

データ拡張に基づくヒト胚性幹細胞を介した毒性予測 Toxicity Prediction via Human Embryonic Stem Cells Using Data Augmentation

小尾 岳人 *
Gakuto Obi

齋藤 彩斗 *
Ayato Saito

佐々木 悠人 *
Yuto Sasaki

加藤 毅†
Tsuyoshi Kato

1. はじめに

新しい化学物質を開発する際、人体にどのような影響が出るか調べるために、動物実験が使われてきた。しかし、近年、動物実験は問題視されるようになった。化粧品分野では、1990 年代～2000 年代初頭に、欧州を中心に Body Shop 運動、PETA キャンペーンなど、大規模な動物実験反対運動が起き、徐々に規制が強化されていった。工業化学・農薬分野では欧州では REACH と呼ばれる制度によって動物試験義務が課されてきたが、倫理的配慮、経済的コスト、時間的コストから動物試験削減に向けた計画が策定された。医薬品分野でも、倫理的側面に加えて、種差の違いにより、動物はヒトの代替モデルとして不十分である場合が問題視されており、これを受けて米国食品医薬品局 (FDA) は 2025 年 4 月に戦略的なロードマップを策定し、動物試験削減に向けて舵を切った。

各分野における動物試験削減への動きは、動物愛護の観点に基づく倫理的懸念の払拭に起因しているが、加えて代替手法の進歩も後押しした。代表的な代替手法には、QSAR、オルガノイド、マイクロ流体チップ、ヒト幹細胞由来細胞などがある。QSAR は化学構造パターンのみから毒性を予測する方法であり、すでに最初のスクリーニングのために、製薬・化粧品・化学業界で広く使われている。しかし、作用機序の無視、新規メカニズムへの対応が困難、代謝変換や複雑な細胞応答を無視といった弱点を抱えている。オルガノイドは長期培養が可能で、慢性毒性の評価にも使えることから注目はされている。しかし、個体差が大きい、血管系や免疫系が不足、栄養・酵素供給に制限がある、大量生産・標準化が難しいといった欠点が導入への壁となっている。マイクロ流体チップは、血管内皮や腎臓尿細管などに近い物理環境を模倣できる。しかし、デバイス作成が難しい、価格が高い、流量制御や培養条件維持が難しいといった障壁がある。ヒト幹細胞由来細胞は、所望の臓器や組織型に分化誘導させたものであり、ヒト由来であるため種差の問題に悩まされずに毒性試験に応用できる。しかし、分化誘導にはばらつきがあり、機能性成熟度が低い、分化誘導に時間・コストがかかるといった弱点に悩まされている。

ヒト多能性幹細胞による毒性評価 [2] は機械学習を用いている。毒性評価の信頼性を確保するには大量の訓練データを用いて高い学習精度を獲得する必要がある。しかるに、遺伝子発現データの測定に用いる RNA-seq や qPCR は費用や時間的コストがかかるため、訓練用例題セットの大規模化は非現実的である。これに対し、Mochida ら [4] は、Mixup データ拡張を導入して、

サポートベクトルマシン (SVM) の汎化性能を向上させた。一方、学習アルゴリズムには、SVM のほかにロジスティック回帰など多くの選択肢がある。毒性といっても、心毒性、肝毒性、腎毒性など多様であり、毒性の種類ごとに最適な学習器械は異なるため、多様な選択肢を試行できるように理論を整備しておくことは重要である。そこで本研究では、適用可能な学習器械のクラスを拡張し、そのいずれの学習器械でも数値的に安定で効率的に学習可能な最適化技術を開発した。

Mixup 拡張を使うと中間値クラスラベルが生じる。勾配降下ベースの最適化アルゴリズムを用いる場合、中間値クラスラベルは問題にならないが、通常の双対問題の最適化を行うアルゴリズムは、一般に直接適用できなくなる。これに対して、筆者らは、より広範な学習器械に適用できる双対問題を見つけた [5]。その理論をさらに詳細に解析した結果、各反復の時間複雑度を維持したまま、反復数を半減できる新しいアルゴリズムを発見した。本論文では、その新しい理論とアルゴリズムを提案するとともに、実データを用いた数値実験により従来法より高速に最適解に収束できることを実証する。なお、証明や導出の詳細については、別稿 [6] に記載してあるので、そちらを参照されたい。

