パターン認識と学習

岡山大学大学院

講師　竹内孔一

本日の内容

- 学習
  - Widrow-hoff learning-rule
  - パーセプトロン再考

Widrow-hoffの学習規則

- 何をするもの?
  - 識別関数を学習させる
- アイデア
  - 学習データxpに対して教師信号upを用意
  - 誤差を最小にする
- 得られるもの
  - 重回帰分析
  - widrow-hoff学習規則(パーセプトロンも含む)

2つの解法

- 解析的手法
  - 誤差最小→ 誤差の式(2乗)を微分=0
  - 直接解が求まる

  (欠点) 式は求まっても実際計算が無理
    - 大次元行列計算
- 数値解析による手法
  - 最急降下法
  - 逐次更新によって最小値を求める

パーセプトロン

- 直交化学習
  - 教師ベクトルと学習データとの二乗誤差最小
  - パーセプトロンの学習規則もこの一種
- ベクトル図からの解釈
  - 重み空間で x のはずれ具合を評価関数とする
  - 最急降下法でパーセプトロンの変更手続きが求まる

練習6

- 図3.2のrがなぜそうなるのか計算で求めよ

回答例

まず求めるrは図1のようにθを仮定して

$$r = \|\mathbf{w}\| \cos\theta \quad \text{(k1)}$$

とおける. このcosθはベクトルxとの関係でいうと

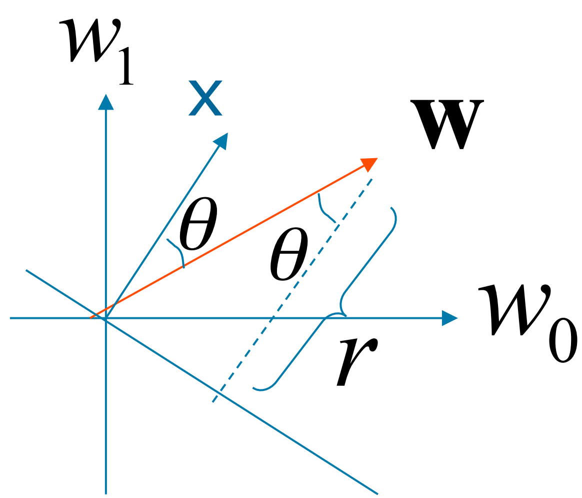
$$\cos\theta = \frac{\mathbf{w} \cdot \mathbf{x}}{\|\mathbf{w}\|\|\mathbf{x}\|}$$

ただしcosθ>0なので

$$\cos\theta = \frac{\|\mathbf{w} \cdot \mathbf{x}\|}{\|\mathbf{w}\|\|\mathbf{x}\|}$$

左を(k1)に代入して

$$r = \frac{\|\mathbf{w} \cdot \mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{w}^t \mathbf{x}\|}{\|\mathbf{x}\|}$$



$$g(\mathbf{x}) = \mathbf{w}^t \mathbf{x} = 0$$

図1 重みw空間

他の回答例

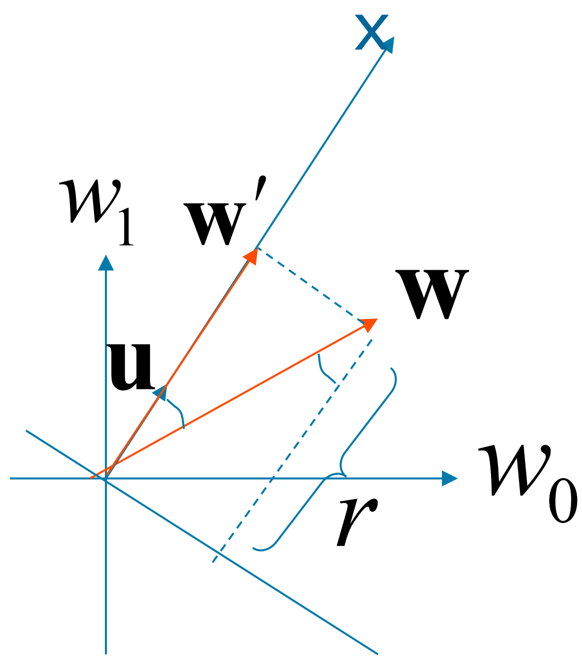
- 正射影で考える

wをxに正射影しよう. x上の単位ベクトルをuとすると

$$\mathbf{w}' = (\mathbf{w}\cdot\mathbf{u})\mathbf{u}$$

ここで $\mathbf{u} = \dfrac{\mathbf{x}}{\|\mathbf{x}\|}$

である. 求めるrはw′のノルムなのでそれを計算する.



