbol{\mu}$ x_i $z_i=j$ | θ_j $\boldsymbol{Q}$ x_i $z_i=j$ x_i $z_i=j$ | θ_j j

$\boldsymbol{\mu}=0$ 10

GP

ME MGP

MGP

GP

ME GP ME

ME [66]

6

MGP

MGP

6.1 MGP

MGP 9

10

[7,61,67-81] [60,62-65,82-88]

[62]

MGP [62]

$$P(\boldsymbol{t}_N \mid \boldsymbol{X}_N, \boldsymbol{\mu}) = \sum_{\boldsymbol{z}_N} \prod_i \Pr(z_i \mid x_i, \tilde{\theta}) \prod_j P(t_k, z_k=j \mid x_k, z_k=j, \theta_j^{GP}) \quad (11)$$

$$P(\boldsymbol{X}_N, \boldsymbol{t}_N \mid \boldsymbol{\theta}) = \sum_{\boldsymbol{z}_N} \prod_i \Pr(z_i \mid \tilde{\theta}) \prod_j P(t_k, z_k=j \mid x_k, z_k=j, \theta_j^{GP}) P(x_k, z_k=j \mid \tilde{\theta}_j) \quad (12)$$

j $P(t_i, z_i=j \mid x_i, z_t=j, \theta_j^{GP})$ 10 Yuan [60] Sun [82]

Pr

$z_i \mid x_i, \tilde{\theta}$ Pr $z_i \mid \tilde{\theta}$

[63-65,73,75,84,86-88]

Dirichlet [60,67-69,78,85,89] Dirichlet

Pitman-Yor [62,70,74,77,79,82-83]

j $P(x_i, z_t=j \mid \tilde{\theta}_j)$

[62-65,82-88] Yuan [60]

MGP

9 Wang Khardon [67]

GP

Kapoor [75] GP

Fox Dunson [90]

GP

MGP

[63-65,67-68,71-72,79,86-88]

[60-62,83,85]

Wishart [60-62,69,82-83,85]

Gamma [60-62,69,74,82-83]

6.2 MGP

MGP

MCMC

VB EM

MGP

MCMC

VB MCMC

[61-62,68-69,77-79,83,85,89]

M-H Gibbs MCMC

[60] MGP

MCMC MGP

[60,71,82] [60]

VB [70,74-75]

[86]

EM

MGP Q

MGP EM

EM

EM EM EM

MCMC-EM

EM E VB

M Q

[60,71,82] [67]

Sun Xu[82] EM

EM M

Tresp[7]

Stachniss Plagemann[91]

Yang Ma[63-64]

Schiegg[76] M Q

EM E

M

Yang Ma[63-64]

EM EM

Nguyen Bo-
nilla[71] E

Yu Chen[84] E

Chen

[65,86] EM

MC-
MC-EM Wu[86] 2015 Chen

[65] E

MCMC

MGP

1

Dong[72] GP MGP

2 Tuong[80]

Liu[81]

GP

GP 3

EP Kapoor[75] MGP

EP

MGP 5

6.3 MGP

MGP

1 ME [66]

[67,74-76,86,91]

2

EM

GP

[65,67,71,84,86,88] Nguyen Bo-
nilla[71,86]

EM

3 MCMC

MCMC

[68-69,83,86]

Meeds Osindero[62]

MCMC

4

Yuan[60] Sun[82] EM

6.4 MGP

MGP

2014 Chen[64] MGP

—— EM Chen

[86] [64] Zhao[88]

[64] Wu[86] E

MCMC

MGP

1

MGP

K

K [7,61,63-65,67-69,71,91] Huang [73]

Akaike

Zhao [88]

K K

K

Dirichlet Pitman-Yor

MGP

Rasmussen [92]

Dirichlet Motorcycle

K

MGP

Shi [93] MCMC

2015 Qiang Ma [89]

K

MGP

MGP

MGP

2

MGP

[60,65]

MGP

3

GP MGP

MGP

MGP

[60,67,82,91] Nguyen Bonil-

la [71] MGP 3

1 5

Ė

GPFR

MGP

7

GP MGP GMM GP

MGP

GMM

[96] GMM

GP GMM MGP

2 IMGP GP

MGP MGP

Dirichlet Pitman-Yor

IGMM GMM

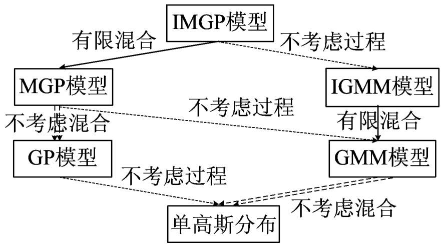
MGP

GP GMM

IMGP IGMM

“ ” IGMM

MGP GMM GP



2 GP MGP GMM

Fig. 2 The hierarchical relationship of GP MGP and GMM models where the arrows denote the direction of degeneration

IMGP

IGMM MGP GMM

1 MGP

GMM GP

IMGP MGP GP

IGMM GMM

GP MGP IMGP

8

GP

GP MGP

1 Williams C K I. Prediction with Gaussian processes From linear regression to linear prediction and beyond M . M. I. Jordan. Learning in Graphical Models. s. l. Netherlands Springer Science & Business Media 1998 599-621.