2. Mixup データを用いた学習問題

本節では、Mixup 法 [3] によって、仮想的に化合物数を増やした際に生じる学習問題を定式化する。Mixup 法は、訓練用データセットから無作為に選択した 2 例題 $(\mathbf{x}_i^o, y_i^o), (\mathbf{x}_j^o, y_j^o) \in \mathbb{R}^d \times \{\pm 1\}$ 、およびベータ分布から生成した $\eta \in [0, 1]$ を使って、新しい例題

$$\mathbf{x}_{\text{new}} = (1 - \eta)\mathbf{x}_i^o + \eta\mathbf{x}_j^o, \quad y_{\text{new}} = (1 - \eta)y_i^o + \eta y_j^o$$

を生成する。これを繰り返すことでデータの水増しを実現する。すると、新たに生成した化合物のクラスラベル y_{new} は 2 値ではなく、中間値になっていることに注意されたい。

Mixup データ拡張によって得たデータセット $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n) \in \mathbb{R}^d \times \mathbb{R}$ から予測器 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ を訓練したい。Mixup データから予測器 f を学習するには、次式で定義される Mixup 損失関数 $\phi_{\text{mixup}}(\cdot; y): \mathbb{R} \rightarrow \mathbb{R}$ を用いる：

$$\phi_{\text{mixup}}(s; y) := \frac{1+y}{2}\phi_0(s) + \frac{1-y}{2}\phi_0(-s). \quad (1)$$

ここで用いられている $\phi_0: \mathbb{R} \rightarrow \mathbb{R}$ は中間値クラスラベルを用いない損失関数であり、 $\phi_0(0) > 0, \forall s \leq 0, \nabla \phi_0(s) < 0$ を満たすとする。たとえば、2 値交差エントロピー損失 (BCE 損失) $\phi_{\text{bce}}(s) := \log(1 + \exp(-s))$ 、2 次ヒンジ損失 $\phi_{\text{qh}}(s) := \frac{1}{2\gamma_{\text{sm}}} \max(0, 1 -$

*群馬大学情報学研究科

†群馬大学情報学部

$s)^2$ などが用いられる。ただし, $\gamma_{sm} \in (0, 1)$ は正の定数である。次節で述べる双対関数は損失関数の凸共役を含んでいる。関数 $\phi: \mathbb{R} \rightarrow \mathbb{R}$ の凸共役は $\phi^*(a) := \sup_{\zeta \in \mathbb{R}} a\zeta - \phi(\zeta)$ と表される。

カーネル法では、**正定値カーネル関数** $\kappa: \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}$ (e.g., RBF カーネル, 多項式カーネル) から構成される再生核ヒルベルト空間 $\mathcal{H}_\kappa$ の中から**正則化経験リスク** $R[f]$ を最小にする予測器 f を探すことになる:

$$R[f] := \frac{\lambda}{2} \|f\|_{\mathcal{H}_\kappa}^2 + \frac{1}{n} \sum_{i=1}^n \phi_{\text{mup}}(f(\mathbf{x}_i); y_i) \quad (2)$$

ただし, $\lambda > 0$ は正則化定数である。カーネル法の学習には、**確率的座標上昇法 (SDCA)** [1] など、双対問題を最大化することによって、最適化のためにチューニングを必要としないアルゴリズムが好まれてきた。ところが、Mixup 拡張を用いる場合、双対問題は効率的に最適化できないという問題が生じることが分かった。次節では、その双対問題を示し、なぜ最適化が複雑化するか述べる。

