

可変長データに対する最大ベイズ境界性学習法の実験的評価

岡内 亮太[†] 片桐 滋[†] 大崎 美穂[†]

[†] 同志社大学

1 はじめに

パターン認識における究極の目標は、最小分類誤り確率（ベイズ誤り）状態に対応する分類境界（ベイズ境界）を構成する分類器パラメータを求めることである。しかし、ベイズ誤りの導出には無限個の標本が必要であるが、実用的な標本数は有限個であるため、その値は推定せざるを得ない。ベイズ誤りの推定手法として、近年、ベイズ境界性に基づく損失の最小化を通してベイズ境界の達成を目指した、最大ベイズ境界性 (MBB: Maximum Bayes Boundary-ness) 学習法¹⁾が提案された。先行研究では、固定次元パターンを対象とした実験を通して、ベイズ誤り状態を達成し得る可能性が示唆されているが、音声などの可変長パターンに対する有用性は未知数である。本稿では、MBB 学習法を音声認識に適用させることで、可変長パターンに対する MBB 学習法の有用性を検証する。

2 最大ベイズ境界性学習法

2.1 学習目標

MBB 学習法は、分類境界のベイズ境界性を最大化することを目的としているため、分類境界のベイズ境界らしさを定義する必要がある。そこで、ベイズ境界性に対応する尺度であるベイズ境界性尺度を定義する。推定された分類境界の評価基準であるベイズ境界性尺度は、以下のシャノンエントロピー関数の式で表される。

$$H_y(\mathbf{X}) = \sum_{j=1}^2 P(C_j|\mathbf{X}) \log_2 P(C_j|\mathbf{X}) \quad (1)$$

式 (1) で表されるベイズ境界性尺度は、入力である分類事後確率が等しい場合、最大値 1 となり、片方のクラスに偏る程、最小値 0 に近づいていく。

2.2 分類器の構成

可変次元である音声パターンを分類対象とするため、クラスモデルとしてマルチプロトタイプ-状態遷移モデル (MP-STM: Multi Prototype-State Transition Model)²⁾を用いる。各音素は複数の状態を持ち、各状態は複数のプロトタイプで表現される (図 1)。

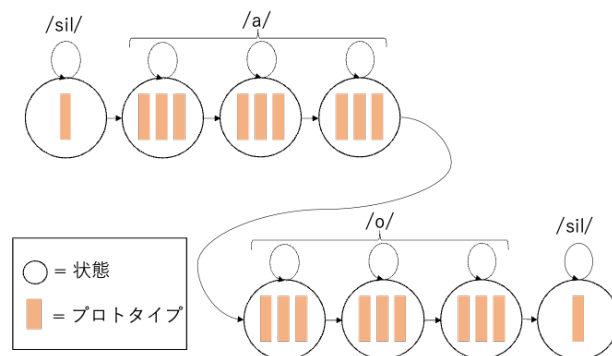


図 1 単語"a o"の MP-STM

2.3 識別関数

入力特徴系列 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$ とクラス C_j の識別関数として、次式で表される動的時間伸縮 (DTW: Dynamic Time Warping) を用いる。

$$g_j(\mathbf{X}; \mathbf{\Lambda}) = -\frac{1}{T} \sum_{t=1}^T \left\| \mathbf{x}_t - \mathbf{r}_{i(\phi_{j,t}, \theta_{j,t}, t)}^{\phi_{j,t}, \theta_{j,t}} \right\|^2 \quad (2)$$

$$i(\phi_{j,t}, \theta_{j,t}, t) = \arg \min_{i=1}^{I_{\phi_{j,t}, \theta_{j,t}, t}} \left\| \mathbf{x}_t - \mathbf{r}_i^{\phi_{j,t}, \theta_{j,t}} \right\|^2 \quad (3)$$

ここで、 t は入力特徴系列の時間インデックス ($t = 1, \dots, T$), $\phi_{j,t}$ および $\theta_{j,t}$ は最小累積距離の経路に沿って選ばれる C_j の時刻 t における音素および状態のインデックス, $i(\phi_{j,t}, \theta_{j,t}, t)$ は時刻 t における入力特徴ベクトルに最も近いプロトタイプのインデックスであり、 $\phi_{j,t}$ 番目の音素の $\theta_{j,t}$ 番目の状態内から選ばれる。この時、入力特徴系列は識別関数値が最も大きいクラス C_j に分類される。

2.4 損失関数

推定された分類境界をベイズ境界に近付けるように分類器パラメータを更新するため、損失は式 (1) を用いて、以下の式で定義する。

$$\hat{U}_y(\mathbf{X}; \mathbf{\Lambda}) = 1 - \hat{H}_y(\mathbf{X}; \mathbf{\Lambda}) \quad (4)$$

2 クラスの場合におけるシャノンエントロピー関数の最大値である 1 と、推定されたベイズ境界性尺度との差を損失と定義することで、損失の最小化がベイズ境界性尺度の最大化につながる。

式 (4) は各標本に対する損失であるため、全入力標本に対する損失を経験的平均損失として、次式で表す。

$$L(\mathbf{\Lambda}) = \frac{1}{|S_{NB}|} \sum_{\mathbf{X} \in S_{NB}} \hat{U}_y(\mathbf{X}; \mathbf{\Lambda}) \quad (5)$$

Experimental Evaluation of Maximum Bayes Boundary-ness Training Method with Variable Length Data
Ryota Okauchi[†] Shigeru Katagiri[†] Miho Ohsaki[†]
[†]Doshisha University

ただし, S_{NB} は境界近傍標本集合を表す.

