多分类问题的线性分类

翟婷婷

扬州大学
信息工程（人工智能）学院
zhtt@yzu.edu.cn

2023年春

课程目标

- 掌握多分类问题转化为二分类问题进行求解的方法：
  - ✓一对多形式(one-versus-all)
  - ✓一对一形式(one-versus-one)
  - ✓最大值形式

多分类问题的线性分类

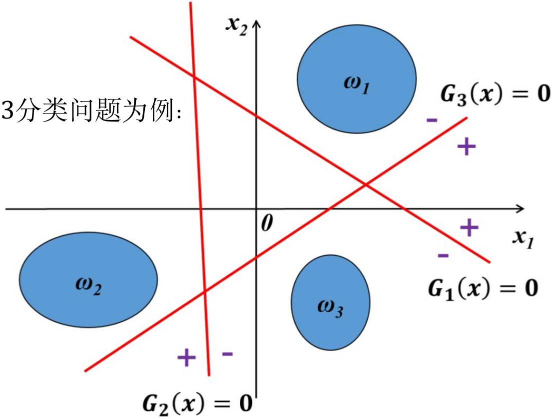
- 一个线性判别函数$G(x)$只能将特征空间划分为**两个**决策域：$G(x)>0$ 和$G(x)<0$，用于二分类问题。
- 多分类问题有多个决策域，用一个线性判别函数无法进行分类。
- 解决办法：训练多个线性判别函数，根据它们的分类决策之间的逻辑关系进行分类。
- 训练方式：
  - ✓ 一对多形式(one-versus-all, one-versus-rest)
  - ✓ 一对一形式(one-versus-one)
  - ✓ 最大值形式

多分类问题的线性分类

- 一对多的训练方式：

  对于训练集中的每一类实例，训练一个线性判别函数，将属于该类的实例和不属于该类的实例分开。

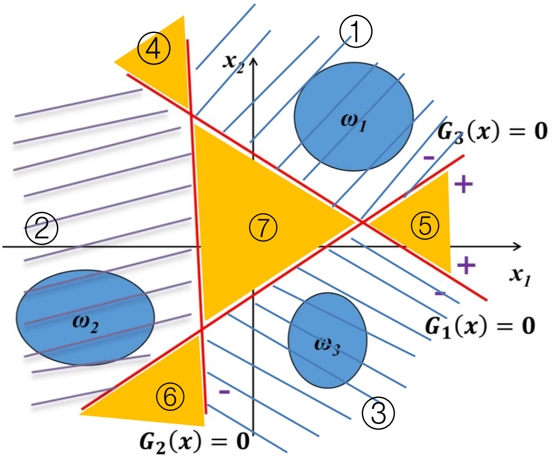
以3分类问题为例：



$G_1(x) = 0$将ω_1类和其余类分开，使得
$G_1(x) > 0, x \in \omega_1$,
$G_1(x) < 0, x \notin \omega_1$

如何确定多分类的决策规则呢？

一对多训练方式对特征空间的划分



编号	$G_1(x)$	$G_2(x)$	$G_3(x)$
①	>0	<0	<0
②	<0	>0	<0
③	<0	<0	>0
④	>0	>0	<0
⑤	>0	<0	>0
⑥	<0	>0	>0
⑦	<0	<0	<0

④⑤⑥⑦都是不可识别区域。

一对多训练方式的分类规则

- 原则：对实例$\boldsymbol{x}$进行分类时，仅当所有判别函数的分类决策没有冲突时才能做出多分类决策，否则无法分类。
- 对于3分类问题，3个判别函数的分类决策没有冲突时的情况：仅当只有一个判别函数的取值> 0，剩余两个判别函数的取值< 0时。
- 计算3个判别函数的取值符号做出最终的多分类决策：

如果$\boldsymbol{G_1}(\boldsymbol{x}) > 0$，$\boldsymbol{G}_2(\boldsymbol{x})$和$\boldsymbol{G}_3(\boldsymbol{x})$均$< 0$，则将$\boldsymbol{x}$归为$\omega_1$类；

如果$\boldsymbol{G}_2(\boldsymbol{x}) > 0$，$\boldsymbol{G}_1(\boldsymbol{x})$和$\boldsymbol{G}_3(\boldsymbol{x})$均$< 0$，则将$\boldsymbol{x}$归为$\omega_2$类；

