第二章证明题

1. $f(x)$ 为凸集 $D \subset R^n$ 上的函数，上图 $epi(f)=\{(x,y) \mid x \in D, y \in R, y \ge f(x)\}$，证明 $f(x)$ 为凸函数的充要条件是 $epi(f)$ 为凸集。

证明： 充分性：对于任意的 $x_1, x_2 \in D, \alpha \in [0,1]$，由 $epi(f)$ 的定义，

$(x_1, f(x_1)), (x_2, f(x_2)) \in epi(f)$，

而 $epi(f)$ 为凸集，得

$$\alpha(x_1, f(x_1)) + (1-\alpha)(x_2, f(x_2)) \in epi(f),$$

即 $(\alpha x_1 + (1-\alpha)x_2, \alpha f(x_1) + (1-\alpha) f(x_2)) \in epi(f)$，

因此 $\alpha f(x_1) + (1-\alpha) f(x_2) \ge f(\alpha x_1 + (1-\alpha)x_2)$，从而 $f(x)$ 为凸函数。

必要性：对于任意的 $(x_1, y_1), (x_2, y_2) \in epi(f), \alpha \in [0,1]$，有 $y_1 \ge f(x_1), y_2 \ge f(x_2)$

由于 $f(x)$ 为凸函数，有

$$f(\alpha x_1 + (1-\alpha)x_2) \le \alpha f(x_1) + (1-\alpha) f(x_2) \le \alpha y_1 + (1-\alpha) y_2$$

所以 $(\alpha x_1 + (1-\alpha)x_2, \alpha y_1 + (1-\alpha) y_2) \in epi(f)$，

即 $\alpha(x_1, y_1) + (1-\alpha)(x_2, y_2) \in epi(f)$，

因此 $epi(f)$ 为凸集。

注：典型错误：充分性的证明中，一开始就取 $(x_1, f(x_1)), (x_2, f(x_2)) \in epi(f)$，凸函数的定义不全；必要性证明中，一开始就取 $x_1, x_2 \in D, \alpha \in [0,1]$

2．考虑规划问题： $\begin{array}{ll} \min & f(x) \\ s.t. & c_i(x) \le 0 \end{array}$，其中，$\boldsymbol{f(x), c_i(x) (i=1,2,\cdots,m): R^n \to R}$ 是凸函数，证明：（1）该问题的可行域是凸集；（2）该问题的最优解的集合 $\boldsymbol{A}$ 是凸集。

证明：（1）设 $\boldsymbol{D_i = \{x \mid c_i(x) \le 0\}}$，可行域 $\boldsymbol{D = \bigcap_{i=1}^{m} D_i}$ 。

对于任意 $\boldsymbol{x_1, x_2 \in D_i = \{x \mid c_i(x) \le 0\}}$，任意 $\boldsymbol{\alpha \in [0,1]}$，有 $\boldsymbol{c_i(x_1) \le 0, c_i(x_2) \le 0}$，

根据 $\boldsymbol{c_i(x)}$ 是凸函数，$\boldsymbol{c_i(\alpha x_1 + (1-\alpha)x_2) \le \alpha c_i(x_1) + (1-\alpha) c_i(x_2) \le 0}$，

因此 $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \boldsymbol{D}_i$ 。

凸集的交集是凸集，因此 $\boldsymbol{D} = \bigcap_{i=1}^{m} \boldsymbol{D}_i$ 为凸集。

（2）设 $\boldsymbol{A}$ 为最优解的集合，若 A 不是空集，任取 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{A}, \alpha \in [0,1]$，有 $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \boldsymbol{D}$ 。

由于 $\boldsymbol{D}$ 是凸集，$\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2 \in \boldsymbol{D}$ 。

$\boldsymbol{f}(\boldsymbol{x})$ 是凸函数，$\boldsymbol{f}(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \le \alpha \boldsymbol{f}(\boldsymbol{x}_1) + (1-\alpha)\boldsymbol{f}(\boldsymbol{x}_2)$。

$\boldsymbol{x}_1, \boldsymbol{x}_2$ 都是最优解，因此 $\boldsymbol{f}(\boldsymbol{x}_1) = \boldsymbol{f}(\boldsymbol{x}_2)$ 为最优函数值。得

$$\boldsymbol{f}(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) \le \alpha \boldsymbol{f}(\boldsymbol{x}_1) + (1-\alpha)\boldsymbol{f}(\boldsymbol{x}_2) = \boldsymbol{f}(\boldsymbol{x}_1)$$

由 $\boldsymbol{x}_1$ 都是最优解，得：$\boldsymbol{f}(\boldsymbol{x}_1) \le \boldsymbol{f}(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2)$

从而 $\boldsymbol{f}(\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2) = \boldsymbol{f}(\boldsymbol{x}_1)$

