姓名：____________

学号：____________

部(院)：____________

______级______班

装

订

)

大 连 理 工 大 学

课程名称：优化方法　　试卷：C　考试形式：闭卷

授课部（院）：数学　考试日期：　年　月　日　试卷共 6 页

	一	二	三	四	五	六					总分
标准分	30	15	15	15	15	10					100
得　分											

得分	

一、　叙述定义：

(1) 精确线搜索，Armijo 线搜索，Goldstein 原则，Wolfe-Powell 原则和强 Wolfe-Powell 原则(10 分)；极点，基本可行解，相对成本向量

(2) 共轭函数，临近映射(proximal mapping)，Moreau 分解，投影(8 分)；

(3) 切锥，线性化可行方向锥，LICQ，MFCQ，临界锥，积极约束指标集(12 分)

(1) 精确线搜索：设 $f(x)$ 连续可微，给定 x^k 和 d^k 满足 $\nabla f(x^k)\neq 0, \nabla f(x^k)^T d^k<0$

如何选择步长 α_k　　$\alpha_k=\arg\min\{\varphi(\alpha)=f(x^k+\alpha d^k)\}$

Armijo 线搜索：取步长 $\alpha_k=\bar{\alpha}\rho^{m_k}$，$m_k$ 是满足下式的第一个非负整数 m：

$f(x^k+\bar{\alpha}\rho^m d^k)-f(x^k)\leq c\bar{\alpha}\rho^m\nabla f(x^k)^T d^k$　其中 $\bar{\alpha}>0, \rho\in(0,1), c\in(0,1)$

Goldstein 原则　$(1-\rho)\alpha_k\nabla f(x^k)^T d^k\leq f(x^k+\alpha_k d^k)-f(x^k)\leq\rho\alpha_k\nabla f(x^k)^T d^k$，$\rho\in(0,\frac{1}{2})$

Wolfe-Powell 原则　$f(x^k+\alpha_k d^k)-f(x^k)\leq\rho\alpha_k\nabla f(x^k)^T d^k$

$\nabla f(x^k+\alpha_k d^k)^T d^k\geq\sigma\nabla f(x^k)^T d^k$　　$\rho\in(0,\frac{1}{2})$　$\sigma\in(\rho,1)$

强 Wolfe-Powell 原则　$f(x^k+\alpha_k d^k)-f(x^k)\leq\rho\alpha_k\nabla f(x^k)^T d^k$

$|\nabla f(x^k+\alpha_k d^k)^T d^k|\leq-\sigma\nabla f(x^k)^T d^k$　$\rho\in(0,\frac{1}{2})$　$\sigma\in(\rho,1)$

(2) 函数 $f:R^n\to R$ 的共轭函数 f^*：$f^*(y)=\sup_x\{y^T x-f(x)\}$

函数 $f:R^n\to\bar{R}$ 的 proximal mapping　$\mathrm{prox}_{tf}(x)=\arg\min_{u\in R^n}(tf(u)+\frac{1}{2}\|u-x\|^2)$ 其中 $t>0$

Moreau 分解　$x=\mathrm{prox}_{tf}(x)+t\,\mathrm{prox}_{f^*/t}(x/t)$

投影

(3) $d\in R^n$ 称为 Ω 在 x 处的切方向，若存在 $\{z^k\}$，$z^k\in\Omega$，$z^k\to x$ 和 $\{t_k\}$，$t_k>0$，$t_k\to 0$，使 $d=\lim_{k\to\infty}\frac{z^k-x}{t_k}$．$\Omega$ 在 x 处所有切方向的集合称为切锥，记为 $T_\Omega(x)$．

$x\in\Omega$ 处的线性化可行方向集合为 $\mathcal{F}(x)=\{d\mid d^T\nabla g_i(x)=0, i\in\mathcal{E}, d^T\nabla g_i(x)\geq 0, i\in A(x)\cap\mathcal{I}\}$

其中 $A(x)=\{i\mid g_i(x)=0, i\in\mathcal{E}\cup\mathcal{I}\}$ 是 x 处的积极约束指标集．

若 $\{\nabla g_i(x), i\in A(x)\}$ 线性无关，称线性无关约束规范(LICQ)在 $x\in\Omega$ 处成立

称 MFCQ 在 x^* 处成立，若 $w\in R^n$ 使 $\nabla g_i(x^*)^T w>0, i\in A(x^*)\cap\mathcal{I}, \nabla g_i(x^*)^T w=0, i\in\mathcal{E}$