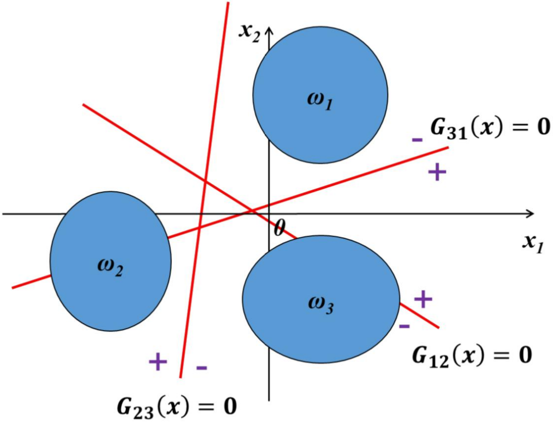
如果$\boldsymbol{G}_3(\boldsymbol{x}) > 0$，$\boldsymbol{G}_1(\boldsymbol{x})$和$\boldsymbol{G}_2(\boldsymbol{x})$均$< 0$，则将$\boldsymbol{x}$归为$\omega_3$类；

在其它的取值组合情况下，无法做出分类决策。

多分类问题的线性分类

- 一对一的训练方式:

  对于训练集中**任意两个类的实例**，训练一个线性判别函数，将这两类的实例分开。



$G_{12}(x) = 0$将ω_1类和ω_2类分开，使得
$G_{12}(x) > 0, x \in \omega_1$
$G_{12}(x) < 0, x \in \omega_2$

如何确定多分类决策呢?

一对一训练方式的分类规则

- **多数投票的原则**：如果绝大多数分类器认为将$\boldsymbol{x}$分为ω_i类，则将$\boldsymbol{x}$分为ω_i类。

- 对于3分类问题，分类规则为：

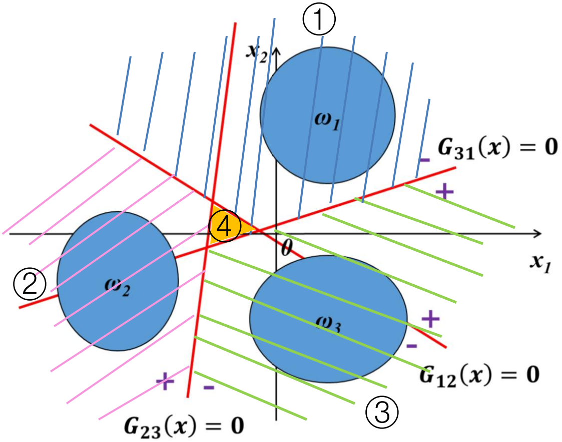
  如果$\boldsymbol{G}_{12}(\boldsymbol{x}) > 0, \boldsymbol{G}_{31}(\boldsymbol{x}) < 0$，无论$\boldsymbol{G}_{23}(\boldsymbol{x})$取值如何，则将$\boldsymbol{x}$归为$\omega_1$类；

  如果$\boldsymbol{G}_{12}(\boldsymbol{x}) < 0$，$\boldsymbol{G}_{23}(\boldsymbol{x}) > 0$，无论$\boldsymbol{G}_{31}(\boldsymbol{x})$取值如何，则将$\boldsymbol{x}$归为$\omega_2$类；

  如果$\boldsymbol{G}_{23}(\boldsymbol{x}) < 0$，$\boldsymbol{G}_{31}(\boldsymbol{x}) > 0$，无论$\boldsymbol{G}_{12}(\boldsymbol{x})$取值如何，则将$\boldsymbol{x}$归为$\omega_3$类；

  其它情况下，无法做出分类决策。

一对一训练方式对特征空间的划分



多数投票

编号	$G_{12}(x)$	$G_{23}(x)$	$G_{31}(x)$
①	>0	-	<0
②	<0	>0	-
③	-	<0	>0
④	<0	<0	<0

④是不可识别区域。

多分类问题的线性分类

- 最大值的训练方式：

  对训练集中的每个类训练一个判别函数，多分类规则：将一个实例分到取值最大的那个判别函数所对应的类中。

- 更准确的语言描述：

  一个k类的分类问题有$\omega_1,\cdots\omega_k$类的实例。每一类训练一个线性判别函数$G_i(\boldsymbol{x}), i=1,2,\cdots k$。

  若对实例$\boldsymbol{x}$进行分类时，$G_i(\boldsymbol{x})=\max\limits_{1\leq j\leq k} G_j(\boldsymbol{x})$，则将实例$\boldsymbol{x}$分为$\omega_i$类。