因此 $\alpha \boldsymbol{x}_1 + (1-\alpha)\boldsymbol{x}_2$ 也是最优解，从而 A 为凸集

3．设 z^*, s^* 分别为下列两个问题 $(I)\ \begin{array}{ll} & \min c^T x \\ s.t. & Ax = b \\ & x \ge 0 \end{array}$ $(II)\ \begin{array}{ll} & \min c^T x \\ s.t. & Ax = b + d \\ & x \ge 0 \end{array}$ 的最优值。y^* 是 (I)的对偶问题的最优解，证明 $z^* + y^{*T} d \le s^*$ 。

证明:(I) 与 (II) 的对偶规划分别为

$$(DI)\ \begin{array}{ll} & \max b^T y \\ s.t. & A^T y \le c \end{array} \quad (DII)\ \begin{array}{ll} & \max (b+d)^T y \\ s.t. & A^T y \le c \end{array}$$

(I) 的最优值与 (DI) 的最优值相同，得：$z^* = b^T y^*$

y^* 是对偶规划 (DI) 的最优解，从而是 (DII) 的可行解。

y^* 在 (DII) 的目标函数值不大于最优值，$(b+d)^T y^* \le s^*$

即 $b^T y^* + d^T y^* \le s^*$，因此 $z^* + y^{*T} d \le s^*$

4 设 $\bar{x}, \bar{y}$ 分别为下列两个问题

$$(I)\ \begin{array}{ll} & \min c^T x \\ s.t. & Ax \ge b \\ & x \ge 0 \end{array} \qquad (II)\ \begin{array}{ll} & \max y^T b \\ s.t. & y^T A \le c^T \\ & y \ge 0 \end{array}$$

的可行解。证明$c^T\bar{x}\ge\bar{y}^Tb$。

证明：若$\bar{x},\bar{y}$分别是(I)与(II)的可行解,则

$$A\bar{x}\ge b,\bar{x}\ge 0,\bar{y}^TA\le c^T,\bar{y}\ge 0$$

由$A\bar{x}-b\ge 0$，$\bar{y}\ge 0$ 得：$\bar{y}^T(A\bar{x}-b)\ge 0$，即$\bar{y}^TA\bar{x}\ge\bar{y}^Tb$

由$\bar{x}\ge 0$，$\bar{y}^TA-c^T\le 0$得：$(\bar{y}^TA-c^T)\bar{x}\le 0$，即$c^T\bar{x}\ge\bar{y}^Tb$

第三章证明题

1. 设G为n阶正定对称矩阵，$u_1,u_2,\cdots,u_n\in R^n$线性无关。$p_k$按如下方式生成：$p_1=u_1$，

$$\boldsymbol{p}_{k+1}=\boldsymbol{u}_{k+1}-\sum_{i=1}^{k}\frac{\boldsymbol{u}_{k+1}^T\boldsymbol{G}\boldsymbol{p}_i}{\boldsymbol{p}_i^T\boldsymbol{G}\boldsymbol{p}_i}\boldsymbol{p}_i(\boldsymbol{k}=1,2,\cdots,\boldsymbol{n}-1),$$

证明$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$关于$\boldsymbol{G}$共轭。

证明：利用“数学归纳法”进行证明

由于$p_1^TGp_2=p_1^TG(u_2-\frac{u_2^TGp_1}{p_1^TGp_1}p_1)=p_1^TGu_2-u_2^TGp_1=0$,

所以$\boldsymbol{p}_1,\boldsymbol{p}_2$关于$\boldsymbol{G}$共轭。

设$\boldsymbol{p}_i,\boldsymbol{p}_j(1\le\boldsymbol{i}<\boldsymbol{j}\le\boldsymbol{k})$关于$\boldsymbol{G}$共轭，即$p_i^TGp_j=0$，下证$\boldsymbol{p}_j(1\le\boldsymbol{j}\le\boldsymbol{k})$与$\boldsymbol{p}_{k+1}$关于$\boldsymbol{G}$共轭

$$p_j^TGp_{k+1}=p_j^TG(u_{k+1}-\sum_{i=1}^{k}\frac{u_{k+1}^TGp_i}{p_i^TGp_i}p_i)=p_j^TGu_{k+1}-\sum_{i=1}^{k}\frac{u_{k+1}^TGp_i}{p_i^TGp_i}p_j^TGp_i,$$

由于$p_i^TGp_j=0$（j≠i），因此

$$p_j^TGp_{k+1}=p_j^TGu_{k+1}-\sum_{i=1}^{k}\frac{u_{k+1}^TGp_i}{p_i^TGp_i}p_j^TGp_i=p_j^TGu_{k+1}-u_{k+1}^TGp_j=0$$

