- 考试时间：
- **2019**年**1**月**7**日晚 **19**：**00—21**：**00**
- 考试地点：
- 六教**6C300:2016210207~2018310437**
- 六教**6C201:2018310445~2018310744**
- 六教**6C202:2018310747~2018312568**及旁听生

- 答疑时间：
- **2019**年**1**月**7**日上午**9**：**00—11**：**30**
- **2019**年**1**月**7**日下午**2**：**30—5**：**00**
- 答疑地点：三教**1101**

- **1.**携带研究生证，以备查对。
- **2**．提前十分钟进入考场。考试开始十五分钟后，不准再进入考场，逾时以旷考论。题卷发出十五分钟后，方可交卷离场。
- **3**．除答卷必需用的文具及教师指定的考试用具外，书包、书籍、笔记、纸张等一律按监考教师要求集中放置。
- **4**．不允许携带具有信息传递或存储功能的工具（如手机等）进入考场。
- **5**．答卷一般用钢笔或圆珠笔（蓝色或黑色，不得用红色），不得用铅笔（画图或外语考试选择题等指定用铅笔除外）。
- **6**．答卷时不准互借文具。
- **7**．严禁以任何理由左顾右盼、交头接耳、抄袭或看别人答卷等各种形式的作弊行为。
- **8**．答卷时，不得中途离场后再行返回。如有特殊原因需离场者，必须经监考教师准许。答卷一经考生带出考场，即行作废。
- **9**．在规定的时间内答卷，不得拖延。交卷时间到，考生须在原座位安静地等候监考教师收卷后，方可离场。

- 禁止使用各种计算器！

总复习

- 一.凸集与凸函数
- 1. 凸集的定义、性质

设S_1和S_2是两个凸集，β实数，则

(1) $\beta S_1=\{\beta x\mid x\in S_1\}$是凸集;

(2) $S_1+S_2=\{x^{(1)}+x^{(2)}\mid x^{(1)}\in S_1, x^{(2)}\in S_2\}$是凸集;

(3) $S_1-S_2=\{x^{(1)}-x^{(2)}\mid x^{(1)}\in S_1, x^{(2)}\in S_2\}$是凸集;

(4) $S_1\cap S_2$是凸集;

2. 极点和极方向的定义

设S是非空集合，$x \in S$，若x不能表示成S中两个不同点的凸组合，即若假设$x = \lambda x^{(1)} + (1-\lambda)x^{(2)}$，必推出$x = x^{(1)} = x^{(2)}$，则称$x$是凸集$S$的极点。

要求：会证明或判断一个点是否是极点.

设S是闭凸集，d为非零向量，如果对S中的每一个x，有$\{x + \lambda d \mid \lambda \geq 0\} \subset S$，则称$d$是$S$的方向；又设$d^{(1)}$和$d^{(2)}$是$S$的两个方向，若对任何正数$\lambda$，有$d^{(1)} \neq \lambda d^{(2)}$，则称$d^{(1)}$和$d^{(2)}$是两个不同的方向，若$S$的方向$d$不能表示成该集合的两个不同方向的正的线性组合，则称d为S的极方向。

要求：会证明或判断一个非零向量是否是方向或极方向.

结论： 设$S=\{x \mid Ax=b, x\geq 0\}$为非空集合，$d$是非零向量，则$d$是$S$的方向的充要条件是$d\geq 0$且$Ad=0$。

了解表示定理

2．凸集分离定理

（1）会应用凸集分离定理

（2）掌握Farkas定理和Gordan定理和证明方法，会应用这两个定理证明相应的题目。

Farkas定理：设A为$m\times n$矩阵，c为n维列向量，则$Ax\le 0$，$c^T x>0$有解的充要条件是$A^T y=c$，$y\ge 0$无解。

Gordan定理：设A为$m\times n$矩阵，那么$Ax<0$有解的充要条件是不存在非零向量$y\ge 0$，使得$A^T y=0$。

3．凸函数(凹函数)