并且 $\{\nabla g_i(x^*), i\in\mathcal{E}\}$ 线性无关．

临界锥 $C(x^*,\lambda^*)=\{w\in\mathcal{F}(x^*)\mid\nabla g_i(x^*)^T w=0$，$\forall i\in A(x^*)\cap\mathcal{I}, \lambda_i^*>0\}$

$\forall i\in\{i\in A(x^*)\cap\mathcal{I}\mid\lambda_i^*>0\}$

得分	

二、考虑线性规划问题

$$
\begin{aligned}
\min \quad & 2x_1 + 5x_2 + 6x_3 \\
\text{s.t.} \quad & 2x_1 + x_2 + 2x_3 \geq 3 \\
& x_1 + 2x_2 + 2x_3 \geq 1 \\
& x_1 + 3x_2 + x_3 \geq 3 \\
& x > 0
\end{aligned}
$$

（1） 写出上述问题的对偶问题（7 分）；

（2） 利用单纯形法求解对偶问题（8 分）。

(1) $\min \ 2X_1 + 5X_2 + 6X_3$

s.t. $2X_1 + X_2 + 2X_3 - X_4 = 3$

$X_1 + 2X_2 + 2X_3 - X_5 = 1$

$X_1 + 3X_2 + X_3 - X_6 = 3$

$X_1, X_2, X_3, X_4, X_5, X_6 > 0$

$C = (2 \ 5 \ 6 \ 0 \ 0 \ 0)^T$

$$A = \begin{pmatrix} 2 & 1 & 2 & -1 & 0 & 0 \\ 1 & 2 & 2 & 0 & -1 & 0 \\ 1 & 3 & 1 & 0 & 0 & -1 \end{pmatrix}$$

$b = (3 \ 1 \ 3)^T$

$\xrightarrow{\text{dual}} \begin{cases} \max & b^T\lambda \\ \text{s.t.} & A^T\lambda \leq C \end{cases}$

$$\begin{cases} \max & 3\lambda_1 + \lambda_2 + 3\lambda_2 \\ \text{s.t.} & 2\lambda_1 + \lambda_2 + \lambda_3 \leq 2 \\ & \lambda_1 + 2\lambda_2 + 3\lambda_3 \leq 5 \\ & 2\lambda_1 + 2\lambda_2 + \lambda_3 \leq 6 \\ & \lambda_1 \geq 0 \\ & \lambda_2 \geq 0 \\ & \lambda_3 \geq 0 \end{cases}$$

(2) $\min \ -3\lambda_1 - \lambda_2 - 3\lambda_2$

s.t. $2\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 2$

$\lambda_1 + 2\lambda_2 + 3\lambda_3 + \lambda_5 = 5$

$2\lambda_1 + 2\lambda_2 + \lambda_3 + \lambda_6 = 6$

$\lambda \geq 0$

$$A = \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 0 & 1 & 0 \\ 2 & 2 & 1 & 0 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}$$

$C = (-3 \ -1 \ -3 \ 0 \ 0 \ 0)^T \qquad B^1 = \{4, 5, 6\} \quad C_{B^1} = (0, 0, 0)^T$

$$M(X^1, B^1) = \left(\begin{array}{cccccc|c} \boxed{2} & 1 & 1 & 1 & 0 & 0 & 2 \\ 1 & 2 & 3 & 0 & 1 & 0 & 5 \\ 2 & 2 & 1 & 0 & 0 & 1 & 6 \\ \hline -3 & -1 & -3 & 0 & 0 & 0 & 0 \end{array}\right) \begin{array}{c} 1 \\ 5 \\ 3 \\ \end{array}$$

$$M(X^2, B^2) = \left(\begin{array}{cccccc|c} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 1 \\ 0 & \frac{3}{2} & \boxed{\frac{5}{2}} & -\frac{1}{2} & 1 & 0 & 4 \\ 0 & 1 & 0 & 1 & 0 & 1 & 4 \\ \hline 0 & \frac{1}{2} & -\frac{3}{2} & \frac{3}{2} & 0 & 0 & 3 \end{array}\right) \begin{array}{c} 2 \\ \frac{8}{5} \\ \\ \end{array} \qquad B^2 = \{1, 5, 6\}$$

