

## 宿題 1

(1)

拡張特徴ベクトル・拡張重みベクトル・正規化を実装し、パーセプトロン (バッチもしくはオンライン) を実装する。また、Linearly separable (linear)・Linearly non-separable (nonlinear)・Skewed linearly separable (slinear) の 3 種類のデータについてパーセプトロンを実行し、その振る舞いを調べる。なお、括弧書きで略称を示した。以降は略称を用いる。

拡張重みの更新式は次のようになる。

$$w^{(t+1)} = w^{(t)} + \rho \sum_{x \in \tilde{X}} x \quad (1)$$

ここで、 $t$  は更新回数、 $\rho$  は学習率、 $\tilde{X}$  は識別誤りとなっているすべてのサンプルである。これを実装すればよい。

データは全て 2 カテゴリからなる二次元点群で、slinear のみ 500 サンプル、他の 2 つは 100 サンプルある。本来、学習の終了条件は点が全て正しいカテゴリに分類されたときであるが、この設定では収束しない場合があるので。更新を 100 回行ったときも学習を終了させる。学習率  $\rho = 0.1$  としたときの結果を以下の表 1 に、各データの点群と境界のプロットを図 1-3 に示す。

結果から、線形分離可能なデータ (linear) や、完全に線形分離するのは不可能だが大半がうまく分けられるようなデータ (slinear) に対しては、パーセプトロンはうまく機能することがわかる。一方、線形分離不可能なデータ (nonlinear) に対しては当然全く機能しない。また、学習中の境界の挙動を観察した結果を以下に記す。linear については、図 1 にある最終的に得られる境界を中心に、それとほぼ平行に減衰振動するような形で境界が移動していき、最終的に収束した。nonlinear については、境界がプロットの範囲中に現れては消えることを繰り返し続け、最後まで振動し続けていた。slinear については、まず図 3 の右上と左上を線対称に置くような位置に境界が引かれ、そこから徐々に立ち上がっていき、最終的に図の中心付近の点を (ある程度) うまく分類できるようになった。linear を減衰振動、nonlinear を振動と表現したことに倣うと、slinear は過減衰のような挙動であった。

表 1: 結果

データの種類	$w_0$	$w_1$	$w_2$	識別誤り数	識別誤り率
linear	-1.002	4.727	2.401	0	0%
nonlinear	-4.002	1.870	-0.1073	38	38%
slinear	-0.2024	20.30	0.01964	2	0.4%