要求：掌握凸(凹)函数的定义、性质及判断方法，会证明或判断一个函数是否是凸(凹)函数 。

凸函数：设S是E^n中的非空凸集，$f(x)$是定义在S上的实函数，如果对于每一对x_1，$x_2 \in S$及每一个a，$0 \le a \le 1$,都有

$$f(ax_1+(1-a)x_2) \le a\, f(x_1)+(1-a)f(x_2)$$

则称函数$f(x)$为S上的凸函数．上式中，若≤变为<，则称为严格凸函数。

若$-f(x)$为S的凸函数，则称$f(x)$为S上的**凹函数**.

凸函数性质

(1) 设$f_1(x)$，$f_2(x)$是凸集S上的凸函数，则函数$f_1(x)+f_2(x)$在S上也是凸函数。

(2) 设$f(x)$是凸集S上的凸函数，则对任意的$a\geq 0$，函数$af(x)$是凸的。

推广：设$f_1(x)$，$f_2(x)$, …, $f_k(x)$是凸集S上的凸函数，$a_i\geq 0$,则$a_1f_1(x)+a_2f_2(x)+ \ldots + a_kf_k(x)$也是凸集$S$上的凸函数.

(3) 设$f(x)$是凸集S上的凸函数，对每一个实数c，则集合

$$S_c=\{x \mid x\in S,\ f(x)\leq c\}$$是凸集。

定理(一阶充要条件): 设S是E^n中非空开凸集，$f(x)$是定义在S上的可微函数，则$f(x)$为凸函数的充要条件是对任意两点$x^{(1)},x^{(2)}\in S$，有

$$f(x^{(2)})\geq f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)});$$

$f(x)$为严格凸函数的充要条件是对任意互不相同两点$x^{(1)},x^{(2)}\in S$，有

$$f(x^{(2)})> f(x^{(1)})+\nabla f(x^{(1)})^T(x^{(2)}-x^{(1)}).$$

定理(二阶充要条件)：设S是E^n中非空开凸集，$f(x)$是定义在S上的二次可微函数，则$f(x)$为凸函数的充要条件是对任意$x\in S$，$f(x)$在x处的$Hessian$矩阵$\nabla^2 f(x)$是半正定的。

凸规划

- 凸规划：求凸函数在凸集上的极小点。

$$
\begin{aligned}
\min \quad & f(x) \\
s.t. \quad & g_i(x) \geq 0,\ i=1,\cdots,m \\
& h_j(x) = 0,\ j=1,\cdots,l
\end{aligned}
$$

若$f(x)$是凸函数，$g_i(x)(i=1,\cdots,m)$是凹函数，$h_j(x)(j=1,\cdots,l)$是线性函数，则原问题为凸规划。

性质：凸规划的局部极小点就是整体极小点，且极小点的集合为凸集。

要求：会判断一个模型是否为凸规划

线性规划部分

LP的标准形式

1、极小化型

2、约束方程为等式

3、所有的决策变量为非负值

4、约束方程的右端项系数为非负值

$$\min \ z=\sum_{j=1}^{n} c_j x_j$$

$$s.t \quad \sum_{j=1}^{n} a_{ij} x_j = b_i \quad i=1,\cdots,m$$

$$x_j \geq 0 \quad j=1,\cdots,n$$

$$\min \ z=cx \qquad c_{1\times n}$$

$$s.t. \quad Ax=b \quad A_{m\times n} \quad b_{m\times 1}\geq 0$$

$$x\geq 0 \qquad x_{n\times 1}$$

- 1．基本概念:
- 可行域(线性规划的可行域是凸集).
- 解的情形:无解(无可行解)、无界解(不存在有限的最优解)、最优解(最优解与最优值的区别)、局部最优解与全局最优解.
- 可行解、基本解、基、基变量、非基变量、基本可行解、非退化（退化）的基本可行解。

- **2**．基本性质：
- 线性规划存在有限最优解的充要条件是所有$cd^{(j)}$为非负数，其中$d^{(j)}$为可行域的极方向。
- 若线性规划问题存在有限最优解，则目标函数的最优值可在某个极点达到。
- 基本可行解与可行域极点之间的关系---等价。
- 基本可行解的存在问题：有可行解，一定有基本可行解。

单纯形法

$$\begin{cases} \min & f(x) = cx \\ s.t. & Ax = b \\ & x \geq 0 \end{cases}$$

1.存在初始基B，使得$B^{-1}b \geq 0$.