3. 従来 SDCA の枠組み

SDCA では、**フェンシエル双対**と呼ばれる、次の双対関数が採用されていた:

$$D_0(\boldsymbol{\alpha}) := -\frac{\lambda}{2} \|f_{\boldsymbol{\alpha}}\|_{\mathcal{H}_\kappa}^2 - \frac{1}{n} \sum_{i=1}^n \phi_{\text{mup}}^*(-\alpha_i; y_i)$$

ただし, $f_{\boldsymbol{\alpha}} := \frac{1}{\lambda n} \sum_{i=1}^n \alpha_i \kappa(\mathbf{x}_i, \cdot)$ とおいた; $\phi_{\text{mup}}(\cdot; y_i): \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ は $\phi_{\text{mup}}(\cdot; y_i)$ の凸共役を表す。本研究で扱う問題の場合、双対関数の値を評価するには Mixup 損失の凸共役を評価する必要がある。解析を進めると Mixup 損失の凸共役は**極小畳み込み** (infimal convolution) で表されることが判明した:

$$\begin{aligned} & \phi_{\text{mup}}^*(-\alpha_i; y_i) \\ &= \inf_{u \in \mathbb{R}} \left\{ \frac{1+y_i}{2} \phi_0^* \left(\frac{2u}{1+y_i} \right) + \frac{1-y_i}{2} \phi_0^* \left(\frac{2(\alpha_i+u)}{1-y_i} \right) \right\}. \end{aligned}$$

一般に極小畳み込みは数値計算を必要とするため、SDCA のようなアルゴリズムには厄介になってしまう。

従来 SDCA は、各反復 t で無作為に例題 $i \in [n]$ を選ぶ。 $\boldsymbol{\alpha}^{(t-1)} \in \mathbb{R}^n$ を第 $(t-1)$ 反復における双対変数の値とする。各反復で

$$J_{0,t}(\Delta \boldsymbol{\alpha}) := D_0(\boldsymbol{\alpha}^{(t-1)} + \Delta \boldsymbol{\alpha} \mathbf{e}_i) - D_0(\boldsymbol{\alpha}^{(t-1)}) \quad (3)$$

の下界となる

$$J_{0,t}(\eta q_{i,t}) \geq \frac{\eta F_{\text{van},i}^{(t-1)}}{n} + \frac{\gamma_{sm} q_{i,t}^2 \eta}{2n} \left(1 - \frac{\eta}{s_t} \right) =: J_{2,t}(\eta) \quad (4)$$

を最大化する $\eta \in [0, 1]$ を見つけて、 $\boldsymbol{\alpha}^{(t)} := \boldsymbol{\alpha}^{(t-1)} +$

$\Delta \boldsymbol{\alpha} \mathbf{e}_i$ のように更新する。ただし、

$$\begin{aligned} z_i^{(t-1)} &:= f_{\boldsymbol{\alpha}^{(t-1)}}(\mathbf{x}_i), \quad u_i^{(t-1)} := -\nabla \phi_{\text{mup}}(z_i^{(t-1)}), \\ q_{i,t} &:= u_i^{(t-1)} - \alpha_i^{(t-1)}, \quad \bar{s}_t := \left(1 + \frac{\kappa(\mathbf{x}_i, \mathbf{x}_i)}{\lambda n \gamma_{sm}} \right)^{-1} \\ F_{\text{van},i}^{(t-1)} &:= \phi_{\text{mup}}(z_i^{(t-1)}; y_i) \\ &\quad + \phi_{\text{mup}}^*(-\alpha_i^{(t-1)}; y_i) + \alpha_i^{(t-1)} z_i^{(t-1)}, \end{aligned}$$

とおいた。

下界 $J_{2,t}(\eta)$ は η の 2 次関数になっているので、 η^2 および η の係数を求めれば、平方完成により下界 $J_{2,t}(\eta)$ を最大化するのは容易である。各反復で下界 $J_{2,t}(\eta)$ を最大化する η を使うことができれば、最適解に近づくまでの必要な反復数が次のように保証される。