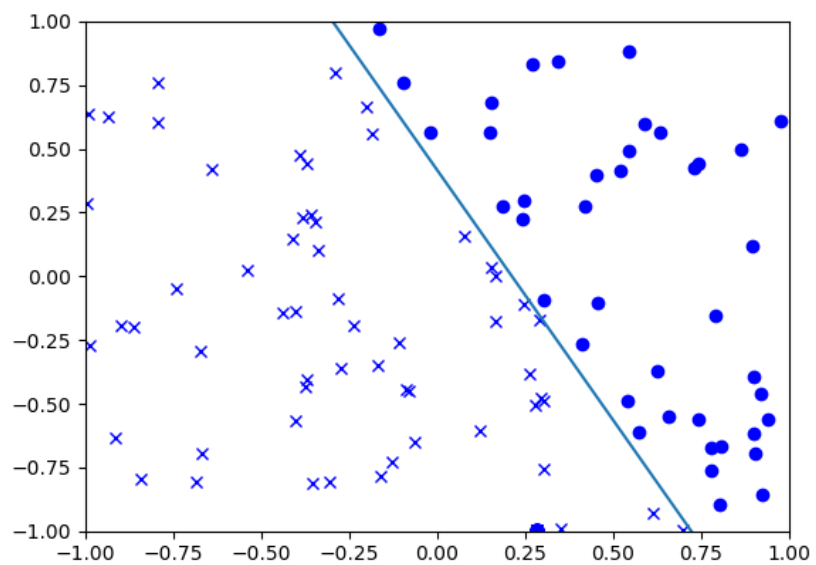


図 1: linear データに対してパーセプトロンを適用したときの点群と境界

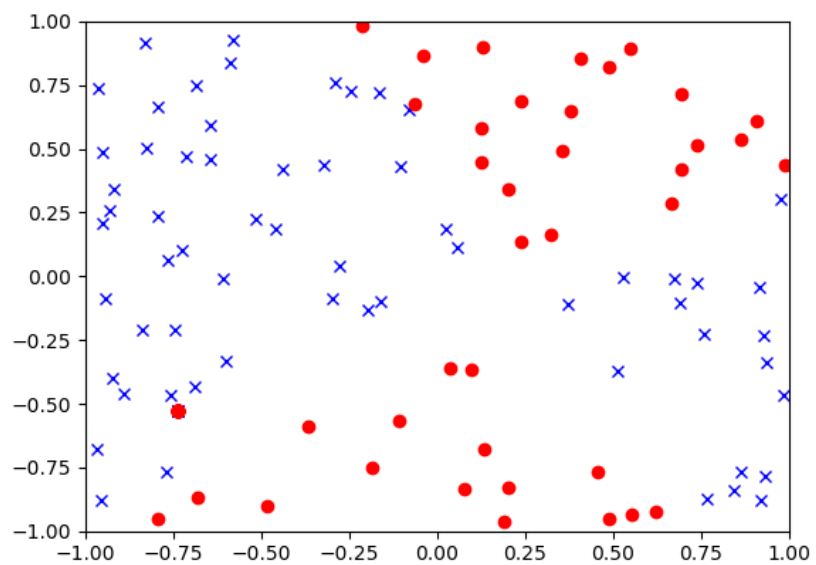


図 2: nonlinear データに対してパーセプトロンを適用したときの点群と境界

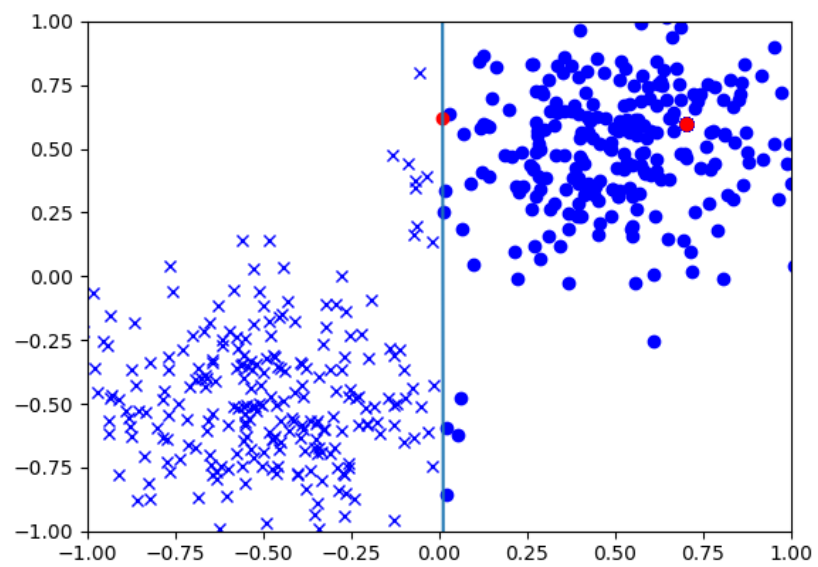


図 3: slinear データに対してパーセプトロンを適用したときの点群と境界

(2)

MSE 法による識別器を実装し、3 種類のデータにより学習させる。また、その振る舞いについて調べる。  
重みの解析解は次のようになる。

$$\mathbf{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2)$$

また、LMS 法を用いるときの、拡張重みの更新式は次のようになる。

$$\hat{\mathbf{w}}^{(t+1)} = \hat{\mathbf{w}}^{(t)} - \rho (\hat{\mathbf{w}}^T \hat{\mathbf{x}} - \mathbf{y}) \hat{\mathbf{x}} \quad (3)$$

学習の終了条件は重みの変化量のノルムが  $1 \times 10^{-5}$  以下となったときとする。今回は学習率をデータごとに変えたため、それも結果にまとめて記すこととする。結果を以下の表 2 に、各データの点群と境界のプロットを図 4-6 に示す。

結果はパーセプトロンのときと同様、linear と slinear に対してはある程度対応でき、nonlinear に対してはうまくいかないというものになった。また、精度はパーセプトロンよりも劣る形となった。しかし、パーセプトロンと異なり、学習の終了条件が実現可能なものになっており、更新回数で切り上げないと収束せず無限ループに陥るということが起きないのは利点である。また、学習中の境界の挙動は、勾配降下法を用いているため、当然減衰振動のようなものとなった。

表 2: 結果

データの種類	学習率	$w_0$	$w_1$	$w_2$	識別誤り数	識別誤り率
linear	0.015	-0.2164	1.379	0.6243	6	6%
nonlinear	0.015	-0.2141	0.5494	0.03766	39	39%
slinear	0.001	-0.007554	1.035	0.7157	13	2.6%

