凸优化第 10 周作业

1 预习作业

下节课没有小测。

2 作业题

1. (编程题) 考虑等式约束优化问题:

$$\begin{aligned} \min \quad & \frac{1}{2}x^\top Px + q^\top x \\ \text{s.t.} \quad & Ax = b, \end{aligned}$$

   其中 $x \in \mathbb{R}^n, \quad P \in \mathbb{S}_+^n, \quad q \in \mathbb{R}^n, \quad A \in \mathbb{R}^{m\times n}, \quad b \in \mathbb{R}^m$。

   使用牛顿方法求解上述优化问题, 采用回溯直线搜索方法, 自己设定参数和初始点. 停止准则为牛顿减量 $\lambda(x)^2 < 10^{-5}$. 求出原问题和对偶问题的最优解, 函数最优值。分别画出对数误差 $\log(f(x_k) - p^*)$ 和下降步长 t_k 关于迭代次数 k 的图像。

   请使用 Q1_data 文件夹中提供的数据求解以上问题。我们给出了 $m = 100, n = 200$ 时对应的矩阵 P, q, A, b.

2. (编程题) 分别用障碍函数法和原对偶内点法求解下述二次规划问题:

$$\begin{aligned} \min \ & \frac{1}{2}x^\top Px + q^\top x \\ \text{s.t.} \ & Ax = b \\ & x \succeq 0, \end{aligned}$$

   其中 $x \in \mathbb{R}^n$, $P \in \mathbb{S}_+^n$, $q \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$.

   (a) 障碍函数法要求:

      i. 阈值误差 $\varepsilon = 10^{-8}$。

      ii. 请画出对数对偶间隙 $\log \frac{n}{t}$ 与 Newton 迭代次数 k 之间的关系图。

      iii. 给出原对偶最优解 x^*, λ^*, v^* 和最优值 p^*。

      障碍函数法中参数 μ 建议选取 $\mu = 10$ 或者自行选取。

   (b) 原对偶内点法要求:

      i. 原误差 $\|r_{\text{pri}}\|_2 \le 10^{-8}$，对偶误差 $\|r_{\text{data}}\|_2 \le 10^{-8}$，代理对偶间隙 $\hat{\eta} \le 10^{-8}$。

ii. 分别画出 $\log\hat{\eta}$ 和 $\log\left\{(\|r_{\text{pri}}\|_2^2+\|r_{\text{dual}}\|_2^2)^{\frac{1}{2}}\right\}$ 与 Newton 迭代次数 k 之间的关系图。

iii. 给出原对偶最优解 x^*,λ^*,v^* 和最优值 p^*。

请使用 Q2_data 文件夹中的数据求解以上两个问题。我们给出了 $m=100, n=200$ 时对应的矩阵 P,q,A,b，以及初始点 x_0，λ，v。

3. (编程题) Consider the following minimization problem:

$$\min f(x_1,x_2)=\frac{1}{2}\left(x_1^2+100x_2^2\right).$$

Suppose the starting point is $x^{(0)}=(100,1)^\top$ and we are using the negative gradient as our descent direction.

Consider the **Heavy ball Method** with $\alpha=4/121$ and $\beta=81/121$, then

(a) plot the corresponding $x^{(k)}$ on the 2D plane and $f(x^{(k)})$ vs k using semi-log plot.

(b) and compare the convergence rate of the Heavy Ball Method to that of standard gradient method by plotting the semi-log plot of $f(x^{(k)})$ vs k.

The algorithms stops when the 2-norm of gradient is less than 10^{-8}.

4. (选做题) Let $a_i\in\mathbb{R}^n$, $i=1,2,\cdots,m$, and assume that

$$A=\begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix}$$

has rank n. Define $\phi_i(x)\triangleq-\log(a_i^\top x-b_i)$, $i=1,2,\cdots,m$. Assume that $D=\{x\in\mathbb{R}^n|Ax-b>0\}$ is non-empty. Define $\phi(x)=\sum_{i=1}^m\phi_i(x)$.

(a) For any $x\in D$, prove that

$$(\nabla\phi(x))^\top(\nabla^2\phi(x))^{-1}\nabla\phi(x)\le m.$$

(b) Let $f_t(x)\triangleq tc^\top x+\phi(x)$, where $c\in\mathbb{R}^n$, and x^* be a minimizer of $f_t(x)$. Let $\mu>1$. Prove that

$$\sqrt{(\nabla f_{\mu t}(x^*))^\top(\nabla^2 f_{\mu t}(x^*))^{-1}\nabla f_{\mu t}(x^*)}\le(\mu-1)\sqrt{m}.$$

(Hint: Use the fact $\begin{bmatrix} vv^\top & v \\ v^\top & 1\end{bmatrix}\succeq 0$ and Schur's complement.)

3 作业说明

1. 编程作业需要撰写报告（包含推导步骤和程序运行结果）。请将报告（pdf 电子版）和代码（编程语言不限）一起打包提交至网络学堂。

2. **请大家务必在截止时间之前提交作业**，迟交一周以内的作业得分是卷面分的 50%，迟交超过一周的作业不得分。

3. 每次作业的满分是 25 分，做选做题有额外加分，但每次作业总分不超过 25 分。