2.5 分類器パラメータの更新

式 (5) を最小化する最適化手法として, 本研究では最急降下法を採用し, 次のようにパラメータ Λ を更新する.

$$\Lambda^{(t+1)} = \Lambda^{(t)} - \varepsilon \frac{\partial L(\Lambda^{(t)})}{\partial \Lambda^{(t)}} \quad (6)$$

$$\frac{\partial L(\Lambda)}{\partial \Lambda} = \frac{1}{|S_{NB}|} \sum_{\mathbf{x} \in S_{NB}} \frac{\partial \hat{U}_y(\mathbf{X}; \Lambda)}{\partial \Lambda} \quad (7)$$

ここで, 式 (6) における ε は学習係数を表す.

3 評価実験

3.1 実験条件

本実験では, 外国話者の方言音声データセットである TIMIT データセットの中から, 母音音素のみを抽出したデータセットを用いて, 音素認識を通じた評価を行う. 入力特徴は 10 次元の LSP(Line Spectral Pairs), ゲイン, および各々の時間変化量からなる 22 次元のベクトルである. 実験条件を表 1 に示す.

表 1 実験条件

音素クラス数	5
学習用標本数	14488
試験用標本数	749
学習係数	1.0
学習回数	10000
再推定間隔	1000
境界近傍標本数	40
状態数	3
プロトタイプ数	3, 10, 20, 30, 50

3.2 実験結果

プロトタイプ数が 30 の場合の実験結果を図 2 に示す. 図 2 において, 横軸は学習回数, 上のパネルが経験的平均損失, 下のパネルが分類誤り率の推移である. また, 下のグラフにおける赤の曲線は学習標本に対する分類誤り率, 橙の曲線は試験用標本に対する分類誤り率, 紫の直線は 10 分割の交差検証 (CV: Cross Validation) 法を用いて求めた, ベイズ誤り参考値である.

3.3 考察

図 2 中の上のグラフより, 学習を経るにつれて損失が減少していることから, ベイズ境界性の最大化に向けて正しく学習できていることがわかる. また, 下のグラフでは試験用標本の分類誤り率がベイズ誤り参考値に漸近していることから, 推定された分類境界がベイズ境界に近付いていることが示唆される. しかし, 学習結果が収束しきっていない点か

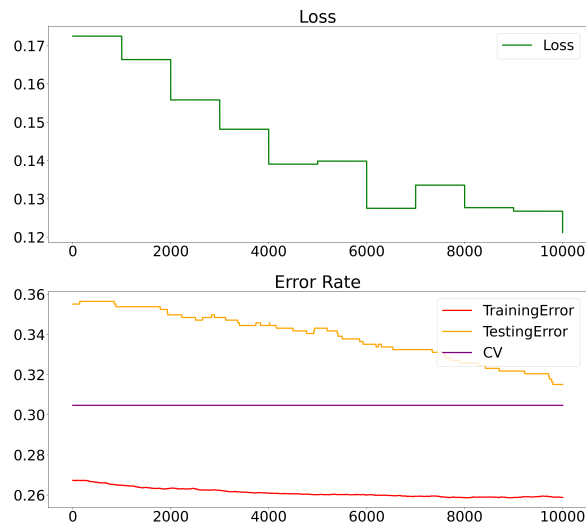


図 2 損失と分類誤り率の推移

ら, 学習回数や学習係数などのハイパーパラメータを調整することで, より良い学習結果が期待できると考える.

4 おわりに

本研究では, 音声パターンに MBB 学習法を適用させ, 可変長パターンに対する MBB 学習法の有用性を検証した. その結果, 可変長パターンを分類対象とした場合でも, 1 回の学習でベイズ境界を正しく推定し得る可能性が示唆された. 今後の展望として, ハイパーパラメータの探索や, 単語認識タスクへの応用などが挙げられる.

謝辞

本研究の一部は科研費 18H03266 の支援を受けて行われた.

参考文献

- 1) Masahiro Senda, David Ha, Hideyuki Watanabe, Shigeru Katagiri, and Miho Ohsaki. Maximum bayes boundary-ness training for pattern classification. In Proceedings of the 2019 2nd International Conference on Signal Processing and Machine Learnings, SPML '19, 2019.
- 2) 松廣達也, 北岡見生代, D. Ha, 渡辺秀行, 片桐滋, 大崎美穂. 大幾何マージン最小分類誤り学習法を用いた音声認識に関する実験的評価. 情報処理学会関西支部支部大会講演論文集, 2016.