如何判断该问题是否有最优解？

如何判断一个基是否为最优基？

如何判断该问题是否有无穷多最优解？

用单纯形表求解问题：

	x_B	x_N	右端
x_B	I_m	$B^{-1}N$	$B^{-1}b$
	0	$c_BB^{-1}N-c_N$	$c_BB^{-1}b$

假设 $\bar{b}=B^{-1}b\geq 0$, 有一基本可行解

$$x=\begin{pmatrix} x_B \\ x_N \end{pmatrix}=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

(1) 若 $c_BB^{-1}N-c_N\leq 0$（极小化问题），则现行基本可行解为最优解

(2) 若存在 $c_BB^{-1}P_j-c_j>0$，用 主元消去法 求改进的基本可行解

2．寻找初始基本可行解

$$\begin{cases} \min \quad f(x) = cx \\ s.t. \qquad Ax = b \\ \qquad\qquad x \geq 0 \end{cases}$$

两阶段法
$$\begin{cases} \min \quad e^T x_a \\ s.t. \qquad Ax + x_a = b \quad e = (1 1 \cdots 1)^T \\ \qquad\qquad x, x_a \geq 0 \end{cases}$$

大**M**法
$$\begin{cases} \min \quad cx + Me^T x_a \\ s.t. \qquad Ax + x_a = b \quad e = (1 1 \cdots 1)^T_{m\times 1} \quad M > 0 \\ \qquad x \geq 0 \end{cases}$$

对偶原理

	min	max	
变量	≥ 0	$\leq$	约束方程
	≤ 0	$\geq$	
	无限制	$=$	
约束方程	$\geq$	≥ 0	变量
	$\leq$	≤ 0	
	$=$	无限制	

- 会写各种形式（对称、非对称、一般形式）的对偶问题；
- 掌握弱对偶定理和强对偶定理及其相关推论；
- 会用互补松弛定理求原问题或对偶问题的解；

小结

原问题(min)	对应关系	对偶问题(max)
有最优解	↔	有最优解
无界解	↔	不可行
不可行	↔	无界解

对偶单纯形法

- 掌握对偶单纯形方法（对偶可行的基本解，如何求初始对偶可行的基本解）。
- 与原单纯形法的区别：
- 原单纯形法保持原问题的可行性，对偶单纯形法保持所有检验数$wP_j-c_j\leq 0$，即保持对偶问题的可行性。
- 特点：先选择出基变量，再选择进基变 量。

- **5**．灵敏度分析
- 改变非基变量和基变量价格系数后，原问题最优解和最优值的改变，会求最优解不变时价格系数的变化范围；
- 右端向量改变对最优解及最优值的影响，会求最优基不变时，右端向量的变化范围；
- 会判断增加新的约束后，原问题最优解是否发生变化

最优性条件

- 无约束问题的极值条件：

定理1:(一阶必要条件)设函数$f(x)$在点$\bar{x}$处可微，若$\bar{x}$是局部极小点，则$\nabla f(\bar{x})=0$.

定理2:(二阶必要条件)设$f(x)$在$\bar{x}$处二阶可微，若$\bar{x}$是局部极小点，则$\nabla f(\bar{x})=0$，且$Hessian$矩阵$\nabla^2 f(\bar{x})$是半正定的。

定理3：设函数$f(x)$在点$\bar{x}$处二次可微，若梯度$\nabla f(\bar{x})=0,$且$Hessian$矩阵$\nabla^2 f(\bar{x})$正定，则$\bar{x}$是严格局部极小点。

推论：对于正定二次函数$f(x)=\frac{1}{2}x^T Ax+b^T x+c$

(A对称正定)，有唯一极小点$x^*=-A^{-1}b$.

