

## ノンパラメトリックベイズ法に基づく時系列データの分節化

## Segmenting Time Series Data Using GP-HSMM with Nonparametric Bayesian Model

長野匡隼 \*1  
Masatoshi Nagano中村友昭 \*1  
Tomoaki Nakamura長井隆行 \*1  
Takayuki Nagai持橋大地 \*2  
Daichi Mochihashi小林一郎 \*3  
Ichiro Kobayashi金子正秀 \*1  
Masahide Kaneko

\*1電気通信大学

The University of Electro-Communications

\*2統計数理研究所

Institute of Statistical Mathematics

\*3お茶の水女子大学

Ochanomizu University

In this paper, we propose a method for dividing continuous time-series data into segments in unsupervised manner. Humans recognize perceived continuous information by dividing it into significant segments such as words and unit motions. To this end, we have been proposed a method based on hidden semi-Markov model with Gaussian process (GP-HSMM). However, it has a big drawback that it requires the number of classes into which time-series data is segmented. To overcome this problem, in this paper, we extend GP-HSMM to nonparametric Bayesian model by introducing hierarchical Dirichlet processes (HDP), and propose hierarchical Dirichlet processes-Gaussian process-hidden semi-Markov model (HDP-GP-HSMM). Hence, the infinite number of classes is assumed and the number of classes are estimated by applying slice sampling. In the experiment, we used the various time-series data and showed that our proposed model can estimate more correct number of classes and achieve more accurate segmentation than baseline methods.

## 1. はじめに

人は知覚した連続的な情報を意味をもつ単位に分節化することで認知している。分節化を行うことで、音声波形から意味を持つ単語を認識することができ、人の動作から単位動作を認識することができる。人は、分節点や単位動作の組み合わせに無数の可能性がある中から、明示的な分節点を与えられことなく自ら学習し、情報を適切に分節化している。ロボットにおいても、言語や動作等を柔軟に学習するために、教師なしで分節化する能力は非常に重要であると言える。

そこで、我々は隠れセミマルコフモデルにおける出力をガウス過程とした Gaussian Process - Hidden Semi-Markov model (GP-HSMM)[Nakamura 17]を用いることで、時系列データの分節化を可能とした。しかし、GP-HSMMでは、事前に分類する単位動作のクラス数を与える必要があった。本稿では、この問題を解決するため GP-HSMM に Hierarchical Dirichlet Process (HDP)[Teh 06]を導入し、ノンパラメトリックベイズモデルへと拡張した HDP-GP-HSMM を提案する。ノンパラメトリックベイズモデルでは、無限個のクラスがあることを仮定しクラスタリングを行う。実際には学習データは有限であり、そのうちの有限個のクラスのみが学習データを表現するのに利用されることで、クラス数の推定を可能としている。すなわち、HDP-GP-HSMM のパラメータを学習することで、教師なしで分節点と各単位系列のクラスを推定するだけでなく、クラス数も決定することができる。しかし、無限個のクラスを仮定しているため、GP-HSMM の学習に用いられている Forward filtering-Backward sampling[Uchiumi 15]をそのまま用いることはできない。そのために、文献[Gael 08]で提案された Slice-sampling を用いることにより、無限個のクラスを確率的に有限個にすることで学習を可能とする。

このような教師無しで動作を分節化する研究はいくつか行われているが、ヒューリスティックな仮定をおいているものが多い。しかしながら、ヒューリスティックな仮定が全ての動作に対して有効に働くかは定かではない。また単位動作の表現に Hidden Markov Model (HMM) を用いることで、教師なしで分節化する手法が提案されている [Beal 01, Taniguchi 11, Fox 11, Matsubara 14]。HMM は、構造がシンプルであり、複雑な動作を表現することが難しいと考えられる。一方、我々は Gaussian Process (GP) により動作を表現する。GP はノンパラメトリックな手法であり、複雑な動作を表現するのに適して

