No.

Date

Assignment 9

1. (a)

证明: ∵ f 是 R^n 上的连续可导凸函数

对一固定的 $x \in R^n$, $g_x(y) = f(y) - \nabla f(x)^T y$

∴ g_x 也是 R^n 上的连续可导函数

∵ $g_x(y) = f(y) + (-\nabla f(x)^T y)$, $\forall y \in R^n$

即 g_x 可表示为两个凸函数之和

∴ g_x 也是凸函数

∵ $\nabla g_x(y) = \nabla f(y) - \nabla f(x)$, $\forall y \in R^n$

∴ $\nabla g_x(x) = 0$, 又∵ g_x 是凸的

∴ x is a minimizer of g_x over R^n

(b) $\forall y, z \in R^n$

$\|\nabla g_x(y) - \nabla g_x(z)\| = \|\nabla f(y) - \nabla f(z)\|$

$\leq L\|y - z\|$ (∇f 是 L-Lipschitz 连续的)

∴ ∇g_x 是 L-Lipschitz 连续的

∴ 由二次上界可知

对 $\forall y, z \in R^n$

$g_x(z) \leq g_x(y) + \nabla g_x(y)^T(z-y) + \frac{L}{2}\|y-z\|^2$

令 $z = y - \frac{1}{L}\nabla g_x(y)$

则 $g_x(z) \leq g_x(y) - \frac{1}{2L}\|\nabla g_x(y)\|^2$

由(a)可知, $g_x(x) \leq g_x(z)$

∴ $g_x(x) \leq g_x(y) - \frac{1}{2L}\|\nabla g_x(y)\|^2$, $\forall x, y \in R^n$, 证毕

(c) 由(b)可知，对 $\forall x, y \in R^n$：$g_x(x) \le g_x(y) - \frac{1}{2L}\|\nabla g_x(y)\|^2$

即 $f(x) - \nabla f(x)^T x \le f(y) - \nabla f(x)^T y - \frac{1}{2L}\|\nabla f(y) - \nabla f(x)\|^2$

即 $f(x) + \nabla f(x)^T(y-x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|^2 \le f(y)$

∴ 证毕

2.

(a) $(F(x) - F(y))^T(x-y) \ge 0, \ \forall x, y$

$\Leftrightarrow \quad (x-y)^T A (x-y) \ge 0, \ \forall x, y$

$\Leftrightarrow \quad A \succeq 0$（A半正定，即A的所有特征值都非负）

(b) $(F(x) - F(y))^T(x-y) > 0, \ \forall x, y \ (x \ne y)$

$\Leftrightarrow \quad (x-y)^T A (x-y) > 0, \ \forall x, y \ (x \ne y)$

$\Leftrightarrow \quad A \succ 0$（A正定，即A的所有特征值都为正）

(c) $(F(x) - F(y))^T(x-y) \ge m\|x-y\|_2^2, \ \forall x, y$

$\Leftrightarrow \quad (x-y)^T A (x-y) \ge m(x-y)^T(x-y), \ \forall x, y$

$\Leftrightarrow \quad (x-y)^T(A - mI)(x-y) \ge 0, \ \forall x, y$

$\Leftrightarrow \quad A - mI \succeq 0$（$A-mI$半正定，即A的所有特征值都大于等于m）

(d) $\|F(x) - F(y)\|_2 \le L\|x-y\|_2, \ \forall x, y$

$\Leftrightarrow \quad \|A(x-y)\|_2 \le L\|x-y\|_2, \ \forall x, y$

$\Leftrightarrow \quad \|A(x-y)\|_2 \le L\|x-y\|_2, \ \forall x, y \ (x \ne y)$

No.
Date

接上页： $\Leftrightarrow \quad \dfrac{\|A(x-y)\|_2}{\|x-y\|_2} \le L, \quad \forall x,y\ (x\neq y)$

$\Leftrightarrow \quad \|A\|_2 \le L$ （即A的最大奇异值小于等于L）

(e) $\quad (F(x)-F(y))^T(x-y) \ge \frac{1}{L}\|F(x)-F(y)\|_2^2, \ \forall x,y$

$\Leftrightarrow \quad (x-y)^T A(x-y) \ge \frac{1}{L}(x-y)^T A^2 (x-y), \ \forall x,y$

$\Leftrightarrow \quad (x-y)^T(A-\frac{1}{L}A^2)(x-y) \ge 0, \ \forall x,y$

$\Leftrightarrow \quad A-\frac{1}{L}A^2 \succeq 0$ （$A-\frac{1}{L}A^2$半正定）

设λ是A的任一特征值，$x\in R^n$ ~~且~~ $(x\neq 0)$是对应特征向量

则 $Ax=\lambda x$

$$(A-\tfrac{1}{L}A^2)x = Ax-\tfrac{1}{L}A^2x = \lambda x-\tfrac{1}{L}A(\lambda x)$$
$$= \lambda x-\tfrac{\lambda}{L}\cdot\lambda x$$
$$= (\lambda-\tfrac{\lambda^2}{L})x$$

$\therefore \ \lambda-\frac{\lambda^2}{L}$ 是 $A-\frac{1}{L}A^2$ 的特征值

$A-\frac{1}{L}A^2$半正定 $\Leftrightarrow$ $A-\frac{1}{L}A^2$的所有特征值都非负

$\Leftrightarrow \ \lambda-\frac{\lambda^2}{L}\ge 0 \ \Leftrightarrow \ 0\le\lambda\le L$，即A的所有特征值都属于$[0, L]$

A+