

# Chapter3 深度学习基础

## 3.1 线性回归

### 基本要素

#### 模型定义

输出与各个输入之间是线性关系

#### 模型训练

通过数据确定模型参数值，使在数据上的误差尽可能小

训练集、样本、标签、特征

损失函数

#### 优化算法

解析解

模型和损失函数形式较为简单，误差最小化问题的解可用公式直接表达出来

数值解

有限次迭代尽可能降低损失函数的值

小批量随机梯度下降

超参数

### 表示方法

神经网络图

矢量计算表达式

尽可能采用矢量计算提升计算效率

## 3.3 线性回归的简洁实现

生成数据集

读取数据集

tensorDataset组合特征和标签

dataloader随机读取小批量

初始化模型参数

定义模型

nn.Sequrntial搭建网络

定义损失函数

定义优化算法

训练模型

## 3.5 图像分类数据集FASHION-MNIST

获取数据集

读取小批量

## 3.2 从零开始实现线性回归

生成数据集

读取数据集

初始化模型参数

定义模型

定义损失函数

定义优化算法

训练模型

## 3.4 SOFTMAX回归

分类问题

回归模型

与线性回归的不同：输出值个数等于标签类别数

输出argmax Oi

softmax将输出值变化为值为正且和为1的概率分布

softmax的形式

单样本分类的矢量计算表达式

小批量样本分类的矢量计算表达式

交叉熵损失函数

平方损失函数太过严格

交叉熵衡量两个概率分布差异

只关心对正确类别的预测概率

模型预测及评价