2 Williams C K I Barber D. Bayesian classification with Gaussian processes J . Pattern Analysis and Machine Intelligence IEEE Transactions on 1998 20 12 1342-1351.

3 Sollich P. Bayesian methods for support vector machines Evidence and predictive class probabilities J . Machine Learning 2002 46 1-3 21-52.

4 Neal R M. Monte Carlo implementation of Gaussian process models for Bayesian regression and classification EB/OL . http //arxiv. org/pdf/physics/9701026v2. pdf 1997/ 2015-05-22.

5 . J . 2005 26 7 96-108.

Zhou Y T Zhang T Y Liu H Y. Kernel-based machine learning method and the applications to multi-user detection a survey J . Journal on Communications 2005 26 7 96-108. in Chinese

6 MacKay D J C. Introduction to Gaussian processes J . NATO ASI Series F Computer and Systems

OL . http //www. gaussianprocess. org 2011-02-23/2015-05-22.

10 Rasmussen C E Williams C K I. Gaussian Processes for Machine Learning M . Cambridge MIT Press 2006 1-248.

11 Seeger M. Gaussian processes for machine learning J . International Journal of Neural Systems 2004 14 2 69-106.

12 . J . 2009 37 6 1153-1158.

Wang X S Zhang Y Y Cheng Y H. Reinforcement Learning for Continuous Spaces Based on Gaussian Process Classifier J . Acta Electronica Sinica 2009 37 6 1153-1158. in Chinese

13 Barber D Williams C K I. Gaussian processes for Bayesian classification via hybrid Monte Carlo J . Advances in Neural Information Processing Systems 1996 9 340-346.

14 Gibbs M N MacKay D J C. Variational Gaussian process classifiers J . IEEE Transactions on Neural Networks 2000 11 6 1458-1464.

15 Opper M Winther O. Gaussian processes for classification Mean-field algorithms J . Neural Computation 2000 12 11 2655-2684.

16 Paciorek C Schervish M. Nonstationary covariance functions for Gaussian process regression J . Advances in Neural Information Processing Systems 2003 16 273-280.

17 Seeger M. Bayesian model selection for support vector machines Gaussian processes and other kernel classifiers C //Proceedings of the 13th Annual Conference on Neural Information Processing Systems. Cambridge MIT Press 2000 EPFL-CONF-161324 603-609.

18 Schwaighofer A Tresp V Yu K. Learning Gaussian process kernels via hierarchical Bayes C //Advances in Neural Information Processing Systems 17. Cambridge MIT Press 2004 1209-1216.

19 Sundararajan S Keerthi S. Predictive approaches for choosing hyperparameters in Gaussian processes J . Neural Computation 2001 13 5 1103-1118.

20 Chatzis S P Demiris Y. Echo state Gaussian process J . Neural Networks IEEE Transactions on 2011 22 9 1435-1445.

21 Soh H Demiris Y. Spatio-Temporal Learning with the Online Finite and Infinite Echo-State Gaussian Processes J . IEEE Trans. Neural Networks and Learning Systems 2015 26 3 522-536.

22 Snelson E Rasmussen C E Ghahramani Z. Warped Gaussian processes J . Advances in Neural Information Processing Systems 2003 16 337-344.

23 Boyle P Frean M. Dependent Gaussian processes J . Advances in Neural Information Processing Systems 2004 17 217-224.

24 Lawrence N D Jordan M I. Semi-supervised learning via Gaussian processes C //Advances in Neural Information Processing Systems 17. Cambridge MIT Press 2004 753-760.

25 . J . 2009 35 7 888-895.

Li H W Liu Y Lu H Q et al. Gaussian Processes Classification Combined with Semi-supervised Kernels J . Acta Automatica Sinica 2009 35 7 888-895. in Chinese

26 Pillonetto G Dinuzzo F De Nicolao G. Bayesian online multitask learning of Gaussian processes J . Pattern Analysis and Machine Intelligence IEEE Transactions on 2010 32 2 193-205.

