书面作业 2

姓名：李正扬　学号：2023010645　日期：2025 年 11 月 19 日

1 优化算法

设一元函数

$$f(x) = 5x^2 - 20x + 25, \quad x_0 = 6$$

(1) 写出 $f'(x)$ 和 $f''(x)$,，并给出最优点 x^*(使目标函数 $f(x)$ 取得最小值的点)。

$f'(x) = 10x - 20$
$f''(x) = 10 > 0$
令 $f'(x) = 0$ 得到 $x^* = 2$

(2) 采用固定步长 $\eta = 0.05$ 的梯度下降法，从 x_0 起做三步迭代，给出 x_1, x_2, x_3。

优化方向 $\Delta x = f'(x) = 10x - 20$
$f'(x_0) = 40$, $x_1 = x_0 - \eta f'(x_0) = 4$
$f'(x_1) = 20$, $x_2 = x_1 - \eta f'(x_1) = 3$
$f'(x_2) = 10$, $x_3 = x_2 - \eta f'(x_2) = 2.5$

(3) 用牛顿法 (Newton，取步长 $t = 1$) 从 x_0 起迭代两步，给出 $x_1^{(N)}$ 和 $x_2^{(N)}$。

牛顿法的优化方向 $\Delta_x = \frac{f'(x)}{f''(x)} = \frac{10x-20}{10} = x - 2$
$\Delta_0 = x_0 - 2 = 2$, $x_1^{(N)} = x_0 - t\Delta_0 = 2$
$\Delta_1 = x_1 - 2 = 0$, $x_2^{(N)} = x_1 - t\Delta_1 = 2$

(4) 比较两种方法在第 3 步 (若已收敛则取收敛的步数) 后与最优点的距离，并说明哪个更快。

固定步长的梯度下降法在第 3 步仍未收敛, 与最优点的距离 $|x_3 - x^*| = 0.5$
牛顿法第 2 步已经收敛, 与最优点的距离为 0
牛顿法的收敛速度更快, 1 步就收敛到了最优点

(5) 在现代机器学习中通常更常用哪一种方法？请说明原因。

更常用梯度下降法。原因：

(a) 牛顿法需要计算二阶导数，计算复杂度高，且 Hessian 矩阵的存储和求逆计算代价大。

(b) 牛顿法是否收敛依赖于起始点的选取，否则容易不收敛或者收敛到鞍点和极大值点。

(c) 梯度下降法可以结合动量、自适应学习率等，收敛速度也得到提升。

2 图像压缩编码

图像压缩的核心思想是在尽量保持视觉质量的前提下，减少冗余信息的存储和传输。根据课上所学知识，回答问题：

给定下列离散符号集及其概率，对信源进行霍夫曼编码，给出每个符号对应的码字，并计算平均码长、编码效率和压缩比。

$$\begin{bmatrix} x_i \\ p_i \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \\ 0.4 & 0.2 & 0.15 & 0.12 & 0.07 & 0.06 \end{bmatrix}$$

(1) 合并 x_5 和 x_6, 概率为 0.13, 重新排序, $\{x_1, x_2, x_3, x_{56}, x_4\}$

(2) 合并 x_{56} 和 x_4, 概率为 0.25, 重新排序, $\{x_1, x_{564}, x_2, x_3\}$

(3) 合并 x_2 和 x_3, 概率为 0.35, 重新排序, $\{x_1, x_{23}, x_{564}\}$

(4) 合并 x_{23} 和 x_{564}, 概率为 0.60, 重新排序, $\{x_{23564}, x_1\}$

(5) 合并 x_{23564} 和 x_1, 概率为 1.00

得到霍夫曼编码为:

x_1: 1
x_2: 000
x_3: 001
x_4: 011
x_5: 0100
x_6: 0101

平均码长 $L = \sum p_i L_i = 2.33$
信息熵 $H = -\sum p_i \log_2 p_i = 2.27\text{bit}$
编码效率 $\eta = \frac{H}{L} = 0.974$
自然二进制码长为 3
压缩比 $k = \frac{3}{L} = 1.29$

3 边缘检测

Sobel 滤波器是一种边缘检测方法，用于图像处理和计算机视觉领域，它通过计算图像灰度值的梯度来检测边缘。根据课上所学知识，回答下列题目：

(1) 写出水平/垂直 Sobel 滤波 S_x，S_y 的 3×3 矩阵形式 (省略归一化因子)。

(2) 设灰度图像

$$I = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 10 & 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 \\ 10 & 10 & 10 & 10 & 10 & 10 \end{bmatrix}$$

