

## 一、填空题

- 闭凸集  $X \subset R^n$ , 任意一点  $y \in R^n$ , 则点  $y$  在集合  $X$  上的投影是否是唯一的 是
- 给定问题  $\min_{x \in R^2} f(x), f(x) = -12x_2 + 4x_1^2 + 4x_2^2 + 4x_1x_2$ , 则 Hessian 矩阵  $H = \begin{bmatrix} 8 & 4 \\ 4 & 8 \end{bmatrix}$ .  
若初始搜索方向  $d_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , 则该方向的共轭方向具有的特征为  $d_2 = \begin{bmatrix} -1 & 2 \end{bmatrix}$ . 假设初始点为  $x^{(1)} = \left(-\frac{1}{2}, 1\right)^T$ , 沿  $d_1$  最小化目标函数得到迭代点  $x^{(2)} = \begin{bmatrix} -1/2 & 1 \end{bmatrix}$ . 然后从该点沿  $d_2$  方向最小化目标函数得  $x^{(3)} = \begin{bmatrix} -1 & 2 \end{bmatrix}$ .
- 目标函数  $f(x_1, x_2) = 3x_1^2 - 4x_1x_2 + x_2^2$  在点  $x^{(0)} = (0, 1)^T$  处的单位最速下降方向为:  $\begin{bmatrix} 2\sqrt{5}/5 & -\sqrt{5}/5 \end{bmatrix}$
- $\min x_1 + x_2, s.t. 2 - x_1^2 - x_2^2 \geq 0$ , 其由非线性乘子法构造的拉格朗日函数在最优点  $x^* = \begin{bmatrix} -1 & -1 \end{bmatrix}$  处对应的  $\nabla_x^2 L(x^*, \lambda^*) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ , 对应的乘子  $\lambda_1^* = 1/2$
- 在 SVM 中的 Hinge 损失和  $L_2$  损失分别如下:

$$f_i(z) = \begin{cases} \max\{0, 1 - y_i z\}, & \text{Hinge 损失} \\ \frac{1}{2} \max\{0, 1 - y_i z\}^2, & L_2 \text{ 损失} \end{cases}$$

$$\text{其梯度 } f'_i(z) = \begin{cases} (0, 1 - y_i z), & \text{Hinge 损失 } -y_i z < 1, 0(y_i z) \geq 1 \\ (0, 0), & L_2 \text{ 损失 } y_i(y_i z - 1)(y_i z < 1), 0(y_i z \geq 1) \end{cases}$$

- 外点法与内点法中, 只能对不等式约束问题进行求解是 内点 法。

## 二、计算证明题

- 给定  $u, v$ ,  $x \equiv (w_1, w_2, b_1, w_3, w_4, b_2, \omega_1, \omega_2, c)$ , 令  $h(u; x) := g(\omega_1 h_1(u; w_1, w_2, b_1) + \omega_2 h_2(u; w_3, w_4, b_2) + c) = g(\omega_1 g(w_1 u_1 + w_2 u_2 + b_1) + \omega_2 g(w_3 u_1 + w_4 u_2 + b_2) + c)$ , 这里:  $g(z) = \frac{1}{1 + \exp(-z)}$ , 并定义  $f(x) = [h(u; x) - v]^2$ 
  - 计算函数  $f$  的 Hessian 矩阵, 并证明函数  $f$  不一定是凸函数;
  - 计算函数  $f$  关于变量  $x$  的梯度  $\nabla f_x$ ;
  - 讨论如何在计算机上有效计算函数  $f$  的梯度;
  - 推导在什么条件下, 函数的梯度是 Lipschitz 连续的, 即满足:  $\|\nabla f(x_1) - \nabla f(x_2)\| \leq L\|x_1 - x_2\|, \forall x_1, x_2 \in R^9$ , 常数  $L > 0$ .
- 问题 P:  $\min f(x); s.t. Ax \leq b; Ex = e; f$  可微,  $A_{m \times n}, E_{l \times n}, x \in R^n, b \in R^m, e \in R^l$ , 证明: 设  $\bar{x}$  是上述问题的一个可行解, 且在点  $\bar{x}$  处有  $A_1 \bar{x} = b_1, A_2 \bar{x} < b_2$ , 其中  $A = [A_1 A_2]^T, b = [b_1 b_2]^T$  则
  - 向量  $d(d \in R^n, d \neq 0)$  是点  $\bar{x}$  处的可行方向的充要条件为  $A_1 d \leq 0, Ed = 0$ ;
  - 若此时  $d$  又满足:  $\nabla f(\bar{x})^T d < 0$ , 则  $d$  是一个可行下降方向.

- 考虑约束最优化问题  $\min f(x) = 4x_1^2 + 5x_1x_2 + x_2^2, s.t. \begin{cases} x_1^2 - x_2 + 2 \leq 0 \\ x_1 + x_2 - 6 \leq 0 \\ x_1, x_2 \geq 0 \end{cases}$  分别写出用罚函数法和内点函数法求解的辅助问题。

- 给定问题(P)  $\begin{cases} \min \left(x_1 - \frac{9}{4}\right)^2 + (x_2 - 2)^2 \\ s.t. \begin{cases} x_2 - x_1^2 \geq 0 \\ x_1 + x_2 \leq 6 \\ x_1, x_2 \geq 0 \end{cases} \end{cases}$ , 请验证  $x^* = \left(\frac{3}{2}, \frac{9}{4}\right)^T$  是 KKT 点, 并证明  $x^*$  是唯一

的最优解。

### 三、问答题

11. 请论述在深度学习优化算法中，是如何一步步对梯度下降法进行改进的。
12. 请论述线性规划算法在优化方法中的重要性。