优化方法作业

计试 61 张翀 2140506063

Week 1

9.4 周二

作业 1

证明：$\forall x_1, x_2 \in P, \theta \in [0,1]$，记 $x_\theta = \theta x_1 + (1-\theta)x_2$，因为 $x_1, x_2 \in \mathbb{R}^n$，有 $x_\theta \in \mathbb{R}^n$.

此时

$$Ax_\theta = A(\theta x_1 + (1-\theta)x_2) = \theta Ax_1 + (1-\theta)Ax_2 \leq \theta b + (1-\theta)b = b$$

且有

$$Bx_\theta = B(\theta x_1 + (1-\theta)x_2) = \theta Bx_1 + (1-\theta)Bx_2 = \theta d + (1-\theta)d = d$$

因此 $x_\theta \in P$. 所以 P 为凸集。

作业 2-1

证明：$\forall x, y \in \mathbb{R}^n, \theta \in [0,1]$，记 $x_\theta = \theta x + (1-\theta)y$.

此时

$$
\begin{aligned}
f(x_\theta) &= \max_{1\le k\le n}(\theta x_k + (1-\theta)y_k) \\
&= \sup_{1\le k\le n}(\theta x_k + (1-\theta)y_k) \\
&\le \sup_{1\le i\le n}(\theta x_i) + \sup_{1\le j\le n}((1-\theta)y_j) \\
&= \theta \max_{1\le i\le n}(x_i) + (1-\theta)\max_{1\le j\le n}(y_j) \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}
$$

$f(x)$ 符合凸函数的定义。

作业 2-2

证明：当且仅当 P 为半正定矩阵时，$f(x)=\frac{1}{2}x^TPx+q^Tx+r$ 为凸函数；P 为正定矩阵时，$f(x)$ 为严格凸函数。此处只给出第一条结论的证明，第二条结论的证明是类似的。

固定 x，并记 $\Delta f(\delta x)=f(x+\delta x)-f(x)$；根据定义，$f(x)$ 为凸函数当且仅当 $\forall\theta\in[0,1],\delta x,\theta\delta x\in\{\delta x:x+\delta x\in domx\},\theta\Delta f(\delta x)-\Delta(\theta\delta x)\ge 0.$

考虑到

$$
\begin{aligned}
\Delta(\delta x) &= f(x+\delta x)-f(x) \\
&= (\frac{1}{2}(x+\delta x)^TP(x+\delta x)+q^T(x+\delta x)+r)-(\frac{1}{2}x^TPx+q^Tx+r) \\
&= \frac{1}{2}(\delta x^TPx+x^TP\delta x+\delta x^TP\delta x)+q^T\delta x \\
&= \frac{1}{2}\delta x^TP\delta x+\frac{1}{2}(\delta x^TPx+x^TP\delta x)+q^T\delta x
\end{aligned}
$$

有

$$\theta\Delta f(\delta x)-\Delta(\theta\delta x)=\frac{1}{2}\theta(1-\theta)\delta x^TP\delta x$$

因为 $\theta\in[0,1]$，所以

$$\theta\Delta f(\delta x)-\Delta(\theta\delta x)\ge 0 \Leftrightarrow \delta x^TP\delta x\ge 0,\forall\delta x\in\{\delta x:x+\delta x\in domx\}$$

考虑到 $domx = \mathbb{R}^n$，所以 $f(x)$ 为凸函数等价于 $\delta x^T P \delta x \geq 0, \forall \delta x \in \mathbb{R}^n$. 右侧的叙述即为 P 为半正定矩阵的定义。

类似地，也可证明 P 为半负定时二次函数为凹函数，P 为负定时二次函数严格凹。

9.6 周四

作业 1

证明：二阶可微的 $f(x)$ 为凸函数，当且仅当 Hessian 矩阵为半正定矩阵。

在以下的证明过程中，我们讨论一种比较简单的情况：$f(x)$ 在 C 上二阶连续可微，其中 C 是开集（任一元素为内点）且为凸集。此时，$\forall x \in C, \exists \delta x > 0, B(x, \delta x) \subset C$，且有

$$f(x+\Delta x) = f(x) + \nabla^T f(x)\Delta x + \frac{1}{2}\Delta x^T \nabla^2 f(\xi)\Delta x, 0 \leq |\xi - x| \leq \Delta x \leq \delta x(*)$$

必要性：

若 $f(x)$ 为凸函数，则固定 x 和 δx.$\forall |\Delta x| < \delta x$，利用 (*) 式有

$$\theta f(x+\Delta x) + (1-\theta)f(x) - f(x+\theta\Delta x) = \frac{\theta}{2}\Delta x^T \nabla^2 f(\xi)\Delta x - \frac{\theta^2}{2}\Delta x^T \nabla^2 f(\xi_\theta)\Delta x \geq 0$$

$$\forall \theta \in [0,1], \xi \in B(x, \Delta x), \xi_\theta \in B(x, \theta\Delta x)$$

取 $\theta = 1$ 可知 $\Delta x^T \nabla^2 f(\xi)\Delta x - \Delta x^T \nabla^2 f(\xi_\theta)\Delta x \geq 0$,而 $\lim\limits_{\Delta x \to 0} \Delta x^T \nabla^2 f(\xi)\Delta x - \Delta x^T \nabla^2 f(\xi_\theta)\Delta x = 0$. 因此，固定 θ 并取 $(\frac{\theta}{2} - \frac{\theta^2}{2})\Delta x^T \nabla^2 f(\xi_\theta)\Delta x$，此式恒取非负值；因为 $\lim\limits_{\Delta x \to 0} \Delta x^T \nabla^2 f(\xi_\theta)\Delta x = \Delta x^T \nabla^2 f(x)\Delta x$，所以由 θ 的任意性知 $\exists B(0, \delta x), \forall |\Delta x| \leq \delta x, \Delta x^T \nabla^2 f(x)\Delta x \geq 0$，即每个点都存在使 Hessian 矩阵半正定的邻域。由 C 是开集知 Hessian 矩阵在 C 上是半正定的。

充分性：

Hessian 的半正定可以保证 (*) 式右侧最后一项非负，因此 $f(x+\Delta x) \geq f(x) + \nabla^T f(x)\Delta x$，满足凸函数的一阶判定条件，所以 $f(x)$ 为凸函数。

作业 2

通过计算特征值易知 P 的正定性，所以优化函数是凸函数，且在可行解集连续可导。另外有 $\nabla f(x) = Px + q$. 对于最优解 $x^* = [1, 1/2, -1]^T$，$\nabla f(x^*) = [-1, 0, 2]^T$，有 $\forall y \in [-1, 1]^3, \nabla^T f(x^*)(y - x^*) = y_1 + 2y_3 + 3 \geq 0$，因此 $f(y) \geq f(x), y \in [-1, 1]^3$，所以 x^* 是此问题的最优解。