使用 3×3Sobel 核，对 I 进行卷积（不使用 padding，步长为 1)，分别计算

$$G_x = S_x \otimes I，G_y = S_y \otimes I.$$

给出 G_x 与 G_y 的数值矩阵；并回答：从结果看，是否成功提取出了水平边缘？请简述理由。

(1) 省略 $\frac{1}{8}$ 归一化因子

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \quad S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

(2)

$$G_x = S_x \otimes I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$G_y = S_y \otimes I = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -36 & -36 & -36 & -36 \\ -36 & -36 & -36 & -36 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

从结果看，成功提取出了水平边缘。原因：

图像亮度在输入的第 3 行与第 4 行之间发生突变，为明显的水平边缘。G_y 在对应位置 (输出的第 2 行与第 3 行) 产生了显著的负响应 (-36)，而其他非边缘区域卷积结果接近 0，说明 Sobel 垂直核 S_y 成功检测出了图像中的水平边缘。

4 分类评价指标

某模型在二分类任务 (正类为 Positive，负类为 Negative) 上的测试结果如下

Truth \ Prediction	positive	negative
positive	50	10
negative	5	35

请计算：准确率 (Accuracy)，精确率 (Precision)，召回率 (Recall)，F1 分数 (F1 Score)。

根据混淆矩阵，有 TP=50, TN=35, FP=5, FN=10，则：
准确率

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = 0.85$$

精确率

$$Precion = \frac{TP}{TP+FP} = 0.909$$

召回率

$$Recall = \frac{TP}{TP+FN} = 0.833$$

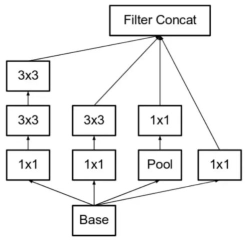
F1 分数

$$F1 \quad Score = \frac{2\times(Presion\times Recall)}{Precion+Recall} = 0.87$$

5 图像分类网络

GoogleNet 是经典的图像分类网络，从 Inception v1-v4 的迭代，反应了神经网络的发展。其中，Inception v2 引入了 BN(Batch Normalization)，同时将 Inception 模块中的 5×5 卷积变成了两个 3×3 卷积的堆叠，如图所示，计算：

(1) 假设输入维度为 $3\times3\times1$，模块内所有卷积核的步长为 1，Max-Pool 的大小为 3×3，经过一个模块后，为使得所有运算结果能直接在通道数



维度上累加, 即输出维度为 $3\times3\times4$，写出不同尺寸卷积核和 Max-Pool 的 padding 长度。

(2) 假设输入如下，假设所有卷积核的每个元素均为 1，在 (1) 的假设下计算经过一个模块后的输出结果 (4 个通道)。

1	2	1
2	4	2
1	2	1

(1) 卷积层输出尺寸公式

$$W_2 = \frac{W_1 - F + 2P}{S} + 1$$
$$H_2 = \frac{H_1 - F + 2P}{S} + 1$$

根据题目有 $W_1 = H_1 = W_2 = H_2 = 3$，$S = 1$

1×1 卷积核，$F = 1$

$$3 - 1 + 2P + 1 = 3 \Rightarrow P = 0$$

3×3 卷积核，$F = 3$

$$3 - 3 + 2P + 1 = 3 \Rightarrow P = 1$$

3×3 MaxPool

为了保持输出仍然为 3×3，同样需要 $P = 1$

(2) 四个通道分别计算如下：

(a) 通道 1 经过 1×1，3×3，3×3 三个卷积核后的输出分别为

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 9 & 12 & 9 \\ 12 & 16 & 12 \\ 9 & 12 & 9 \end{bmatrix}, \begin{bmatrix} 49 & 70 & 49 \\ 70 & 100 & 70 \\ 49 & 70 & 49 \end{bmatrix}$$

(b) 通道 2 经过 1×1，3×3 两个卷积核后的输出分别为

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 9 & 12 & 9 \\ 12 & 16 & 12 \\ 9 & 12 & 9 \end{bmatrix}$$

(c) 通道 3 经过 1×1 卷积核和 3×3 MaxPool 后的输出分别为

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 4 & 4 & 4 \\ 4 & 4 & 4 \\ 4 & 4 & 4 \end{bmatrix}$$

(d) 通道 4 经过三个 1×1 卷积核后的输出为

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$