定理 1. [1] ϕ_0 が $1/\gamma_{sm}$ 平滑のとき、双対関数 $D_0(\boldsymbol{\alpha})$ の最大化に従来 SDCA を適用すると、反復数が

$$t \geq \frac{1}{\beta} \log \left(\frac{h_D^{(0)}}{\epsilon_P \beta} \right), \quad \left(\beta := \frac{\lambda \gamma_{sm}}{R_{mx}^2 + \lambda \gamma_{sm} n} \right) \quad (5)$$

に達したとき、 $\mathbb{E}[R[f^{(t)}] - R[f^*]] \leq \epsilon_P$ を満たす。ただし、 $f^{(t)}$ は反復 t の主変数の値、 f^* は主変数の最適解、 $h_D^{(0)}$ は初期値における双対関数の目的ギャップ、 $R_{mx} := \max_{i \in [n]} \sqrt{\kappa(\mathbf{x}_i, \mathbf{x}_i)}$ である。

ここで問題となるのは $F_{\text{van},i}^{(t-1)}$ に含まれる凸共役 $\phi_{\text{mup}}^*(-\alpha_i^{(t-1)}; y_i)$ の算出である。前述の通り、この凸共役は極小畳み込みとなり、一般に閉形式では表せないため、各反復ごとに数値探索を要求することになってしまう。

4. 提案法: MixupSDCA_{approx}

SDCA の更新は、2 次関数 $J_{2,t}(\eta)$ の最大化によって行われること、さらに、各反復で 2 次関数の最大化を実現すれば、線形収束が保証されることを述べた。しかし、Mixup 拡張したデータから学習を行う場合、係数に含まれる $F_{\text{van},i}^{(t-1)}$ の計算には数値探索が必要になってしまう。本節では、線形収束性の理論保証を保持しつつも、 $F_{\text{van},i}^{(t-1)}$ をある近似値 $\tilde{F}_i^{(t-1)}$ に置き換えることで数値探索を回避する方法を提案する。

$y_i \in \{\pm 1\}$ の場合は $F_{\text{van},i}^{(t-1)}$ を直接計算できるので、 $\tilde{F}_i^{(t-1)} := F_{\text{van},i}^{(t-1)}$ とする。これ以降、 $|y_i| < 1$ の場合を考える。この場合、 $\tilde{F}_i^{(t-1)}$ の値は

$$\begin{aligned} \tilde{F}_i^{(t-1)} &:= \phi_{\text{mup}}(z_i^{(t-1)}; y_i) - \tilde{\alpha} \tilde{\zeta}_t \\ &\quad + \alpha_i^{(t-1)} z_i^{(t-1)} - \phi_{\text{mup}}(\tilde{\zeta}_t; y_i) \end{aligned} \quad (6)$$

と決定する。ただし、 $\tilde{\alpha}$ 及び $\tilde{\zeta}_t$ の値は Algorithm 1 によって計算する。

Algorithm 1: Determine $\tilde{F}_i^{(t-1)}$.

```

1 begin
2   Let  $\alpha_i^\diamond := -\nabla\phi_{\text{mup}}(0; y_i)$ ;
3   if  $\alpha_i^\diamond < \alpha_i^{(t-1)}$  then
4     Find  $\zeta_t \in \mathbb{R}$  such that
        $\tilde{\zeta}_t \leq 0$  and  $-\nabla\phi_{\text{mup}}(\tilde{\zeta}_t; y_i) \leq \alpha_i^{(t-1)}$ 
5   else
6     Find  $\tilde{\zeta}_t \in \mathbb{R}$  such that
        $\tilde{\zeta}_t \geq 0$  and  $-\nabla\phi_{\text{mup}}(\tilde{\zeta}_t; y_i) \geq \alpha_i^{(t-1)}$ 
7   end
8   Let  $\tilde{\alpha} := -\nabla\phi_{\text{mup}}(\tilde{\zeta}_t; y_i)$ ; Use (6) to
     compute  $\tilde{F}_i^{(t-1)}$ ;
9 end

```

次の補助定理は、この近似法の線形収束性を証明するのに用いる。

補助定理 2. *Algorithm 1* で算出される $\tilde{F}_i^{(t-1)}$ は不等式 $\tilde{F}_i^{(t-1)} \leq F_{\text{van},i}^{(t-1)}$ を満たす。

