（装订线内不要答题）

# 复旦大学大数据学院
## 2020年春季学期课程期末考试卷

☒A 卷　　□B 卷　　□C 卷

**课程名称：最优化方法**
**课程代码：DATA130026.01**
**开课院系：大数据学院 考试形式：闭卷**

姓　名：__________　学　号：_________　专　业：______________

声明：我已知悉学校对于考试纪律的严肃规定，将秉持诚实守信宗旨，严守考试纪律，不作弊，不剽窃；若有违反学校考试纪律的行为，自愿接受学校严肃处理。

学生（签名）：_____________

年　　月　　日

| 题　目 | 1 | 2 | 3 | 4 | 5 | 6 | 总　分 |
|---|---|---|---|---|---|---|---|
| 得　分 | | | | | | | |

1. (20 points) Please answer true or false. (You may use the notation "T" for "true" and "F" for "false".) No explanation is needed. A correct answer is worth 2 points, no answer 0 points, a wrong answer $-1$ points.

   (1) Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous differentiable real valued function. Then $f$ is convex if and only if $f(y) \geq f(x) + \nabla f(x)^T(y-x)$ holds for $x, y \in \text{dom}(f)$

   (2) Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is a continuous differentiable real valued function. Then $x$ is a global minimizer of $f(x)$ if $\nabla f(x) = 0$.

   (3) Suppose $C$ is a closed and convex set. Then the subdifferential of the indicator function

   $$I_C(x) := \begin{cases} 0, & \text{if } x \in C \\ \infty, & \text{otherwise,} \end{cases}$$

   is always equivalent to the normal cone

   $$N_C(x) = \{g \in \mathbb{R}^n : g^T x \geq g^T y, \forall y \in C\}.$$

   (4) The Lagrangian dual of a quadratically constrained quadratic programming (QCQP) problem is equivalent to the Lagrangian dual of the semidefinite relaxation of the same QCQP problem.



   (5) A convex QCQP problem has the same optimal value (assuming it exists) with its SDP relaxation if the Slater condition holds.

   (6) The set $\{(x,y) : x > 0, y > 0, xy > 1\}$ is nonconvex.

   (7) Newton's method may not converge for unconstrained convex optimization problems (assuming the objective function is twice-continuously differentiable and its Hessian is Lipschitz continuous).

   (8) The function $f(x) = -\sqrt{x}$ with $\text{dom}(f) := \{z : z \geq 0, z \in \mathbb{R}\}$, is not subdifferentiable at $x = 0$.

   (9) For a nonlinear optimization problem, if the gradient descent method converges, then it converges to a local minimum.

   (10) The feasible set of the standard semidefinite program may be a nonconvex set.

2. (20 points)

   (1) (7 points) Write down the subdifferential of

   $$f(x) := \|Ax + b\|_2,$$

   where $\text{dom}(f) = \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

   (2) (8 points) Let $A \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^n$ be given. Consider the following problem:

   $$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax \geq 0. \end{aligned}$$

   Show that the optimal value of the above problem is 0 if and only if there exists an $y \geq 0$ such that $A^T y = c$.

   (3) (7 points) Suppose you want to compute the maximal eigenvalue of a symmetric matrix $A \in \mathbb{R}^{n \times n}$. Write down an SDP problem for this target. (You only need to write down the formulation. Do **NOT** solve the problem.)

3. (20 points) Suppose you are using the proximal gradient method to solve the following problem,

   $$\min f(x) := g(x) + h(x)$$

   where $g(x) = x_1^2 + 2x_1x_2 + x_2^2 - 2(x_1 + x_2)$, and $h(x) = |x_1|$. Use $x_0 = (0,0)$ as the initial point.

   (1) (15 points) Now suppose you are using the following line search rule in your method: find the smallest nonnegative integer $s$ such that

   $$g(y) \leq g(x) + \nabla g(x)^T(y-x) + \frac{1}{2t}\|y-x\|^2$$

   where $y = \text{prox}_{th}(x - t\nabla g(x))$, and $t = \beta^s \hat{t}$ (by setting $\hat{t} = 1$, $\beta = 0.5$). Write down the proximal gradient method iteration for computing $x_1$. You need to write the both the value and calculation of $x_1$.

(装订线内不要答题)

(2) (5 points) Show that if the $x_1$ computed in (a) is an optimal solution or not. Write down your derivation.

4. (20 points ) Consider the following linear program, with bounds and a single linear equality constraint:

$$\min_x -\sum_{i=1}^{2020} c_i x_i \quad \text{s.t.} \sum_{i=1}^{2020} a_i x_i = b,\ 0 \le x_i \le u_i,\ i = 1, 2, \ldots, 2020,$$

where $c_i, a_i, u_i$ $(u_i > 0)$, $i = 1, \ldots, 2020$ and $b$ are given constants.

(a) (10 points) Write down the KKT optimality conditions for this problem.

(b) (10 points) Assume that $a_i = 1$ for all $i$, and that the variables $c_i$ are ordered such that

$$c_1 > c_2 > \ldots > c_{2020}.$$

Suppose further that

$$\sum_{i=1}^{2000} u_i + \frac{1}{2} u_{2001} = b.$$

Using this information, find the primal solution $x$ and the Lagrange multiplier vectors that satisfy the KKT conditions.

5. (20 points) Consider the convex optimization problem

$$\min_{x \in X} f(x),$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable (on $\mathbb{R}^n$) function which is convex on the set $X$; the set $X \subset \mathbb{R}^n$ is convex and non-empty. We suppose that the set $X$ is compact so that the problem is guaranteed to have a non-empty set of optimal solutions, denoted by $X^*$.

Prove that if $x^*$ and $\hat{x}$ (assuming $x^* \neq \hat{x}$) both are optimal solutions (that is, are in the set $X^*$), then we have the following two results:

(1) (10 points) $\nabla f(x^*)^T(\hat{x} - x^*) = \nabla f(\hat{x})^T(\hat{x} - x^*) = 0$; (Hint: Recall the optimality condition $\nabla f(x^*)^T(y - x^*) \ge 0,\ \forall y \in X$.)

(2) (10 points) $\nabla f(x^*) = \nabla f(\hat{x})$ holds. (Hint: If $f$ is continuously differentiable on $\mathbb{R}^n$, then the subdifferential $\partial f(x) = \{\nabla f(x)\}$ for all $x \in \mathbb{R}^n$.)