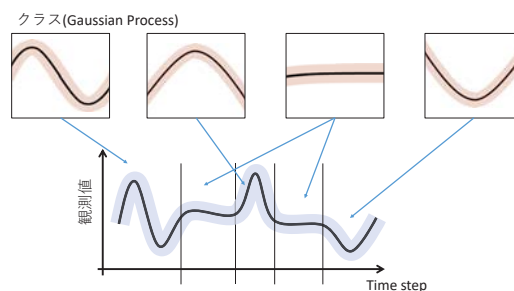


図 1: ガウス過程による単位系列表現

いると考えている。

本稿では、人工データとモーションキャプチャデータを用いた実験により、他の教師なしの分節化手法よりもクラス数推定及び分節化精度が高いことを示す。

## 2. ノンパラメトリックベイズ

本稿では、ノンパラメトリックベイズモデルを用いることで単位動作のクラス数を推定する。パラメトリックな混合モデルではクラス数は有限であり、有限次元のディリクレ分布から多項分布のパラメータが生成され、この多項分布から隠れ状態が決定される。一方、ノンパラメトリックベイズモデルでは、無限次元の多項分布を用いることで無限個のクラスを仮定する。しかし、実際にはその中の有限個のクラスのみが使用されるため、結果としてクラス数の推定が可能となる。無限次元の多項分布の構成法の 1 つが Stick Breaking Process (SBP) である。

## 2.1 Stick Breaking Process

Stick Breaking Process (SBP)[Gael 08] は、無限次元の多項分布を生成可能なディリクレ過程である。SBP では、図 2 のように長さ 1 の棒をベータ分布からサンプリングされる値  $v_k$  により、折り続けることで無限次元の多項分布  $\beta$  を構成している。

$$v_k \sim \text{Beta}(1, \gamma) \quad (k = 1, \dots, \infty) \quad (1)$$

$$\beta_k = v_k \prod_{i=1}^{k-1} (1 - v_i) \quad (k = 1, \dots, \infty) \quad (2)$$

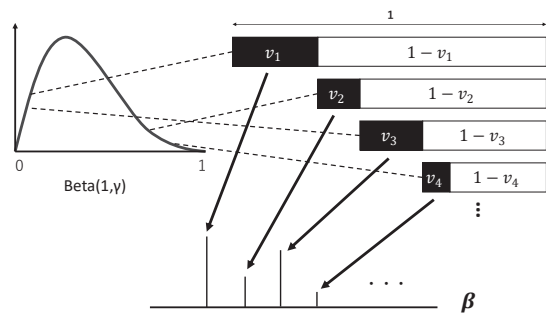


図 2: Stick Breaking Process

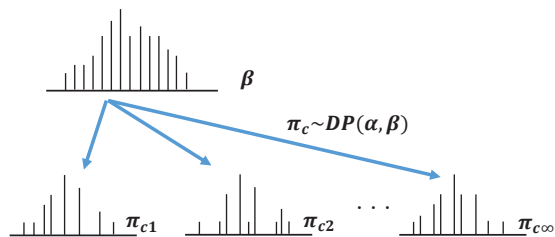


図 3: Hierarchical Dirichlet Process

このプロセスは一般的に  $\beta \sim \text{GEM}(\gamma)$  と表記される [Pitman 02].

## 2.2 Hierarchical Dirichlet Process

ディリクレ過程を用いることで、無限の状態数をもった HMM の遷移確率を構成することができる。しかし、単に SBP によって各状態の遷移確率を構成するだけでは、状態毎に可能な遷移先が異なることになる。HMM のようなモデルの場合、全ての状態において可能な遷移先が共有されており、その遷移確率が状態毎に異なる分布を構成しなければならない。そのために、図 3 で示すように、SBP によって生成した分布  $\beta$  を基底測度とし共有し、状態  $c$  毎に異なる遷移確率  $\pi_c$  を DP によって生成する。

$$\pi_c \sim \text{DP}(\alpha, \beta) \quad (3)$$

このような、二段階の DP によって無限次元の確率分布を構成する手法は、Hierarchical Dirichlet Process (HDP) と呼ばれている。

## 3. HDP-GP-HSMM

図 4 が提案手法のグラフィカルモデルである。  $c_j (j = 1, 2, \dots, J)$  が単位系列のクラスを表しており、無限個のクラス数を仮定している。  $\pi_c$  は、クラス  $c$  から各クラスへの遷移確率を表しており、  $\gamma$  をパラメータとする GEM 分布 (Stick Breaking Process) によって生成された  $\beta$  と、  $\alpha$  をパラメータとする Dirichlet Process (DP) から生成される。

$$\beta \sim \text{GEM}(\gamma) \quad (4)$$

$$\pi_c \sim \text{DP}(\alpha, \beta) \quad (5)$$

$j$  番目のクラス  $c_j$  は、  $j-1$  番目のクラス  $c_{j-1}$  と遷移確率  $\pi_c$  によって決定される。ここでは、クラスと対応した  $\mathbf{X}_c$  をパラメータとするガウス過程から単位系列  $\mathbf{x}_j$  が生成されることを仮定している。

