

課題 1: 例からの学習

提出締切 5月20日(金)正午

1. 2次元の信号 $x = (x_1, x_2)^T$ (Tは転置)を受け取り

$$z = \sum_{i=1}^2 w_i x_i - h \quad (1)$$

を出力するニューロンのモデルを考えよう. ここで係数 $w = (w_1, w_2)^T = (-0.5, 1.0)^T$, ニューロンの閾値(しきいち) $h = 2.0$ とする. 平面(横軸 x_1 , 縦軸 x_2)を描き, このニューロンの出力が正になる領域を示しなさい(手書きでよい).

2. 講義の web ページ <http://www.cs.miyazaki-u.ac.jp/~date/lectures/bis/> より, 2次元のデータが10個書き込まれているファイル `apples.dat` (りんごのデータ)と `oranges.dat` (みかんのデータ)をダウンロードし, 2次元平面上に, それぞれ異なる記号(と×など)でプロットせよ. ちなみに `apples.dat` は, 以下の内容ではじまる10行のデータである.

```
1.0627 1.1438
0.4268 1.5955
1.5946 0.9812
```

⋮

3. リンゴのデータとみかんのデータを正確に区別できるような機械を作りたい. 2次元の信号 $x = (x_1, x_2)^T$ を受け取り

$$z = f\left(\sum_{i=1}^2 w_i x_i - h\right) = f\left(\sum_{i=0}^2 w_i x_i\right) \quad (2)$$

を出力するニューロンのモデルを考えよう. ここで $f(u) = \frac{1}{1 + e^{-\lambda u}}$ とする(λ は入出力特性を記述するパラメータ. ここでは $\lambda = 1$ とする). 常に1という入力 $x_0 = 1$ があると考えれば, $w_0 = -h$ と書き直して記述が簡単になる. $w = (w_0, w_1, w_2)^T$ の各要素 w_i は区間 $[0, 1)$ の一様乱数に設定する(時間 $t = 0$ での初期値). このモデルが正の値を返すと「りんご」, そうでなければ「みかん」と判別することにしよう. この初期状態のモデルで, 20個のデータ(例題 x^1, x^2, \dots, x^{20})のうち, 何個, 正しく判別できるか計算機で実験せよ(w の値をさまざまな乱数値に変えて1000回試し, 平均を求めてみる).

4. 例題を一つ入力するごとに, 各係数 $w_i, i = 0, 1, 2$ を, 以下の式にしたがい更新せよ.

$$w_i := w_i + \eta(y^\alpha - z)z(1 - z)x_i^\alpha, \quad i = 0, 1, 2 \quad (3)$$

ここで z は入力 x^α に対する出力, y^α は, その例題に対する望ましい答え (0 もしくは 1), η は学習係数で $\eta = 0.1$ とする. ひと通り例題を入力した後で, 20 個の例題データのうち, 何個, 正しく判別できるようになっているか, 計算機実験で求めよ.

5. 4. で示した手順 (学習) を繰り返し, 横軸に学習回数, 縦軸に正解率 (またはエラー率) をプロットした図 (学習曲線) を作成せよ.

6. 学習係数 η , シグモイド関数のパラメータ λ の値を変更し, 学習が進む過程がどう変化するか観察し, パラメータの意味を考察しなさい. 学習の進み具合は, 誤差

$$E = \sum_{\alpha=1}^{20} E^\alpha, \quad \text{where } E^\alpha = \left[y^\alpha - f \left(\sum_{i=0}^2 w_i x_i^\alpha \right) \right]^2 \quad (4)$$

など, 自分が自然だと思う量を用いなさい.

7. データファイル名 `apples3.dat` と `oranges3.dat` をダウンロードし, 同様に実験せよ. 先の例題より分類が難しくなっていることを確認せよ.

8. 与えられた例題に対してだけ, 正解を出せるようになるだけでは面白くない. 問題を無数に生成して機械を評価しよう. 次のような問題を考えてみる. 平面上の 2 点をランダムに決め, これを「りんご」「みかん」の中心座標とし, それぞれ $x^a = (x_1^a, x_2^a)$, $x^o = (x_1^o, x_2^o)$ とする. これに平均 0, 分散 σ^2 の正規分布にしたがう乱数 n_1, n_2 を加え, $x^\alpha = (x_1^a + n_1, x_2^a + n_2)$ として, それぞれの例題を 10 個作る (学習用の例題, σ は適当な値に決める. σ が小さいと問題が易くなる). 機械の性能を評価するときは, 学習用の例題とは関係なく, 評価用だけに使う例題をその場で作る (例えば 1000 個). 学習用に例題 20 個を用い, 評価用の例題を 1000 個生成し, 学習曲線を描きなさい (例題の数 20 を変えて実験すると面白いかもしれない).

注意: 課題が, 手に負えない場合は, はやめに A-334 か A-423 に相談にくること.

補足: 細かい点 (学習回数など) は自分で適当に決めて進めればよい. その他, 課題に曖昧な点があった場合も適当に解釈していい. その場合, どこをどう解釈したか, そう解釈した理由等をレポートに記述すること. レポートの最後には, 感想, 質問などを記述してほしい. 今後のため, 「理解できなかった」という記述よりは, 具体的にどこの部分から分からなくなったか記述してもらえれば助かる. レポートの L^AT_EX を使った簡単な書き方は

<http://www.cs.miyazaki-u.ac.jp/~date/lectures/latex/latexreport.html>

を参照.