定理4：设$f(x)$是定义在E^n上的可微凸函数，$\bar{x}\in E^n$，则$\bar{x}$为整体极小点的充要条件是$\nabla f(\bar{x})=0$.

约束极值问题的最优性条件

对$\min\limits_{x\in E^n} f(x)$，设$\bar{x}\in E^n$是任给一点，

$d\neq 0$，若存在$\delta>0$，使得对任意的$\lambda\in(0,\delta)$，

有$f(\bar{x}+\lambda d)<f(\bar{x})$，则称$d$为$f(x)$在点$\bar{x}$处的

下降方向(descent direction)。

$$F_0=\left\{d \,\middle|\, \nabla f(\bar{x})^T d<0\right\}$$

称为点$\bar{x}$处的下降方向集。

定义： 设集合$S \subset E^n, \bar{x} \in clS, d$为非零向量，若存在数$\delta > 0$,使得对任意$\lambda \in (0, \delta)$，都有

$$\bar{x} + \lambda d \in S$$

则称d为集合S在$\bar{x}$的可行方向(feasible direction)。

$$D = \left\{ d \middle| \, d \neq 0, \bar{x} \in clS, \exists \delta > 0, \forall \lambda \in (0, \delta), 有\bar{x} + \lambda d \in S \right\}$$

$\bar{x}$处的可行方向锥。

定理：设x是问题(A)的可行解，$I=\{i \mid g_i(x)=0\}$是在x处起作用约束下标集，又设$f(x)$，$g_i(x)(i\in I)$在x处可微，$g_i(x)(i\notin I)$在x处连续，如果 $\nabla f(x)^T d<0$，$\nabla g_i(x)^T d>0(i\in I)$，则$d$是可行下降方向。

$$(A)\quad \begin{cases} \min\ f(x) \\ s.t.\quad g_i(x)\geq 0,\ i=1,2,\cdots,m \end{cases}$$

其中$f(x), g_i(x)$均为可微函数。

定理2(KKT必要条件）考虑问题$\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\ge 0, i=1,\cdots,m \\ \qquad\quad h_j(x)=0,\ j=1,\cdots,l\end{cases}$

设$\overline{x}$为可行点,$I=\{i\mid g_i(\overline{x})=0\}.f,g_i(i\in I)$在$\overline{x}$处可微，

$g_i(i\notin I)$在$\overline{x}$连续，$h_j(j=1,\cdots,l)$在$\overline{x}$连续可微，向量集

$$\{\nabla g_i(\overline{x}),\ \nabla h_j(\overline{x})\mid i\in I, j=1,\cdots,l\}$$

线性无关，若$\overline{x}$是局部最优解，则存在数$w_i,i\in I$

和$v_j(j=1,\cdots,l)$，使得

$$\nabla f(\overline{x})-\sum_{i\in I}w_i\nabla g_i(\overline{x})-\sum_{j=1}^{l}v_j\nabla h_j(\overline{x})=0.$$

$$w_i\ge 0\quad (i\in I).$$

定理2'(*KKT*必要条件）考虑问题$\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\ge 0, i=1,\cdots,m \\ \qquad\quad h_j(x)=0,\ j=1,\cdots,l\end{cases}$

设$\bar{x}$为可行点,$I=\{i\mid g_i(\bar{x})=0\}$.$f,g_i$在$\bar{x}$处可微，$h_j(j=1,\cdots,l)$在$\bar{x}$连续可微，向量集

$$\{\nabla g_i(\bar{x}),\ \nabla h_j(\bar{x})\mid i\in I, j=1,\cdots,l\}$$

线性无关，若$\bar{x}$是局部最优解，则存在数$w_i, i\in I$和$v_j(j=1,\cdots,l)$，使得

$$\nabla f(\bar{x})-\sum_{i=1}^{m}w_i\nabla g_i(\bar{x})-\sum_{j=1}^{l}v_j\nabla h_j(\bar{x})=0$$