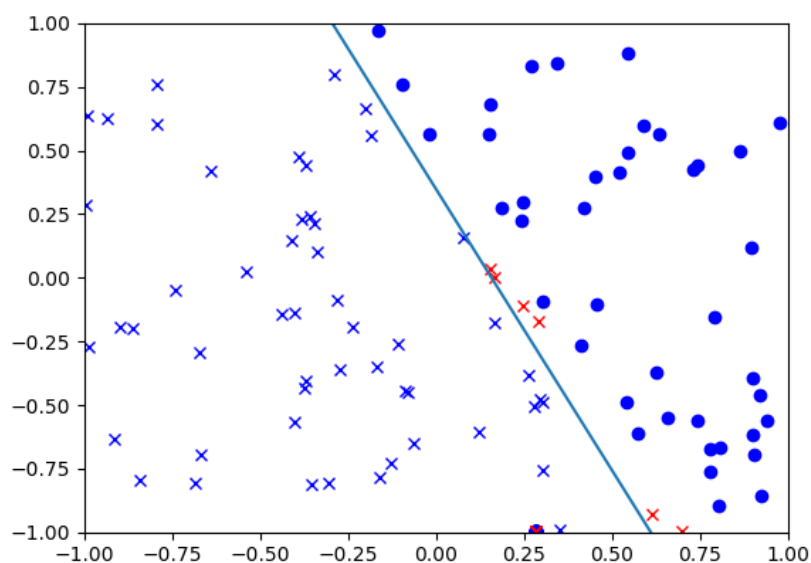


図 4: linear データに対して MSE 法を適用したときの点群と境界

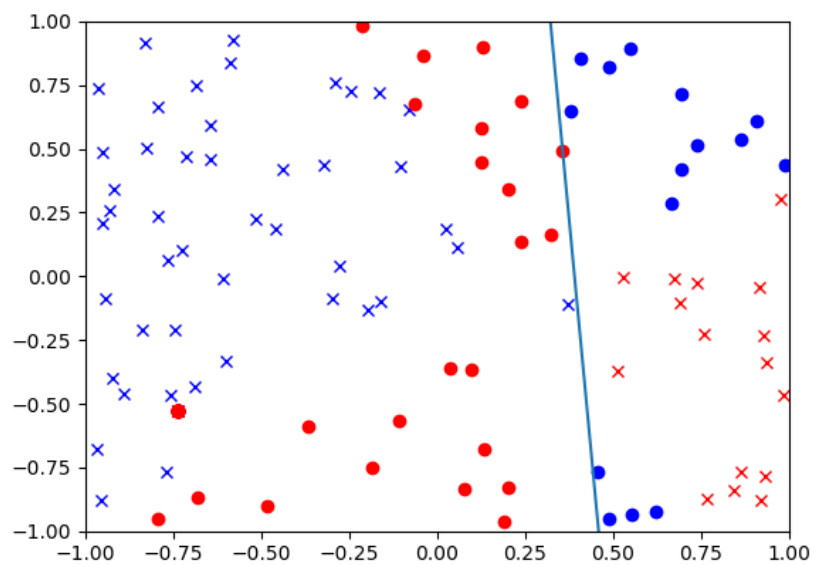


図 5: nonlinear データに対して MSE 法を適用したときの点群と境界

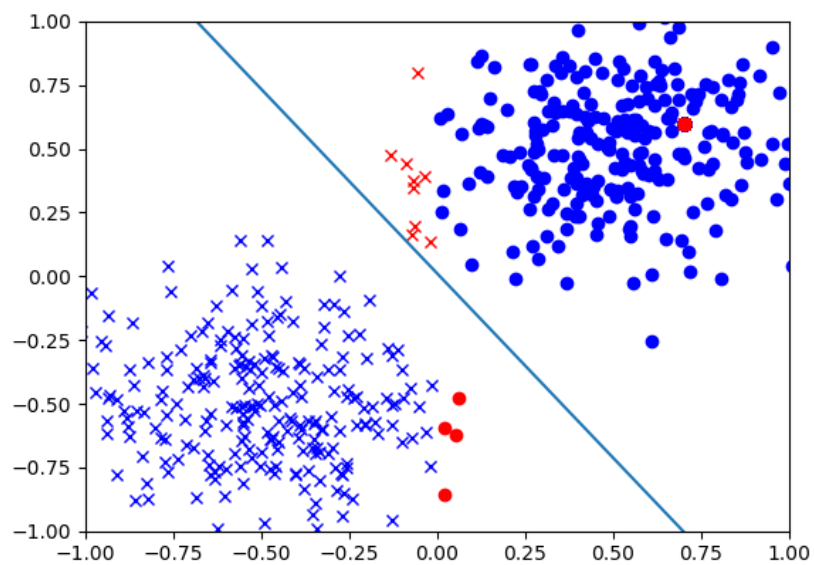


図 6: slinear データに対して MSE 法を適用したときの点群と境界