$$c_j \sim P(c|c_{j-1}, \beta, \alpha) \quad (6)$$

$$\mathbf{x}_j \sim \mathcal{GP}(\mathbf{x}|\mathbf{X}_{c_j}) \quad (7)$$

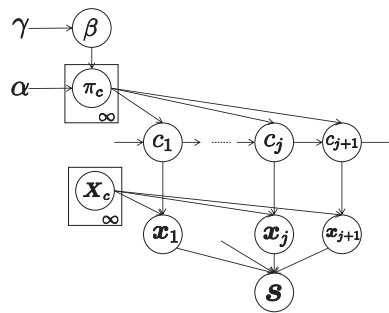


図 4: HDP-GP-HSMM のグラフィカルモデル

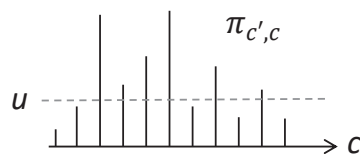


図 5: Slice sampling

ただし、  $\mathbf{X}_c$  はクラス  $c$  に分類された単位系列の集合である。これらの単位系列を結合することで、観測系列  $\mathbf{S} = \mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_J$  となる。

## 4. 学習

提案モデルでは、観測系列内の単位系列とそのクラスをサンプリングすることで推定する。効率的に計算するために、1つの観測系列内の単位系列とそのクラスを一括でサンプリングする Blocked Gibbs Sampler を用いる。まず、全ての観測系列をランダムに分節・分類することで初期化する。次に、ある観測系列  $\mathbf{s}_n$  を分節して得られた単位系列  $\mathbf{x}_{nj} (j = 1, 2, \dots, J_n)$  を、それらが分類されているクラスからすべて削除し、各クラスのガウス過程のパラメータ  $\mathbf{X}_c$  と遷移確率  $P(c|c')$  を更新する。次に、単位動作の長さとする組み合わせを効率的に計算可能な Forward filtering-Backward sampling を用い、単位系列  $\mathbf{x}_{nj} (j = 1, 2, \dots, J_n)$  とそれらのクラス  $c_{nj} (j = 1, 2, \dots, J_n)$  を推定し、各クラスと対応したガウス過程のパラメータ  $\mathbf{X}_c$  と遷移確率  $P(c|c')$  を更新する。以上の操作を全ての系列に対して繰り返すことで、観測系列の分節点とガウス過程のパラメータを交互に最適化することができる。しかし、このままでは生成された遷移確率  $\pi_c$  の遷移先のクラス数は無限個存在するため、単位動作が各クラスに分類される確率を計算することができない。本稿では、この問題を解決するために、Slice Sampling を用いる。Slice Sampling では、パラメータ  $u$  を閾値として遷移先を確率的に切り捨てることで、遷移確率  $\pi_c$  の遷移先のクラス数を有限にする。パラメータ  $u$  は、タイムステップ  $j$  に対して、遷移確率  $\pi_c$  から次式のように計算する。  $\pi_{c',c}$  はクラス  $c'$  から  $c$  への遷移確率である。

$$p(u_j|c_{j-1}, c_j, \pi_c) = \frac{\delta(0 < u_j < \pi_{c_{j-1}, c_j})}{\pi_{c_{j-1}, c_j}} \quad (8)$$

図 5 のようにパラメータ  $u$  を用いて、  $\pi_{c',c} \geq u_j$  となるクラスのみ遷移可能とすることで、使用されるクラスを有限個にする。その結果、forward filtering-backward sampling を利用することが可能となる。

## 5. 実験

提案手法の有効性を検証するために、複数のデータに対して分節化を行い、HDP-HMM [Beal 01], HDP-

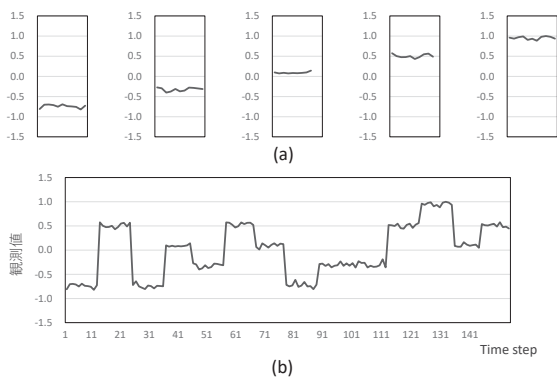


図 6: 人工データ 1: (a) 5つの単位動作, (b) 観測系列.