ステップサイズは、 $F_{\text{van},i}^{(t-1)}$ を近似値 $\tilde{F}_i^{(t-1)}$ に入れて次のように決定する：

$$\eta_{\text{approx},t} := \min \left\{ 1, \bar{s}_t \cdot \max \left\{ 1, \frac{\tilde{F}_i^{(t-1)} + \frac{1}{2}\gamma_{\text{sm}}q_{i,t}^2}{\gamma_{\text{sm}}q_{i,t}^2} \right\} \right\}. \quad (7)$$

近似値 $\tilde{F}_i^{(t-1)}$ を使って、このようにステップサイズを決定する SDCA を **MixupSDCA_{approx}** と呼ぶことにする。

Algorithm 1 の Line 4 および Line 6 の実装例は A 節に記載した。この実装を用いると、各反復の計算時間はデータセットのサイズに比例するので、Mixup しない従来 SDCA の各反復と同じ時間複雑度を維持できる。さらに、必要な反復数の上界に関しても、Mixup データからの学習にも関わらず、従来 SDCA の各反復の上界を維持することが判明した。

定理 3. ϕ_0 が $1/\gamma_{\text{sm}}$ 平滑のとき、双対関数 $D_0(\alpha)$ の最大化に近似 MIXUPSDCA (Algorithm 2) を適用しても、反復数が

$$t \geq \frac{1}{\beta} \log \left(\frac{h_D^{(0)}}{\epsilon_P \beta} \right), \quad \left(\beta := \frac{\lambda \gamma_{\text{sm}}}{R_{mx}^2 + \lambda \gamma_{\text{sm}} n} \right) \quad (8)$$

に達したとき、 $\mathbb{E}[R[f^{(t)}] - R[f^*]] \leq \epsilon_P$ が保証される。

Algorithm 2: 近似 MIXUPSDCA.

```

1 begin
2   Choose  $\alpha^{(0)} \in \text{dom}(-D_0)$ ;
3    $f^{(0)} := \frac{1}{\lambda n} \sum_{i=1}^n \alpha_i^{(0)} \kappa(\mathbf{x}_i, \cdot)$ ;
4   for  $t := 1$  to  $T$  do
5     Select  $i$  at random from  $\{1, \dots, n\}$ ;
6      $z_i^{(t-1)} := f^{(t-1)}(\mathbf{x}_i)$ ; Compute  $\bar{s}_t$ ;
7     Run Algorithm 1 to compute  $\tilde{F}_i^{(t-1)}$ ;
8     Use (7) to compute  $\eta_{\text{approx},t}$ ;
9      $\alpha^{(t)} := \alpha^{(t-1)} + q_{i,t} \eta_{\text{approx},t} \mathbf{e}_i$ ;
10     $f^{(t)} := f^{(t-1)} + \frac{1}{\lambda n} \kappa(\mathbf{x}_i, \cdot) q_{i,t} \eta_{\text{approx},t}$ ;
11  end
12 end

```

5. 毒性予測実験

ヒト胚性幹細胞を用いた毒性予測において、Mixup によるデータ拡張が SVM 以外の学習器械にも有効か調査するため、[4] の方法に倣って評価実験を実施した。すなわち、Yamane ら [2] が用いた 24 種類の化学物質をヒト胚性幹細胞に暴露させて得られる遺伝子発現データを用いた。この 24 化学物質に対して、神経毒性の有無および心毒性の有無は判明している。これらを 2 クラス分類の正解データとして、2 種類のベンチマークデータセットを用意した。神経毒性の正例は 13 個、負例は 11 個、心毒性の正例は 15 個、負例は 9 個であった。

