最优化 Homework2

学号	姓名
20337025	崔璨明

证明 proximal gradient descent algorithm Convergence in O(1/ε) steps 。

证明过程：

证明：设函数 $f(x) = g(x) + h(x)$

$g(x)$是凸函数且可微，$h(x)$是凸函数

由近端梯度下降：$x_{t+1} = prox_{h,\gamma}(x_t - \gamma \nabla g(x_t))$

则 generalized gradient：

$$G_\gamma(x) = \frac{1}{\gamma}(x - prox_{h,\gamma}(x - \gamma \nabla g(x)))$$

∵ $g(x)$是凸函数

∴ $\|\nabla g(x) - \nabla g(y)\| \leq L\|x-y\|,\ \forall x, y$ ……①

又∵ Quadraic Upper Bound Lemma

∴ $g(y) \leq g(x) + \nabla g(x)^T(y-x) + \frac{L}{2}\|y-x\|^2$ ……②

令 $y = x - \gamma G_\gamma(x)$，得

$$g(x - \gamma G_\gamma(x)) \leq g(x) - \nabla g(x)^T G_\gamma(x) + \frac{\gamma^2 L}{2}\|G_\gamma(x)\|^2$$ ……③

使 $0 < \gamma \leq \frac{1}{L}$，有

$$g(x - \gamma G_\gamma(x)) \leq g(x) - \nabla g(x)^T G_\gamma(x) + \frac{\gamma}{2}\|G_\gamma(x)\|^2$$ ……④

又∵ $g(x)$和$h(x)$都为凸函数，∴对 $\forall z \in dom f$，有：

$h(z) \geq h(x - \gamma G_\gamma(x)) + (G_\gamma(x - \nabla f(x)))^T(z - x + \gamma G_\gamma(x))$ ……⑤

$g(z) \geq g(x) + \nabla f(x)^T(z - x)$ ……⑥

④+⑤+⑥，得：

$$f(x - \gamma G_\gamma(x)) \leq f(z) + G_\gamma(x)^T(x - z) - \frac{\gamma}{2}\|G_\gamma(x)\|^2$$ ……⑦

$\therefore$ 在每一次迭代中 $x' = x - \gamma G_\gamma(x)$

令 $z = x^*$，即最优解，有：

$$f(x') - f^* \le G_\gamma(x)^T(x - x^*) - \frac{\gamma}{2}\|G_\gamma(x)\|^2 \quad \cdots ⑧$$

$$= \frac{1}{2\gamma}(\|x - x^*\|^2 - \|x - x^* - \gamma G_\gamma(x)\|^2) \quad \cdots ⑨$$

$$= \frac{1}{2\gamma}(\|x - x^*\|^2 - \|x' - x\|^2) \quad \cdots ⑩$$

令 $x = x_{i-1}$，$x' = x_i$，$i = 1, 2, \dots, k$，有：

$$\sum_{i=1}^{k}(f(x_k) - f^*) \le \frac{1}{2\gamma}\sum_{i=1}^{k}(\|x_{i-1} - x^*\|^2 - \|x_i - x^*\|^2) \cdots ⑪$$

$$= \frac{1}{2\gamma}(\|x_0 - x^*\|^2 - \|x_k - x^*\|^2) \quad \cdots ⑫$$

$$\le \frac{1}{2\gamma}\|x_0 - x^*\|^2 \quad \cdots ⑬$$

由(7)，得 $f(x') \le f(x) - \frac{t}{2}\|G_\gamma(x)\|^2$

$\therefore$ $f(x_i)$ 是单调递减的。

$\therefore$

$$f(x_k) - f^* \le \frac{1}{k}\sum_{i=1}^{k}(f(x_i - f^*)) \le \frac{1}{2k\gamma}\|x_0 - x^*\|^2$$

$\therefore$ 当给定精度 ε 时，近端梯度下降将在 k 次迭代后收敛到 $O(\frac{1}{\varepsilon})$