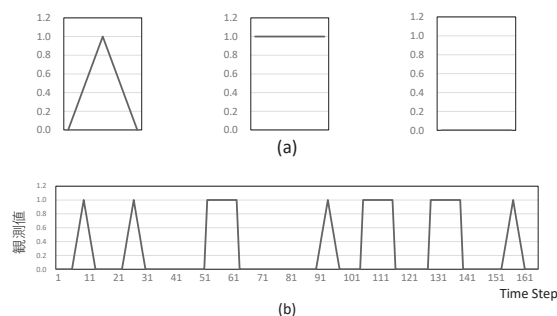


図 7: 人工データ 2: (a) 5つの単位動作, (b) 観測系列.

HMM+NPYLM[Taniguchi 11], BP-HMM[Fox 11], Autoplait[Matsubara 14] と比較した. 評価方法は文献[Nakamura 17]と同様に, 正規化ハミング距離及び適合率, 再現率, F 値を用いた. 分節点の誤差の許容範囲は観測系列の全フレーム数の  $\pm 5\%$  とした.

### 5.1 人工データ

以下の二つの人工データに対して分節化を行った.

- **人工データ 1:** まず, 5つの値からランダムに選択し, 一様乱数により系列長をランダムに選択する.

$$k \sim \text{uniform}(9, 13). \quad (9)$$

さらに, 正規乱数によりノイズを加える.

$$x \sim x + \text{normal}(x, 0.05). \quad (10)$$

図 6 のように, これらの単位系列をランダムに結合して観測系列を作成した.

- **人工データ 2:** 図 7(a) の単位系列をランダムに結合することで, 図 7(b) に示す観測系列を作成した.

それぞれ分節化の結果を表 1 と表 2 にまとめた. 表 1 と表 2 より, 提案手法を用いて分節化を行った結果, ハミング距離は 0.1 未満かつ適合率, 再現率, F 値は 1.0 となった. この結果より, HDP-GP-HSMM では正しく分節化がされていることが分かる. さらに, クラス数も人工データ 1 では 5, 人工データ 2 では 3 と正しく推定された.

また, 人工データ 1 は HDP-HMM においても正しく分節化されていることが分かる. これは, 人工データ 1 が, ガウス分布に従う単純なデータであったためだと考えられる. そのため, 人工データ 2 はガウス分布に従わないデータであるために正しく分節化ができなかったと考えられる. 一方, 人工データ 2 のような複雑なデータは, HDP-GP-HSMM や BP-HMM のような複雑なモデルでは正しく分節化されていることが分かる.

表 1: 人工データ 1 の分節化結果

	ハミング距離	適合率	再現率	F 値	推定されたクラス数
HDP-GP-HSMM	0.016	1.0	1.0	1.0	5
HDP-HMM	0.011	1.0	1.0	1.0	5
HDP-HMM+NPYLM	0.48	0.53	1.0	0.69	35
BP-HMM	0.36	1.0	0.94	0.97	4
Autoplait	0.79	0	0	0	3

表 2: 人工データ 2 の分節化結果

	ハミング距離	適合率	再現率	F 値	推定されたクラス数
HDP-GP-HSMM	0.037	1.0	1.0	1.0	3
HDP-HMM	0.13	0.43	1.0	0.60	8
HDP-HMM+NPYLM	0.45	0.57	1.0	0.72	19
BP-HMM	0.038	1.0	1.0	1.0	3
Autoplait	0.35	0.17	0.15	0.16	1

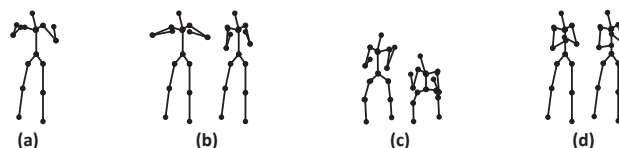


図 8: チキンダンスに含まれる単位動作: (a) beaks, (b) wings, (c) tail feathers, and (d) claps.

