

230323 t4 p5 softmax回归

230323 t4 p5 softmax回归

1 softmax简介

回归vs分类

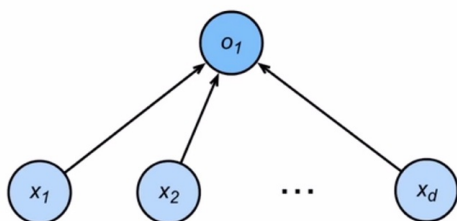
- 回归估计一个连续值
- 分类预测一个离散的类别
 - MNIST 手写数字识别 (10类)
 - ImageNet 自然物体分类 (1000类)
 - kaggle上的蛋白质分类(28) 恶意软件分类(9)

从回归到多类分类



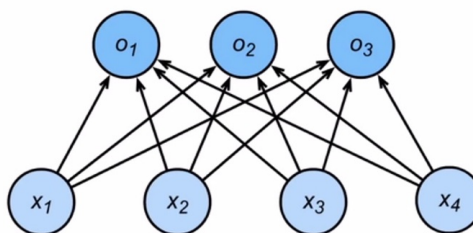
回归

- 单连续数值输出
- 自然区间 \mathbb{R}
- 跟真实值的区别作为损失



分类

- 通常多个输出
- 输出 i 是预测为第 i 类的置信度



动手学深度学习 v2 • <https://courses.d2l.ai/zh-v2>

对类别进行编码-独热编码 (one-hot encoding)

独热编码是一个向量，它的分量和类别一样多。类别对应的分量设置为1，其他所有分量设置为0。
对于一个判断是猫、狗、鸡的三分类问题，可编码为：
(1,0,0)代表猫，(0,1,0)代表狗，(0,0,1)代表鸡

对于这三个类别输出，需要与其同样多的仿射函数 (affine function)。

若上述例子的一个样本有四个输入特征，则可由如下仿射函数得到3个未被规范化为概率(非负且和为1)的预测(logit)

$$o_1 = x_1 w_{11} + x_2 w_{12} + x_3 w_{13} + x_4 w_{14} + b_1,$$

$$o_2 = x_1 w_{21} + x_2 w_{22} + x_3 w_{23} + x_4 w_{24} + b_2,$$

$$o_3 = x_1 w_{31} + x_2 w_{32} + x_3 w_{33} + x_4 w_{34} + b_3.$$

想要将预测转化为各类别的概率（置信度），需使用softmax运算

$$\text{softmax}(\mathbf{X})_{ij} = \frac{\exp(\mathbf{X}_{ij})}{\sum_k \exp(\mathbf{X}_{ik})}$$

\mathbf{X} 是由样本数* $[o_1, o_2, \dots, o_k]$ 构成的矩阵，softmax对于每一个样本，求其在各输出类别上的置信度向量

$$[\hat{y}_1, \hat{y}_2, \dots, \hat{y}_j, \dots, \hat{y}_k], \text{其中} \hat{y}_j = \frac{\exp(o_j)}{\exp(o_1) + \exp(o_2) + \dots + \exp(o_k)}$$

损失函数

L2 Loss 平方差损失函数 $l(y, y') = \frac{1}{2}(y - y')^2$

L1 Loss 绝对值损失函数 $l(y, y') = |y - y'|$

Huber's Robust Loss

在 $|y - y'| > 1$ 时采用L1函数-1/2, ≤ 1 时采用L2函数