27 Gilboa E Saatci Y Cunningham J. P. Scaling Multidimensional Inference for Structured Gaussian Processes J . IEEE Trans. on Pattern Analysis and Machine Intelligence 2015 37 2 424-436.

28 Zhao X Cheung L W K. Multiclass Kernel-Imbedded Gaussian Processes for Microarray Data Analysis J . IEEE/ACM Transactions on Computational Biology and Bioinformatics TCBB 2011 8 4 1041-1053.

29 Dallaire P Besse C Chaib-Draa B. An approximate inference with Gaussian process to latent functions from uncertain data J . Neurocomputing 2011 74 11 1945-1955.

30 Williams C K I Vivarelli F. Upper and lower bounds on the learning curve for Gaussian processes J . Machine Learning 2000 40 1 77-102.

31 Vivarelli F Opper M. General bounds on Bayes errors for regression with Gaussian processes J . Advances in Neural Information Processing Systems 1999 11 302-308.

32 Sollich P Halees A. Learning curves for Gaussian process regression Approximations and bounds J . Neural Computation 2002 14 6 1393-1428.

33 Opper M Malzahn D. Learning curves for Gaussian Processes regression A framework for good approximations J . Advances in Neural Information Processing Systems 14 2001 14 273-279.

34 Kakade S Seeger M Foster D. Worst-case bounds for Gaussian process models C //Proc. of the 18th Annual Conference on Neural Information Processing Systems.

Cambridge MIT Press 2005 619-626. EPFL-CONF-161315 .

35 Gibbs M. Bayesian Gaussian processes for classification and regression D . Cambridge University of Cambridge Department of Physics 1997.

36 Shen Y Ng A Seeger M. Fast Gaussian process regression using kd-trees C //Proceedings of the 18th Annual Conference on Neural Information Processing Systems. Cambridge MIT Press 2005 1225-1232.

37 Williams C K I Rasmussen C E Scwaighofer A et al. Observations on the Nyström method for Gaussian process prediction R . London University of Edinburgh and University College London 2002 1-9.

38 Chalupka K Williams C K I Murray I. A framework for evaluating approximation methods for Gaussian process regression J . The Journal of Machine Learning Research 2013 14 1 333-350.

39 Quiñonero-Candela J Rasmussen C E. A unifying view of sparse approximate Gaussian process regression J . The Journal of Machine Learning Research 2005 6 1939-1959.

40 Snelson E Ghahramani Z. Sparse Gaussian Processes using pseudo-inputs C //Advances in Neural Information Processing Systems 18. Cambridge MIT Press 2005 1257-1264.

41 Sollich P. Probabilistic Methods for Support Vector Machines C //Advances in Neural Information Processing Systems 12. Cambridge MIT Press 1999 349-355.

42 Van Gestel T Suykens J A K Lanckriet G et al. Bayesian framework for least-squares support vector machine classifiers Gaussian processes and kernel Fisher discriminant analysis J . Neural Computation 2002 14 5 1115-1147.

43 Rasmussen C E Kuss M. Gaussian processes in reinforcement learning C //Advances in Neural Information Processing Systems 16. Cambridge MIT Press 2003 751-759.

44 Engel Y Mannor S Meir R. Reinforcement learning with Gaussian processes C //Proceedings of the 22nd International Conference on Machine Learning. New York ACM 2005 201-208.

45 Ko J Klein D J Fox D et al. Gaussian processes and reinforcement learning for identification and control of an autonomous blimp C //Robotics and Automation Jinan Shandong China 2007 IEEE International Conference on. IEEE 2007 742-747.

46 Wang J M Fleet D J Hertzmann A. Gaussian process dynamical models for human motion J . Pattern Analysis and Machine Intelligence IEEE Transactions on 2008 30 2 283-298.

47 . J . 2011 16 8 1511-1515.

Lv P Zhang M M Xu M L et al. Rhythmical motion synthesis based on Gaussian process dynamical model J . Journal of Image and Graphics 2011 16 8 1511-1515. in Chinese

48 Deisenroth M P Rasmussen C E Peters J. Gaussian process dynamic programming J . Neurocomputing 2009 72 7 1508-1524.

49 Amoto C Chowdhary G Liu M et al. Off-policy reinforcement learning with Gaussian processes J . IEEE/CAA Journal of Automatica Sinica 2014 1 3 227-238.

50 Gao X Wang X Tao D et al. Supervised Gaussian process latent variable model for dimensionality reduction J . Systems Man and Cybernetics Part B Cybernetics IEEE Transactions on 2011 41 2 425-434.

51 Lawrence N. Probabilistic non-linear principal component analysis with Gaussian process latent variable models J . The Journal of Machine Learning Research 2005 6 1783-1816.