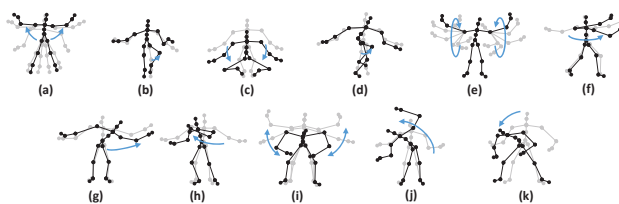


図 9: 体操に含まれる単位動作: (a) jumping jack, (b) jogging, (c) squatting, (d) knee raise, (e) arm circle, (f) twist, (g) side reach, (h) boxing, (i) arm wave, (j) side bend, and (k) toe touch.

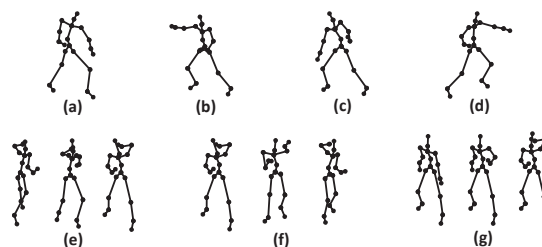


図 10: 空手の型に含まれる単位動作: (a) 下段払い (左), (b) 追い突き (右), (c) 下段払い (右), (d) 追い突き (左), (e) 上げ受け (左), (f) 上げ受け (右), (g) 添え手.

### 5.2 モーションキャプチャデータ

以下の三つのモーションキャプチャデータに対して分節化を行った.

- **チキンダンス:** CMU のモーションキャプチャデータに含まれるチキンダンスのモーションキャプチャデータを使用した\*1. 図 8 のようにチキンダンスには 4 個の単位動作が含まれている. 分節化には, 文献 [Matsubara 14] で用いられている両手両足の加速度を用いた.

\*1 <http://mocap.cs.cmu.edu/>

表 3: チキンダンスの分節化結果

	ハミング距離	適合率	再現率	F 値	推定された クラス数
HDP-GP-HSMM	0.038	1.0	1.0	1.0	4
HDP-HMM	0.46	0.39	0.71	0.50	4
HDP-HMM+NPYLM	0.65	0.63	0.71	0.67	8
BP-HMM	0.16	0.70	1.0	0.82	4
Autoplait	0.026	1.0	1.0	1.0	4

表 4: 体操の分節化結果

	ハミング距離	適合率	再現率	F 値	推定された クラス数
HDP-GP-HSMM	0.29	0.44	0.93	0.60	10
HDP-HMM	0.80	0.02	0.89	0.04	17
HDP-HMM+NPYLM	0.64	0.28	1.0	0.44	23
BP-HMM	0.23	0.33	0.95	0.50	13
Autoplait	0.89	0.67	0.11	0.20	3

- **体操:** CMU のモーションキャプチャデータの subject14 から体操のモーションキャプチャデータを用いた。図 9 のように体操のデータには 11 個の単位動作が含まれている。分節化には、計測した動作の両手両足の奥行きを除いた左右 ( $x$ ) と上下 ( $y$ ) の二次元の座標を用いた。
- **空手の型:** より複雑な動作として、空手のモーションキャプチャデータを用いた\*2。図 10 のように空手の型には 7 個の単位動作が含まれている。分節化には、計測した動作の両手の奥行きを除いた左右 ( $x$ ) と上下 ( $y$ ) の二次元の座標を用いた。

これらのデータに対する前処理としてダウンサンプリングと正規化を行った。HDP-GP-HSMM では、毎秒 4 フレームにダウンサンプリングを行った。また、min-max 法を用いて最小値を  $-1$ 、最大値を  $1$  にするように正規化を行った。一方、他の分節化の手法においては、毎秒 8 フレーム、4 フレーム、2 フレームでダウンサンプリングを行い、min-max 法と z-score を用いて正規化を行い精度の良いものを選択した。z-score は、平均値及び標準偏差を  $0$  及び  $1$  とする正規化である。

表 3, 4, 5 は各データの分節化の結果を表す。これらの表より、提案手法では正しく分節化がされていることが分かる。また、チキンダンスのデータは BP-HMM と autoplait でも高精度で正しく分節化出来た。これは、チキンダンスのデータが単位動作を繰り返すだけの最も単純なデータであるからだと考えられる。

さらに、提案手法ではチキンダンスと空手の型のデータのクラス数を正しく推定できた。一方、体操のデータにおいて提案手法で分節化を行った結果、クラス数は正解のクラス数 11 個より少ない 10 個と推定された。これは、図 9 の動作 F と動作 H が似ている動作であるので、同じクラスに分類されたためだと考えられる。しかし、他の手法に比べ提案手法が最も正解のクラス数に近いクラス数を推定できたことが分かる。