$$w_i g_i(\bar{x})=0, i=1,2,\cdots,m$$

$$w_i\ge 0\quad i=1,2,\cdots,m.$$

定义广义的Lagrange函数:

$$L(x,w,v)=f(x)-\sum_{i=1}^{m}w_i g_i(x)-\sum_{j=1}^{l}v_j h_j(x)$$

$$=f(x)-w^T g(x)-v^T h(x)$$

其中$w=(w_1,w_2,\cdots,w_m)^T$

$$v=(v_1,v_2,\cdots,v_l)^T$$

乘子向量

$$g(x)=(g_1(x),\cdots,g_m(x))^T$$

$$h(x)=(h_1(x),\cdots,h_l(x))^T.$$

定理2'(KKT必要条件）考虑问题$\begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \geq 0, i=1,\cdots,m \\ \qquad\quad h_j(x)=0,\ j=1,\cdots,l \end{cases}$

设$\overline{x}$为可行点,$I=\{i \mid g_i(\overline{x})=0\}$.$f, g_i$在$\overline{x}$处可微，$h_j(j=1,\cdots,l)$在$\overline{x}$连续可微，向量集

$$\{\nabla g_i(\overline{x}),\ \nabla h_j(\overline{x}) \mid i \in I, j=1,\cdots,l\}$$

线性无关，若$\overline{x}$是局部最优解，则存在乘子向量$\overline{w} \geq 0, \overline{v}$,使得

$$\nabla_x L(\overline{x}, \overline{w}, \overline{v})=0$$

定理3.(一阶充分条件)

设问题$\begin{cases}\min\ f(x) \\ s.t.\quad g_i(x)\ge 0\ \ i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0\ \ j=1,2,\cdots,l\end{cases}$

中，f是凸函数，$g_i(i=1,2,\cdots,m)$是凹函数，

$h_j(j=1,2,\cdots,l)$是线性函数，S为可行域，

$\bar{x}\in S$，$I=\{i\mid g_i(\bar{x})=0\}$。$f$和$g_i(i\in I)$在点$\bar{x}$可微，

$h_j(j=1,2,\cdots,l)$在点$\bar{x}$连续，$g_i(i\notin I)$在点$\bar{x}$连续，

且在$\bar{x}$处KKT条件成立，则$\bar{x}$为整体极小点。

推论1：设(NP)是凸规划，则$\bar{x}\in S$是整体最优解

$\Leftrightarrow\bar{x}$是KKT点。

定理（二阶必要条件）：设$\bar{x}$是(NP)的局部最优解，f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$二次连续可微，且在$\bar{x}$处，$\{\nabla g_i(\bar{x}), i\in I, \nabla h_j(\bar{x}), j=1,\cdots,l\}$为线性无关组，则存在$\bar{w},\bar{v}$，使$(\bar{x},\bar{w},\bar{v})$为$(LM)$的解且矩阵$\nabla^2 L_x(\bar{x},\bar{w},\bar{v})$在子空间$\bar{G}$上是半正定的，其中

$$\bar{G}=\left\{d\left|\begin{array}{l}\nabla g_i(\bar{x})^T d=0, i\in I\text{且}\bar{w}_i>0\\ \nabla g_i(\bar{x})^T d\geq 0, i\in I\text{且}\bar{w}_i=0\\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l\end{array}\right.\right\}.$$

定理（二阶充分条件）：设f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$是二次连续可微函数，$\overline{x}$为可行解，若存在$\overline{w},\overline{v}$，使$(\overline{x},\overline{w},\overline{v})$为$(LM)$的解且矩阵$\nabla^2 L_x(\overline{x},\overline{w},\overline{v})$在子空间$G$上是正定的，则$\overline{x}$是严格局部极小点。