52 Lawrence N D Moore A J. Hierarchical Gaussian process latent variable models C //Proceedings of the 24th International Conference on Machine Learning. Corvallis OR USA ACM 2007 481-488.

53 Urtasun R Darrell T. Discriminative Gaussian process latent variable model for classification C //Proceedings of the 24th International Conference on Machine Learning. Corvallis OR USA ACM 2007 927-934.

54 Cheng F Yu J Xiong H. Facial expression recognition in JAFFE dataset based on Gaussian process classification J . Neural Networks IEEE Transactions on 2010 21 10 1685-1690.

55 Brahim-Belhouari S Bermak A. Gaussian process for nonstationary time series prediction J . Computational Statistics & Data Analysis 2004 47 4 705-712.

56 Sun S Zhong P Xiao H et al. Active Learning With Gaussian Process Classifier for Hyperspectral Image Classification J . IEEE Transactions on Geoscience and Remote Sensing 2015 53 4 1746-1760.

57 Jacobs J. P Koziel S. Two-Stage Framework for Efficient Gaussian Process Modeling of Antenna Input Characteristics J . IEEE Transactions on Antennas and Propagation 2014 62 2 706-713.

58 Chen N Qian Z Nabney I T et al. Wind Power Forecasts Using Gaussian Processes and Numerical Weather Prediction J . IEEE Transactions on Power System 2014 29 2 656-665.
59 Markov K Matsui T. Music Genre and Emotion Recognition Using Gaussian Processes J . IEEE Access 2014 2 688-697.
60 Yuan C Neubauer C. Variational mixture of Gaussian process experts C //Advances in Neural Information Processing Systems 21. Cambridge MIT Press 2008 1897-1904.
61 Gramacy R B Lee H K H. Bayesian treed Gaussian process models with an application to computer modeling J . Journal of the American Statistical Association 2008 103 483 1119-1130.
62 Meeds E Osindero S. An alternative infinite mixture of Gaussian process experts C //Advances in Neural Information Processing Systems 18. Cambridge MIT Press 2005 883-890.
63 Yang Y Ma J. An efficient EM approach to parameter learning of the mixture of gaussian processes C //Advances in Neural Networks-ISNN 2011. Berlin Heidelberg Springer 2011 165-174.
64 . EM D . 2011.
Yang Y. Study of the EM algorithms for the Mixture of Experts Architecture D . Beijing Peking University. School of Mathematical Sciences 2011. in Chinese
65 Chen Z Ma J Zhou Y. A Precise Hard-Cut EM Algorithm for Mixtures of Gaussian Processes C //Intelligent Computing Methodologies. Switzerland Springer International Publishing 2014 68-75.
66 Yuksel S E Wilson J N Gader P D. Twenty years of mixture of experts J . Neural Networks and Learning Systems IEEE Transactions on 2012 23 8 1177-1193.
67 Wang Y Khardon R. Sparse Gaussian Processes for multitask learning C //Machine Learning and Knowledge Discovery in Databases. Berlin Heidelberg Springer 2012 711-727.
68 Shi J Q Murray-Smith R Titterington D M. Bayesian regression and classification using mixtures of Gaussian processes J . International Journal of Adaptive Control and Signal Processing 2003 17 2 149-161.
69 Shi J Q Murray-Smith R Titterington D M. Hierarchical Gaussian process mixtures for regression J . Statistics and Computing 2005 15 1 31-41.
70 Ross J Dy J. Nonparametric mixture of Gaussian processes with constraints C //Proceedings of the 30th International Conference on Machine Learning ICML-13 . s. l. s. n. 2013 1346-1354.
71 Nguyen T Bonilla E. Fast allocation of Gaussian process experts C //Proceedings of the 31st International Conference on Machine Learning ICML-14 . s. l. s. n. 2014 145-153.
72 Lu Z. The Laplace Approximation of Gaussian Process Mixture EB/OL . http //snowbird. djvuzone. org/2007/abstracts/144. pdf 2007/2015-05-22.
73 Huang M Li R Wang H et al. Estimating Mixture of Gaussian Processes by Kernel Smoothing J . Journal of Business & Economic Statistics 2014 32 2 259-270.
74 Platanios E A Chatzis S P. Mixture Gaussian Process Conditional Heteroscedasticity J . IEEE Transactions on Pattern Analysis and Machine Intelligence 2014 36 5 888-900.
75 Kapoor A Ahn H Picard R W. Mixture of Gaussian processes for combining multiple modalities C //Multiple Classifier Systems. Berlin Heidelberg Springer 2005 86-96.
76 Schiegg M Neumann M Kersting K. Markov Logic Mixtures of Gaussian Processes Towards Machines Reading Regression Data C //Proceedings of the 15th International Conference on Artificial Intelligence and Statistics. s. l. s. n. 2012 1002-1011.
77 Wei H Lu W Zhu P et al. Camera control for learning nonlinear target dynamics via Bayesian nonparametric Dirichlet-process Gaussian-process DP-GP models C //Intelligent Robots and Systems IROS 2014 . Chicago IL 2014 IEEE/RSJ International Conference on. IEEE 2014 95-102.
78 Hernández S Sallis P. Distributed Minimum Temperature Prediction Using Mixtures of Gaussian Processes C //Environmental Software Systems. Infrastructures Services and Applications. s. l. Springer International Publishing 2015 484-491.
79 Ouyang R Low K H Chen J et al. Multi-robot active sensing of non-stationary Gaussian process-based environmental phenomena C //Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems. s. l. International Foundation for Autonomous Agents and Multi-agent Systems 2014 573-580.
80 Nguyen-Tuong D Seeger M Peters J. Model learning with local Gaussian Process regression J . Advanced Robotics 2009 23 15 2015-2034.
81 Liu Z Zhou L Leung H et al. Kinect Posture Recon-

