

最小分類誤り学習法における損失関数の平滑度制御による局所解回避効果の検証

千田 将大[†] 松廣 達也[†] 渡辺 秀行[‡] 片桐 滋[†] 大崎 美穂[†]
[†] 同志社大 [‡] ATR

1. はじめに

最小分類誤り(MCE: Minimum Classification Error)学習法における平滑化分類誤り数損失の平滑度は、学習の未知標本耐性を制御し、かつ誤分類尺度空間における分布関数のパルツェン推定における窓(核)幅によって求められる[1]。また、この平滑性は原理的に、損失最小化における局所解回避にも有益であるが、その効果は明らかではない。本稿では、追加的な回避効果が期待できる疑似焼き鈍し法にヒントを得た平滑度制御法を導入し、それを用いて損失平滑性が持つ局所解回避効果を分析する。

2. 損失関数の平滑度の役割

MCE 学習において、最小化の対象である経験的平均損失は次のように表される。

$$L(\Lambda) = \sum_{n=1}^N \ell(\mathbf{x}_n; \Lambda) = \sum_{n=1}^N \frac{1}{1 + \exp(-\alpha_y d_y(\mathbf{x}_n; \Lambda))} \quad (1)$$

ここで、 $\ell(\mathbf{x}_n; \Lambda)$ は、標本 \mathbf{x}_n (クラス y に所属)をクラスモデルパラメータ Λ を用いて分類する際の平滑化分類誤り数損失であり、 N は学習標本の数、 d_y と α_y は、それぞれ y の誤分類尺度と損失関数平滑度である。また、 α_y は、誤分類尺度空間における分布関数のパルツェン推定からも定式化され、ガウスカネルを用いる場合、その窓幅 h_y との間に次の関係が成り立つ[2]。

$$\alpha_y = 4 / \sqrt{2\pi} h_y \quad (2)$$

式(2)より、窓幅 h_y が大きいとき、 α_y は小さく、損失 L は平滑になり、学習の未知標本耐性を向上させると共に、その損失最小化における局所解(局小点)を回避し易くする効果が期待される。

3. 損失関数の平滑度制御

パルツェン推定と式(2)に基づいて得られる平滑な L 学習に潜在する極小点を避けるため、 α_y の推定直後に一旦その値を推定値より強制的に小さくし(疑似焼き鈍し法におけるスケジューリングと概念的に近い)、より平滑な L の最小化を行う。なお、 α_y の自動設定法[2]と理想的なベイズリスクとの関連を重視し、この α_y の制御は、次式に基づいて自動設定法の繰り返しの間の M エポック(経験的平均損失の勾配に基づくパラメータ更新)において実行する。

$$\hat{\alpha}_y^{(m)} = \hat{\alpha}_y^{(0)} + (\dot{\alpha}_y - \hat{\alpha}_y^{(0)}) \sin\left(\frac{m}{M} \cdot \frac{\pi}{2}\right) \quad (3)$$

ここで、 $\dot{\alpha}_y$ は自動設定法によって得られた値であり、

$\hat{\alpha}_y^{(0)}$ は M エポック内の更新の初期値として $\dot{\alpha}_y$ 値を強制的に小さくした初期値、 $\hat{\alpha}_y^{(m)}$ はその第 m ステップにおける値である。

4. 評価実験

ETL-WD-Iデータベース(492単語、男女合計20人分)の音声データから、学習用に6人分音声データを、試験用に14人分を用いる単語音声認識を行い、損失最小化の状況を調査した。損失 L の最小化には最急降下法を用いた。学習エポックの回数は200とし、その更新に伴う学習係数値は50とした。自動設定法の実行間隔は $M = 50$ とした。また、 $\hat{\alpha}_y^{(0)} = \dot{\alpha}_y / 2$ とした(減衰率: 0.5)。比較のため、スケジューリングを行わず $\hat{\alpha}_y^{(m)} = \dot{\alpha}_y$ とする、自動設定法のみ(50エポック間隔で実行)を用いた学習も行った。

表1. 分類精度(%)

学習過程	学習用標本	試験用標本
α の制御なし	97.49	85.09
α の制御あり	97.29	84.64

表1に単語分類率を示す。 α_y 制御の有無にかかわらず、ほぼ同様の損失最小化がなされ、また式(3)を用いる制御法の達成精度がわずかながら低い。自動設定法による $\dot{\alpha}_y$ が、それ以上の平滑化を要さない程度まで L 平滑化し、一方で式(3)による平滑度の追加が最小化を遅らせた可能性が示唆される。

5. まとめ

平滑化分類誤り数損失の平滑度を追加的に制御することで、損失最小化における局所解の回避効果を調査した。実験を通し、誤分類尺度空間における平滑度自動設定法が局所解回避に一定程度の効果を持っていることなどを示唆する結果が得られた。一方、追加的な制御法には一層の改善が必要であることも示唆された。

謝辞: 本研究の一部は科研費(JP26280063)および文科省 H26年度私学戦略的研究基盤形成支援事業「ドライバ・イン・ザ・ループ」の支援を受けて行われた。また、ETL-WD-IデータベースはNII音声資源コンソーシアムからの提供を受けた。

参考文献

- [1] E. McDermott, et al.: Comput. Speech Lang., vol.18, pp.107-122, Apr. 2004.
 [2] Hideyuki Watanabe, et al.: IEEE, Proc. MLSP, Sep. 2011.