其中
$$G=\left\{ d\neq 0 \left| \begin{array}{l} \nabla g_i(\overline{x})^T d=0, i\in I \text{且} \overline{w}_i>0 \\ \nabla g_i(\overline{x})^T d\geq 0, i\in I \text{且} \overline{w}_i=0 \\ \nabla h_j(\overline{x})^T d=0, j=1,2,\cdots,l \end{array} \right. \right\}.$$

对偶问题（要求：会写对偶问题）

Lagrange对偶问题

$$
\begin{aligned}
&\min\ f(x) \\
&s.t.\quad g_i(x) \geq 0,\ i=1,\cdots,m \\
&\qquad\ \ h_j(x) = 0,\ j=1,\cdots,l \qquad\qquad (1) \\
&\qquad\ \ x \in D
\end{aligned}
$$

定义(1)的对偶问题：

$$
\begin{aligned}
&\max\ \theta(w,v) \\
&s.t.\quad w \geq 0 \qquad\qquad (2)
\end{aligned}
$$

其中$\theta(w,v)=\inf\left\{f(x)-\sum_{i=1}^{m} w_i g_i(x)-\sum_{j=1}^{l} v_j h_j(x) \middle| x \in D\right\}$

若上式不存在有限下界时，令$\theta(w,v)=-\infty$.

$\theta(w,v)$称为$Lagrange$对偶函数。

定理1(弱对偶定理)

设x和(w,v)分别是原问题和对偶问题的可行解，则

$$f(x) \geq \theta(w,v)。$$

推论1: 对于原问题和对偶问题，必有

$$\inf\{f(x) \mid g(x) \geq 0, h(x) = 0, x \in D\} \geq \sup\{\theta(w,v) \mid w \geq 0\}.$$

推论2: 若$f(\bar{x}) \leq \theta(\bar{w},\bar{v})$，其中$\bar{x}$为原问题的可行解，$\bar{w} \geq 0$，则$\bar{x}$和$(\bar{w},\bar{v})$分别是原问题和对偶问题的最优解。

推论3: 若$\inf\{f(x) \mid g(x) \geq 0, h(x) = 0, x \in D\} = -\infty$，则对$\forall w \geq 0$，有$\theta(w,v) = -\infty$。

推论4: 如果$\sup\{\theta(w,v) \mid w \geq 0\} = +\infty$，则原问题没有可行解。

算法

- 闭映射的定义
- 会判断一个具体的算法映射在某个点是否具有闭的性质。
- 二次终止性的定义：若某个算法对于任意的正定二次函数，从任意的初始点出发，都能经有限步迭代达到其极小点，则称该算法具有二次终止性。

一维搜索

- 一维搜索的定义
- 一维搜索的性质：

设目标函数$f(x)$具有一阶偏导数，$x^{(k+1)}$由下列规则产生：

$$\begin{cases} f(x^{(k)}+\lambda_k d^k) = \min\limits_{\lambda} f(x^{(k)}+\lambda d^k) \\ x^{(k+1)} = x^{(k)} + \lambda_k d^k \end{cases}$$

则有$\nabla f(x^{(k+1)})^T d^k = 0$。

使用导数的最优化方法

- 最速下降方向的定义：

$$d^{(k)} = -\nabla f(x^{(k)})$$

- 牛顿方向：

$$d^{(k)} = -\nabla^2 f(x^{(k)})^{-1} \nabla f(x^{(k)})$$

共轭方向法

共轭方向

定义： 设A是$n\times n$对称正定矩阵，若E^n中的两个方向$d^{(1)}$和$d^{(2)}$满足

$$(d^{(1)})^T A d^{(2)} = 0$$

则称这两个方向关于A共轭，或称它们关于A正交。

性质

设A是n阶对称正定矩阵，$d^{(1)}, d^{(2)}, \cdots, d^{(k)}$是$k$个$A$共轭的非零向量，则这$k$个向量线性无关。

定理2：设有二次函数

$$f(x)=\frac{1}{2}x^T Ax+b^T x+c,$$

其中$A_{n\times n}$是对称正定矩阵，$d^{(0)},d^{(1)},\cdots,d^{(n-1)}$是$A$共轭的非零向量，从任意一点$x^{(0)}\in E^n$出发，依次沿这组向量进行一维搜索，