struction based on a Local Mixture of Gaussian Process Models J . IEEE Transactions on Visualization and Computer Graphics 2015 PP 99 .

82 Sun S Xu X. Variational inference for infinite mixtures of Gaussian processes with applications to traffic flow prediction J . IEEE Transactions on Intelligent Transportation Systems 2011 12 2 466-475.

83 Sun S. Infinite mixtures of multivariate Gaussian processes C //International Conference on Machine Learning and Cybernetics. Tianjin 2013 IEEE International Conference on. IEEE 2013 1011-1016.

84 Yu J Chen K Rashid M M. A Bayesian model averaging based multi-kernel Gaussian process regression framework for nonlinear state estimation and quality prediction of multiphase batch processes with transient dynamics and uncertainty J . Chemical Engineering Science 2013 93 19 96-109.

85 Ohishi Y Mochihashi D Kameoka H et al. Mixture of Gaussian process experts for predicting sung melodic contour with expressive dynamic fluctuations C //Acoustics Speech and Signal Processing ICASSP . Florence Italy 2014 IEEE International Conference on. IEEE 2014 3714-3718.

86 Wu D Chen Z Ma J. An MCMC based EM algorithm for mixtures of Gaussian processes C //Advances in Neural Networks-ISNN 2015. Berlin Heidelberg Springer 2015 327-334.

87 Chen Z Ma J. The Hard-Cut EM Algorithm for Mixture of Sparse Gaussian Processes C //Intelligent Computing Methodologies. Switzerland Springer International Publishing 2015 13-24.

88 Zhao L Chen Z Ma J. An Effective Model Selection Criterion for Mixtures of Gaussian Processes C //Advances in Neural Networks-ISNN 2015. Berlin Heidelberg Springer 2015 345-354.

89 Qiang Z Ma J. Automatic Model Selection of the Mixtures of Gaussian Processes for Regression C //Advances in Neural Networks-ISNN 2015. Berlin Heidelberg Springer 2015 335-344.

90 Fox E B Dunson D B. Multiresolution Gaussian Processes C //Advances in Neural Information Processing Systems 25. Cambridge MIT Press 2012 737-745.

91 Stachniss C Plagemann C Lilienthal A J et al. Gas Distribution Modeling using Sparse Gaussian Process Mixture Models C //Robotics Science and Systems. Cambridge MIT Press 2008 310-317.

92 Rasmussen C E Ghahramani Z. Infinite mixtures of Gaussian process experts C //Advances in Neural Information Processing Systems 14. Cambridge MIT Press 2001 881-888.

93 Shi J Q Wang B. Curve prediction and clustering with mixtures of Gaussian process functional regression models J . Statistics and Computing 2008 18 3 267-283.

94 Shi J Q Wang B Murray-Smith R et al. Gaussian process functional regression modeling for batch data J . Biometrics 2007 63 3 714-723.

95 Shi J Q Wang B Will E J et al. Mixed-effects Gaussian process functional regression models with application to dose-response curve prediction J . Statistics in medicine 2012 31 26 3165-3177.

96 Ma J Liu J. The BYY annealing learning algorithm for Gaussian mixture with automated model selection J . Pattern Recognition 2007 40 7 2029-2037.

1973

2013 9 ~2014 6

E-mail zyt@hebut. edu. cn

1990 2015

E-mail kazy90@126. com

1962 1992

E-mail jwma@math. pku. edu. cn