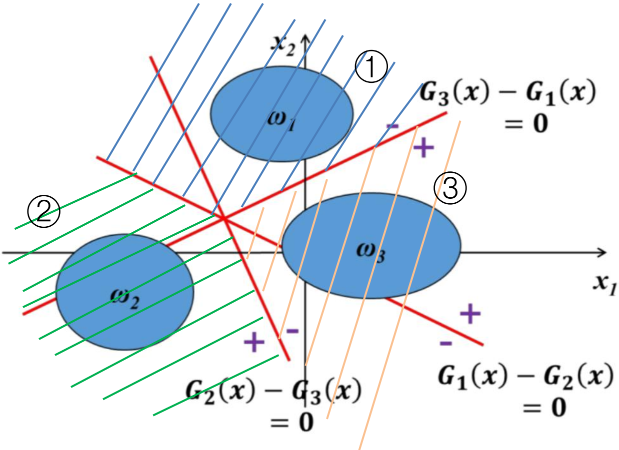
$G_i(\boldsymbol{x})=\max\limits_{1\leq j\leq k} G_j(\boldsymbol{x})$等价于：对$\forall j\neq i$，有$G_i(\boldsymbol{x})>G_j(\boldsymbol{x})$。

最大值训练方式的分类规则

- 以3分类问题为例，一个实例$\boldsymbol{x}$的分类决策规则是：
  - ✓ 如果$G_1(\boldsymbol{x}) > G_2(\boldsymbol{x})$且$G_1(\boldsymbol{x}) > G_3(\boldsymbol{x})$，则判$\boldsymbol{x}$为$\omega_1$类；
  - ✓ 如果$G_2(\boldsymbol{x}) > G_1(\boldsymbol{x})$且$G_2(\boldsymbol{x}) > G_3(\boldsymbol{x})$，则判$\boldsymbol{x}$为$\omega_2$类；
  - ✓ 如果$G_3(\boldsymbol{x}) > G_1(\boldsymbol{x})$且$G_3(\boldsymbol{x}) > G_2(\boldsymbol{x})$，则判$\boldsymbol{x}$为$\omega_3$类；
- 令$G_{12}(\boldsymbol{x}) = G_1(\boldsymbol{x}) - G_2(\boldsymbol{x})$，$G_{23}(\boldsymbol{x}) = G_2(\boldsymbol{x}) - G_3(\boldsymbol{x})$，$G_{31}(\boldsymbol{x}) = G_3(\boldsymbol{x}) - G_1(\boldsymbol{x})$，则上述分类规则变为：
  - ✓ 如果$G_{12}(\boldsymbol{x}) > 0$且$G_{31}(\boldsymbol{x}) < 0$，则判$\boldsymbol{x}$为$\omega_1$类；
  - ✓ 如果$G_{12}(\boldsymbol{x}) < 0$且$G_{23}(\boldsymbol{x}) > 0$，则判$\boldsymbol{x}$为$\omega_2$类；
  - ✓ 如果$G_{31}(\boldsymbol{x}) > 0$且$G_{23}(\boldsymbol{x}) < 0$，则判$\boldsymbol{x}$为$\omega_3$类；
- 观察发现：

  3条线性决策边界相交于一点！！！

最大值构建方式对特征空间的划分



①满足$G_1(\boldsymbol{x}) > G_2(\boldsymbol{x})$且$G_1(\boldsymbol{x}) > G_3(\boldsymbol{x})$;

②满足$G_2(\boldsymbol{x}) > G_1(\boldsymbol{x})$且$G_2(\boldsymbol{x}) > G_3(\boldsymbol{x})$;

③满足$G_3(\boldsymbol{x}) > G_1(\boldsymbol{x})$且$G_3(\boldsymbol{x}) > G_2(\boldsymbol{x})$.

多分类问题的线性分类-对比

- 对于一个k分类问题，比较3种训练方式：

一对多：需要k个判别函数，不可识别区域较多。

一对一：需要$C_k^2=k(k-1)/2$个判别函数，不可识别区域数目减小，但随着k增大所需判别函数的数量会大大增加，不适用于k较大的情况。

最大值：需要k个判别函数，没有不可识别区域，就识别效果来，性能最好，但是学习得到判别函数的过程最复杂!

小结

- 一个线性判别函数$G(x)$只能将特征空间划分为**两个**决策域：$G(x) > 0$ 和$G(x) < 0$，因此只能用于二分类问题。
- 多分类问题的线性分类可以通过构建多个线性判别函数，根据一定的逻辑关系进行分类，训练方式包括：
  - ✓ 一对多方式
  - ✓ 一对一方式
  - ✓ 最大值方式

难点