$$\min_{\lambda\geq 0} f(x^{(k)}+\lambda d^{(k)})=f(x^{(k)}+\lambda_k d^{(k)})$$

$$x^{(k+1)}=x^{(k)}+\lambda_k d^{(k)}\quad k=0,1,\cdots,n-1$$

则$\nabla f(x^{(k+1)})^T d^{(j)}=0,\ j=0,1,\cdots,k$。

定理（扩张子空间定理,expanding subspace theorem）

设有函数 $f(x)=\frac{1}{2}x^T Ax+b^T x+c$

其中A为n阶对称正定矩阵，$d^{(1)},d^{(2)},\cdots,d^{(k)}$是$A$共轭的非零向量。以任意的$x^{(1)}\in E^n$为初始点，依次沿$d^{(1)},d^{(2)},\cdots,d^{(k)}$进行一维搜索，得到点$x^{(2)},x^{(3)},\cdots,x^{(k+1)}$，则$x^{(k+1)}$是$f(x)$在线性流形

$$M_k\left(x^{(1)};\left\{d^{(1)},d^{(2)},\cdots,d^{(k)}\right\}\right)=\left\{x\,\middle|\,x=x^{(1)}+\sum_{i=1}^{k}\mu_i d^{(i)},\mu_i\in R\right\}$$

$$=x^{(1)}+B_k$$

上的唯一极小点。特别的，当$k=n$时，$x^{(n+1)}$是$f(x)$在E^n上的唯一极小点。

可行方向法：

$$\begin{cases} \min\ f(x) \\ s.t. \quad Ax \geq b \\ \qquad\quad Ex = e \end{cases} \qquad (1)$$

定理： 设$\bar{x}$是问题(1)的可行解，在$\bar{x}$点处，有

$$A_1\bar{x} = b_1,\ A_2\bar{x} > b_2$$

其中$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$，则非零向量$d$是$\bar{x}$处的可行方向的充要条件是$A_1 d \geq 0, Ed = 0$.

$$(2)\quad \begin{cases} \min \nabla f(x)^T d \\ s.t. \quad A_1 d \geq 0 \\ \qquad\quad Ed = 0 \\ \qquad\quad |d_j| \leq 1,\ j = 1,2,\cdots,n \end{cases}$$

在问题(1)中，设x是可行解，在点x处有

$A_1 x = b_1$，$A_2 x > b_2$，其中

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix},\quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix},$$

则x是KKT点的充要条件是问题(2)的目标函数最优值$=0$。

定理： 设$\alpha_1,\alpha_2,\cdots,\alpha_r$是子空间$U$的一组基，记

$$M=\begin{bmatrix}\alpha_1^T\\ \alpha_2^T\\ \vdots\\ \alpha_r^T\end{bmatrix}_{r\times n}$$

则

(1)由E^n到U的正交投影矩阵Q可表示为

$$Q=M^T(MM^T)^{-1}M。$$

(2)由E^n到$U^\perp$的正交投影矩阵P可表示为

$$P=I-M^T(MM^T)^{-1}M。$$

推论1: 若Q是正交投影矩阵，则$Q^T=Q$，$QQ=Q$.

推论2: 若Q是正交投影矩阵，则Q是半正定的.

定理: 设x是问题(1)的可行解，在点x处，有$A_1x=b_1$，$A_2x>b_2$，其中

$$A=\begin{bmatrix}A_1\\A_1\end{bmatrix}\quad b=\begin{bmatrix}b_1\\b_2\end{bmatrix}$$