一つ抜き交差検証法を用いて予測性能を評価した。一つ抜き交差検証法では、24 化学物質のうち 1 個だけを評価用とし、23 個を訓練に用いる。この 23 個を Mixup データ拡張によって、50 個例題を増やした。損失関数には、平滑ヒンジ損失 (SM loss)、2 値交差エントロピー損失 (BCE loss)、2 次ヒンジ損失 (QH loss) の 3 種類を用いた。カーネル関数には RBF カーネルを用いた。正則化パラメータ λ および RBF カーネルのカーネル幅のパラメータは 23 個の訓練用例題内でさらに一つ抜き交差検証法を行うことで最適値を推定した。

予測性能の定量評価には AUROC を用いた。Mixup によるデータ拡張は乱数を用いるので、試行ごとに結果が異なる。そこで、5 回試行して得られる 5 個の AUROC の平均で評価することとした。図 1 に神経毒性および心毒性の予測性能を示す。いずれの毒性においても、またいずれの損失関数を用いた場合も、Mixup によって予測性能が向上することが確認できた。神経毒性および心毒性においては、平滑ヒンジ損失より BCE 損失のほうが高い AUROC を得た。以上より、Mixup は多くの学習器械に対して汎化性能の向上をもたらすこと、また、SVM が最適とは限らないことが分かった。

6. 収束実験

提案する最適化アルゴリズム MIXUPSDCA_{approx} の最適解への収束の速さを評価するため、数値実験を行っ

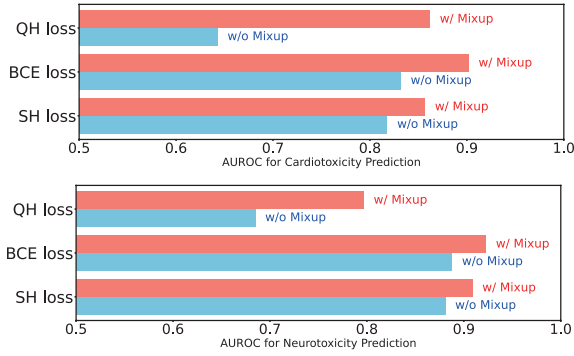


図 1: 毒性予測性能.

表 1: 学習に要した時間の実測値.

アルゴリズム	計算時間 [sec]
MIXUPSDCA _{naïve} [1]	$9.06 \cdot 10^2$
MIXUPSDCA _{decomp} [5]	$2.12 \cdot 10^2$
MIXUPSDCA _{approx} (提案法)	$1.67 \cdot 10^2$
SGD w/ $\eta = 10^{-1}$	N/A
SGD w/ $\eta = 10^{-2}$	$1.48 \cdot 10^3$
SGD w/ $\eta = 10^{-3}$	N/A
SGD w/ $\eta = 10^{-4}$	N/A

た. 各反復で数値探索を行う従来の SDCA を使った場合 (MIXUPSDCA_{naïve}), 小尾らが提案するアルゴリズム (MIXUPSDCA_{decomp}) [5], 確率的勾配降下法 (SGD) と比較を行った. SGD のステップサイズ η には 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} の 4 通りを試した. 各アルゴリズムの $R[f] - R_* < 10^{-5}$ に達するまで時間を比較した. 訓練用例題には, 毒性予測実験で用いた Mixup データを用い, BCE 損失を採用した. 正則化パラメータ λ 3 通り, RBF カーネルの幅パラメータ 4 通りを直列に実行したときの合計時間を測定した. その結果を表 1 に示す. N/A はいずれかの条件で 5000 エポック内に収束しなかったことを示す. MIXUPSDCA_{approx} が最も早く収束した. SGD は $\eta = 10^{-2}$ 以外では 5000 エポック内に収束しない場合があった.

7. 結論

本論文では, 代表的な機械学習アルゴリズムである SDCA を一般化して Mixup データ拡張した例題から学習できるようにする理論を整備した. SDCA の各反復は, 損失関数の凸共役の値の計算を要求していた. Mixup データ拡張した場合, 損失関数の凸共役は極小値のみで表されるため, 値の算出は数値探索が必要となる. これに対して, 筆者らは近似的に値を求める方法を見つけた. 特筆すべき点は, 近似したとしても, 同じ反復数の理論保証を維持できる点である. このアルゴリズムを使って, ヒト多能性幹細胞を用いた化合物の毒性予測器の学習に適用したところ, Mixup 拡張によって汎化性能を向上できることを確かめた. また, 我々の近似解法により, 各反復で数値探索を行うより, 有意に高速に最適解に収束できることを確認した.