以上の分節化及びクラス数推定の結果から、HDP-GP-HSMM が様々なデータに対して有効であると考えられる。

## 6. まとめ

本稿では、ガウス過程を出力分布とした隠れセミマルコフモデルにディリクレ過程を導入することで、動作のクラス数を推定しつつ動作を分節・分類する手法を提案した。提案手法では、動作の切れ目と、動作のクラスを Forward filtering-Backward sampling により同時に推定し、クラス数をノンパラメトリックベイズモデルにより推定することで動作の分節化を行った。実験により、動作のクラス数の推定をしつつ、分節・分類が可能であることが示された。また、様々な連続的な時系列データに関しても分節化が可能であることが示された。

しかし、ガウス過程を用いているため、他の手法と比較すると計算時間が長いといった問題がある。この問題のため、HDP-GP-HSMM は大規模データに用いることができない。計算時

\*2 <http://www.mocapdata.com/>

表 5: 空手の型の分節化結果

	ハミング距離	適合率	再現率	F 値	推定された クラス数
HDP-GP-HSMM	0.21	0.70	0.90	0.80	7
HDP-HMM	0.43	0.47	0.71	0.57	9
HDP-HMM+NPYLM	0.57	0.28	0.67	0.40	20
BP-HMM	0.32	0.40	1.0	0.57	12
Autoplait	0.60	0	0	0	3

間の削減については、文献 [Nguyen-Tuong 09, Okadome 14] で提案されているガウス過程の近似的手法を用いることができると考えている。さらに、今回は特徴量を両手両足の  $xy$  座標のみ用いたが、今後全身の関節位置を用いた、より複雑な動作を扱うことも考えている。

## 参考文献

- [Nakamura 17] T. Nakamura, T. Nagai, D. Mochihashi, I. Kobayashi, H. Asoh, and M. Kaneko, "Segmenting Continuous Motions with Hidden Semi-Markov Models and Gaussian Processes," *Frontiers in Neurorobotics*, vol.11, no.67, 2017.
- [Teh 06] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet processes," *Journal of the American Statistical Association*, vol.101, no.476, pp.1566–1581, 2006.
- [Uchiumi 15] K. Uchiumi, T. Hiroshi, and D. Mochihashi, "Inducing Word and Part-of-Speech with Pitman-Yor Hidden Semi-Markov Models," *Joint Conference of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, pp.1774–1782, 2015.
- [Gael 08] J. V. Gael, Y. Saatchi, Y. W. Teh, and Z. Ghahremani, "Beam Sampling for the Infinite Hidden Markov Model," *In Proceedings of the 25th International Conference on Machine Learning*, pp.1088–1095, 2008.
- [Beal 01] M. J. Beal, Z. Ghahremani, and C. E. Rasmussen, "The infinite hidden Markov model," *In Advances in Neural Information Processing Systems*, pp.577–584, 2001.
- [Taniguchi 11] Taniguchi. T, and Nagasaka. S, "Double articulation analyzer for unsegmented human motion using Pitman-Yor language model and infinite hidden Markov model," *In IEEE/SICE International Symposium on System Integration*, pp.250–255, 2011.
- [Fox 11] E. B. Fox, E. B. Sudderth, M. I. Jordan, and A. S. Willsky, "Joint modeling of multiple related time series via the beta process," *arXiv preprint arXiv:1111.4226*, 2011.
- [Matsubara 14] Y. Matsubara, Y. Sakurai, and C. Faloutsos, "Autoplait: Automatic mining of co-evolving time sequences," *ACM SIGMOD International Conference on Management of Data*, pp.193–204, 2014.
- [Pitman 02] J. Pitman, "Poisson-Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition," *Combinatorics, Probability and Computing*, vol.11, pp.501–514, 2002. MR1930355
- [Nguyen-Tuong 09] D. Nguyen-Tuong, J. R. Peters, and M. Seeger, "Local Gaussian process regression for real time online model learning and control," *In Advances in Neural Information Processing Systems*, pp.1193–1200, 2009.
- [Okadome 14] Y. Okadome, K. Urai, Y. Nakamura, T. Yomo, and H. Ishiguro, "Adaptive lsh based on the particle swarm method with the attractor selection model for fast approximation of Gaussian process regression," *Artificial Life and Robotics*, vol.19, pp.220–226, 2014.