又设$M=\begin{bmatrix}A_1\\E\end{bmatrix}$为满秩矩阵，令

$$P=I-M^T\left(MM^T\right)^{-1}M$$

若$P\nabla f(x)\neq 0$，令$d=-P\nabla f(x)$，则d为下降可行方向。

定理2: 设x是问题(1)的可行解，在点x处，有$A_1x = b_1$，$A_2x > b_2$，其中

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

又设$M = \begin{bmatrix} A_1 \\ E \end{bmatrix}$为满秩矩阵，令

$$P = I - M^T \left(MM^T \right)^{-1} M$$

$$W = \left(MM^T \right)^{-1} M \nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}$$

其中u和v分别对应于A_1和E。设$P\nabla f(x) = 0$，则

1. 若$u \geq 0$,则x为KKT点；

2. 如果u中含有负分量，不妨设$u_j < 0$，这时从A_1中去掉u_j对应的行，得到$\hat{A}_1$。令

$$\hat{M} = \begin{bmatrix} \hat{A}_1 \\ E \end{bmatrix} \qquad \hat{P} = I - \hat{M}^T \left(\hat{M}\hat{M}^T \right)^{-1} \hat{M}$$

$$d = -\hat{P}\nabla f(x)$$

则d为x处的下降可行方向。

$$\boxed{W = \left(MM^T \right)^{-1} M\nabla f(x) = \begin{bmatrix} u \\ v \end{bmatrix}}$$

Wolfe既约梯度法

$$\min \ f(x)$$

$$s.t. \quad Ax = b, \quad x \geq 0$$

$A_{m\times n}, r(A) = m, b_{m\times 1}, f$是$E^n$上的连续可微函数

$$A=(B,N),\quad x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$$

$$\begin{aligned}&\min\ f(x_B,x_N)\\&s.t.\quad Bx_B+Nx_N=b\\&\qquad\ \ x_B,x_N\geq 0\end{aligned}\quad\Longleftrightarrow\quad\begin{aligned}&\min F(x_N)\\&s.t.\quad x_B,x_N\geq 0\end{aligned}$$

$$x_B=B^{-1}b-B^{-1}Nx_N$$

$$\Rightarrow F(x_N)=f(x_B(x_N),x_N)$$

***f*(*x*)的既约梯度为**

$$r(x_N)=\nabla F(x_N)$$

$$=\nabla_{x_N}f(x)-(B^{-1}N)^T\nabla_{x_B}f(x)$$

定理 设x是可行解，$A=(B,N)$是$m\times n$矩阵，B为m阶可逆矩阵，$x=\left(x_B{}^T, x_N{}^T\right)^T$，$x_B>0$，函数$f$在点$x$处可微，又设

$$d=\begin{pmatrix} -B^{-1}Nd_N \\ d_N \end{pmatrix},$$

其中 $$d_{N_j}=\begin{cases} -x_{N_j} r_j(x_N) & r_j(x_N)>0 \\ -r_j(x_N) & r_j(x_N)\le 0 \end{cases}$$

如果$d\neq 0$，则d是下降可行方向，而且$d=0$的充要条件是x为KKT点。

惩罚函数法

- 外点法的定义及相关性质。

$$\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0\ i = 1,2,\cdots,m \\ \qquad\quad h_j(x) = 0\ j = 1,2,\cdots,l \end{cases}$$

$$\min\ F(x,\sigma) = f(x) + \sigma\left[\sum_{i=1}^{m} \varphi\left[g_i(x)\right] + \sum_{j=1}^{l} \psi\left[h_j(x)\right]\right]$$

函数φ和ψ的典型取法：

$$\varphi\left[g_i(x)\right] = \left[\max\{0, -g_i(x)\}\right]^{\alpha} \quad \psi\left[h_j(x)\right] = \left|h_j(x)\right|^{\beta}$$

其中$\alpha \geq 1, \beta \geq 1$均为给定常数，通常取$\alpha = \beta = 2$。

- 内点法的定义及性质

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ \ i = 1,2,\cdots,m \end{cases}$$

$$\begin{cases} \min G(x,r) = f(x) + rB(x) \\ s.t. \quad x \in \operatorname{int} S \end{cases}$$

$$B(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)} \qquad B(x) = -\sum_{i=1}^{m} \ln g_i(x)$$