A. Algorithm 1 の Line 4 及び Line 6 の実装例

筆者らは Algorithm 1 の Line 4 及び Line 6 を次のように実装した:

$$\mathcal{L}_i := \{\zeta \in \mathbb{R} \mid \phi_{\text{mup}}(\zeta; y_i) \leq n\phi_0(0)\}, \quad (9)$$

$$b_{u,i} := \inf \mathcal{L}_i \text{ and } b_{\ell,i} := \sup \mathcal{L}_i$$

とおく. Algorithm 1 で定義したように $\alpha_i^\diamond := -\nabla \phi_{\text{mup}}(0; y_i)$ とおく. 2つの場合がある: (i) $\alpha_i^{(t-1)} < \alpha_i^\diamond$ および (ii) $\alpha_i^{(t-1)} \leq \alpha_i^\diamond$. ケース (i) では, $\tilde{\zeta}_t$ の値を

$$\tilde{\zeta}_t := \min \left\{ \zeta \in \bar{\mathcal{Z}}_{p,i} \mid \alpha_i^{(t-1)} \leq -\nabla \phi_{\text{mup}}(\zeta; y_i) \right\} \quad (10)$$

とする. ただし,

$$\bar{\mathcal{Z}}_{p,i} := \left\{ \exp \left(\frac{k}{n} (4 + \log b_{\ell,i}) - 4 \right) \mid k \in \{0\} \cup [n] \right\} \quad (11)$$

とした. ケース (ii) では, $\tilde{\zeta}_t$ の値を

$$\tilde{\zeta}_t := \max \left\{ \zeta \in \bar{\mathcal{Z}}_{n,i} \mid -\nabla \phi_{\text{mup}}(\zeta; y_i) \leq \alpha_i^{(t-1)} \right\} \quad (12)$$

とする. ただし,

$$\bar{\mathcal{Z}}_{n,i} := \left\{ -\exp \left(\frac{k}{n} (4 + \log(-b_{u,i})) - 4 \right) \mid k \in \{0\} \cup [n] \right\} \quad (13)$$

とした.

謝辞: 本研究は JSPS 科学研究費 22K04372 および 23H03555 の助成, および JST ASTEP JPMJTR234E の支援を受けたものである.

参考文献

- [1] S. Shalev-Shwartz & T. Zhang. Stochastic dual coordinate ascent methods for regularized loss. *JMLR*, 14(1):567–599, 2013.
- [2] J. Yamane et al., StemPanTox: A fast and wide-target drug assessment system for tailor-made safety evaluations using personalized iPS cells. *iScience*, 25(7):104538, 2022.
- [3] H. Zhang et al, mixup: Beyond empirical risk minimization. In *ICLR*, 2018.
- [4] Rikuto Mochida, Miya Nakajima, Haruki Ono, Takahiro Ando, and Tsuyoshi Kato, “Mixup SVM Learning for Compound Toxicity Prediction Using Human Pluripotent Stem Cells,” *IEICE Transactions on Information & Systems*, vol. E107-D, no. 12, pp. 1542–1545, Dec. 2024, doi: 10.1587/transinf.2024ed18040.
- [5] 小尾 岳人, 齋藤 彩斗, 加藤 毅, 「ヒト多能性幹細胞による化合物毒性予測のための機械学習アルゴリズムの検討」, 情報処理学会 第 87 回全国大会論文集 (第 4 分冊), pp. 671–672, 2025 年 3 月 13 日, 大阪.
- [6] Gakuto Obi, Ayato Saito, Yuto Sasaki, and Tsuyoshi Kato, “Linearly Convergent Mixup Learning,” 2025, <https://arxiv.org/abs/2501.07794>