$$M(X^3, B^3) = \left(\begin{array}{cccccc|c} 1 & \frac{1}{5} & 0 & \frac{3}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ 0 & \frac{3}{5} & 1 & -\frac{1}{5} & \frac{2}{5} & 0 & \frac{8}{5} \\ 0 & 1 & 0 & -1 & 0 & 1 & 4 \\ \hline 0 & \frac{7}{5} & 0 & \frac{3}{5} & \frac{3}{5} & 0 & \frac{21}{5} \end{array}\right)$$

得分	

三、（15 分）设 f 是 $\mathbf{R}^n$ 上的二阶连续可微函数，$\nabla^2 f$ 在点 x^* 的一领域内满足 Lipschitz 条件（其常数是 L），其中 x^* 是满足二阶充分条件的孤立局部极小点，设 $\{x^k\}$ 由经典牛顿法生成，证明

（i）当初始点 x^1 充分接近 x^* 时，序列 $\{x^k\}$ 有定义且收敛到 x^*;

（ii）$\{x^k\}$ 二次收敛到 x^*;

（iii）$\{\|\nabla f(x^k)\|\}$ 二阶收敛到 0.

得分	

四、若x^*是非线性规划问题的局部最优解，并且线性无关约束规范在该点成立，**证明**：一阶必要性条件（8 分）和二阶必要性条件（7 分）都成立。

得分	

五、（15 分）设 f 是 $\mathbf{R}^n$ 上的连续可微下有界函数， ∇f 是一致连续的。$\{x^k\}$ 由 $x^{k+1} = x^k + \alpha_k d^k$ 生成，其中 d^k 满足角条件，α_k 满足 Wolfe-Powell 条件。证明 $\{\nabla f(x^k)\}$ 收敛到 0.

.

得分	

六、（10 分）简述求解非线性规划问题的增广拉格朗日方法。

考虑问题 $\min_{x\in R^n} \|Ax-b\|_2^2+\lambda\|x\|_1$，其中 $A\in R^{m\times n}$，$\lambda>0$

简要叙述求解上述问题的两种方法，包括方法基本步骤和关键的迭代公式

Lasso

- 考虑问题 (Lasso)

$$\min_{x\in\mathbb{R}^n} \frac{1}{2}\|Ax-b\|_2^2+\gamma\|x\|_1$$

其中 $A\in\mathbb{R}^{m\times n}$ 和 $\gamma>0$

- (accelerated) proximal gradient method:

$$x^{k+1}=\mathbf{prox}_{t_k\gamma\|\cdot\|_1}(x^k-t_kA^T(Ax^k-b))$$

Lasso

- 考虑问题

$$\min_x \frac{1}{2}\|Ax-b\|^2+\lambda\|x\|_1$$

- 等价地写为

$$\begin{array}{ll}\min_{x,z} & \frac{1}{2}\|Ax-b\|^2+\lambda\|z\|_1 \\ \text{subject to} & x-z=0\end{array}$$

- ADMM:

$$\begin{aligned}x^{k+1}&=(A^TA+\rho I)^{-1}(A^Tb+\rho z^k-y^k)\\ z^{k+1}&=\mathbf{prox}_{\lambda/\rho\|\cdot\|_1}(x^{k+1}+y^k/\rho)\\ y^{k+1}&=y^k+\tau\rho(x^{k+1}-z^{k+1})\end{aligned}$$

简要叙述序列二次规划方法

写出下述问题的对偶问题

$$\min_{x\in R^n} \|Ax-b\|_2^2+\lambda\|x\|,$$ 其中 $A\in R^{m\times n}$，$\lambda>0$

证明：$f: R^n \to R$ 是凸函数当且仅当，对于任意的 $x^1, x^2, \cdots, x^m \in R^n\ (m \geq 2)$ 和 $a_1, a_2, \cdots, a_m \geq 0\ \left(\sum_{i=1}^{m} a_i = 1\right)$

有 $f(a_1 x^1 + \cdots + a_m x^m) \leq a_1 f(x^1) + \cdots + a_m f(x^m)$