$$g(\mathbf{x}) = \mathbf{w}^t\mathbf{x} = 0$$

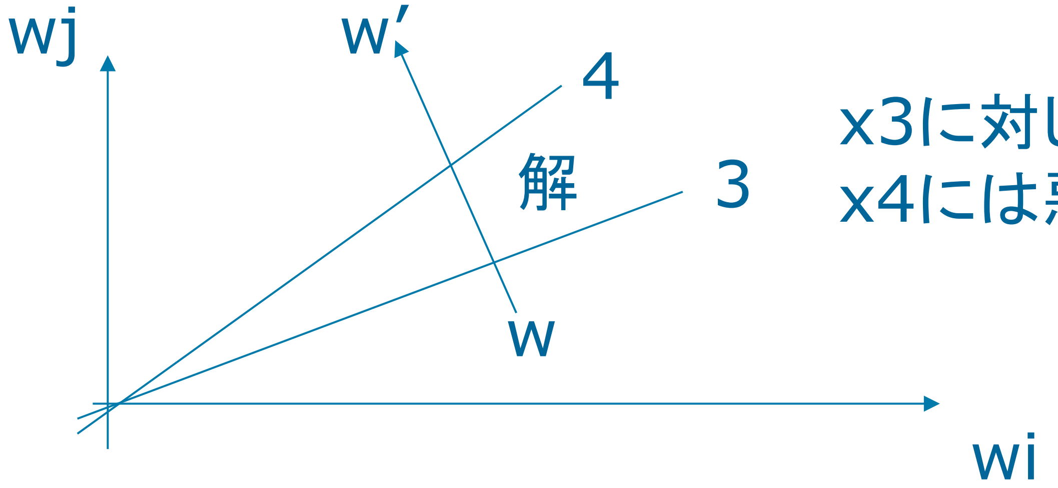
$$r = \|\mathbf{w}'\| = \|(\mathbf{w}\cdot\mathbf{u})\mathbf{u}\| = \|\mathbf{w}\cdot\mathbf{u}\| = \frac{\|\mathbf{w}\cdot\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\|\mathbf{w}^t\mathbf{x}\|}{\|\mathbf{x}\|}$$

uは長さ1なので||u||は消える

注意点

- Widrow-hoff の学習規則とパーセプトロン
  - wの更新の式 $\mathbf{w}'_i = \mathbf{w}_i - \rho(g_{ip} - b_{ip})\mathbf{x}_p$ はJpの最小化であってJではない
  - よってこの学習規則の更新は1つのxpに最適でも全体のJでは悪くなることがある
  - よって常にJ=ΣJp が最小になるようにする

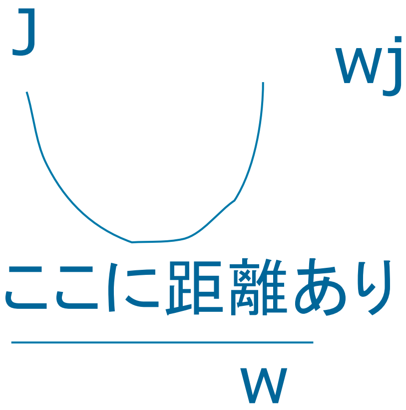
例)

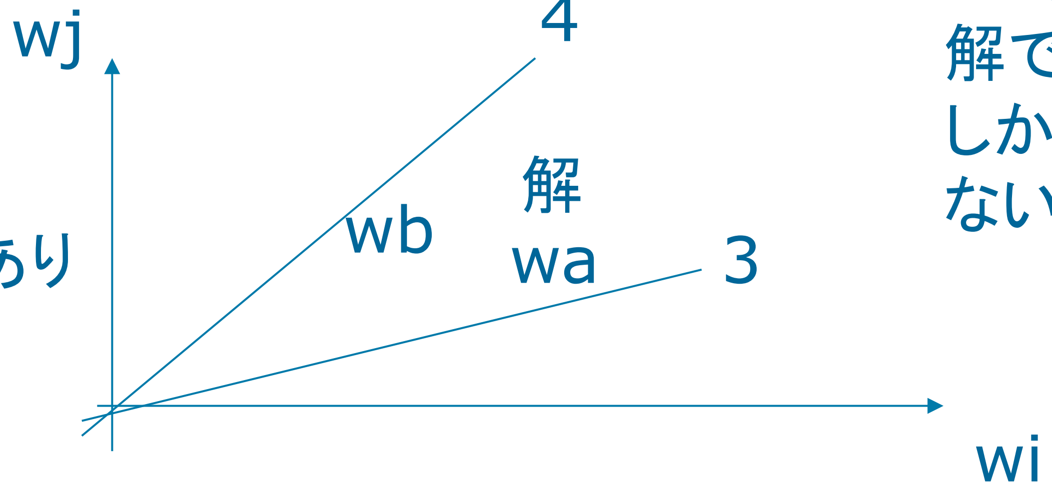


x3に対して修正しても
x4には悪くなる

注意点

- Widrow-hoff の学習規則とパーセプトロン
  - パーセプトロンは量(識別面からの離れ具合)を見ていない
  - Widrow-hoffの方は教師信号との差を見ている
  - Widrow-hoffはたとえ判別が全て正しくても誤差は残る





waとwbはどちらも解で学習はとまる
しかし誤差量は0ではない