故$\boldsymbol{p}_1,\boldsymbol{p}_2,\cdots,\boldsymbol{p}_n$关于$\boldsymbol{G}$共轭

2. 设A为n阶对称正定矩阵，非零向量$p_1,p_2,\cdots,p_n$关于矩阵A共轭，证明

$$A^{-1}=\sum_{k=1}^{n}\frac{p_kp_k^T}{p_k^TAp_k}。$$

证明：设 $B=\sum_{k=1}^{n}\frac{p_k p_k^T}{p_k^T A p_k}$，

由于非零向量 $p_1,p_2,\cdots,p_n$ 关于矩阵 A 共轭，所以 $p_k^T A p_i=0(k\neq i)$，

对于任意 $i(1\le i\le n)$，

$$BAp_i=\sum_{k=1}^{n}\frac{p_k p_k^T Ap_i}{p_k^T Ap_k}=\frac{p_i p_i^T Ap_i}{p_i^T Ap_i}=p_i$$

因此 $BA[p_1,p_2,\cdots,p_n]=[BAp_1,BAp_2,\cdots,BAp_n]=[p_1,p_2,\cdots,p_n]$。

非零向量 $p_1,p_2,\cdots,p_n$ 关于矩阵 A 共轭，所以 $p_1,p_2,\cdots,p_n$ 线性无关，从而

$[p_1,p_2,\cdots,p_n]$ 可逆，所以有 $BA=I,B=A^{-1}.$

3. 用 PRP 共轭梯度法求解无约束优化问题 $\min\ f(x)$,若一维搜索是精确的，且在求解过程中，每一步的梯度都是非零向量，证明 PRP 方法产生的搜索方向都是下降方向。

证明：（1） $p_0=-g_0$，$p_0^T g_0=-g_0^T g_0<0$,所以 p_0 是下降方向。

（2）$k\ge 1$ 时，$p_k=-g_k+\beta_{k-1}p_{k-1}$，其中 $\beta_{k-1}=\frac{g_k^T(g_k-g_{k-1})}{g_{k-1}^T g_{k-1}}$，

由于一维搜索是精确的，所以 $g_k^T p_{k-1}=0$

近而，$g_k^T p_k=-g_k^T g_k+\beta_{k-1}g_k^T p_{k-1}=-g_k^T g_k<0$

所以 p_k 是下降方向。

4. 设 A 为 n 阶对称正定矩阵，非零向量 $p_1,p_2,\cdots,p_n$ 关于矩阵 A 共轭，$x\in R^n$.证明 $x=\sum_{k=1}^{n}\frac{p_k^T Ax}{p_k^T Ap_k}p_k$。

证明：非零向量 $p_1,p_2,\cdots,p_n$ 关于矩阵 A 共轭，因此 $p_1,p_2,\cdots,p_n$ 线性无关，进而，

$p_1,p_2,\cdots,p_n$ 构成 R^n 的一组基底。

设 x 在这组基底下的线性表示为 $x=\sum_{k=1}^{n}\alpha_k p_k$。

对任意$\boldsymbol{i} \in \{1, 2, \cdots, n\}$，有 $p_i^T Ax = \sum_{k=1}^{n} \alpha_k p_i^T Ap_k$ 。

$p_1, p_2, \cdots, p_n$关于矩阵A共轭，当$k \neq i$时 $p_j^T Ap_k = 0$，

因此　　$p_i^T Ax = \alpha_i p_i^T Ap_i$，

从而$\alpha_i = \dfrac{p_i^T Ax}{p_i^T Ap_i}.$

因此$x = \sum_{k=1}^{n} \dfrac{p_k^T Ax}{p_k^T Ap_k} p_k$

第四章证明题

1. $f : R^n \to R$为可微凸函数，证明x^*为优化问题$\begin{array}{ll} \min & f(x) \\ s.t. & x \geq 0 \end{array}$的最优解的充要条件是

$x^* \geq 0, \nabla f(x^*) \geq 0, (\nabla f(x^*))^T x^* = 0$ 。

证明：$\begin{array}{ll} \min & f(x) \\ s.t. & x \geq 0 \end{array}$是凸规划问题，KT 条件是最优解的充要条件。

问题有n个约束，约束函数的梯度为$e_i (i = 1, 2, \cdots, n)$，显然线性无关，该问题的 KT 条件是存在乘子向量$\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_n)$，使得

$$\nabla f(x^*) - \sum_{i=1}^{n} \lambda_i e_i = 0\text{，}\ \lambda_i x_i = 0\text{，}\ \lambda_i \geq 0\text{。}$$

结合解的可行性条件，得最优解的充要条件是

$$x^* \geq 0\text{，}\ \lambda = \nabla f(x^*)\text{，}\ \lambda_i x_i = 0\text{，}\ \lambda_i \geq 0\text{。}$$

充分性：若$x^* \geq 0, \nabla f(x^*) \geq 0, (\nabla f(x^*))^T x^* = 0$，

令$\lambda = \nabla f(x^*)$，则有$x^* \geq 0, \lambda = \nabla f(x^*) \geq 0, \lambda^T x^* = 0$。

即，x^*, λ两个非负向量的内积为零，

显然有$\lambda_i x_i = 0 (i = 1, 2, \cdots, n)$。

必要性：若$x^* \geq 0$，$\lambda = \nabla f(x^*)$，$\lambda_i x_i = 0$，$\lambda_i \geq 0$。

则 $x^* \ge 0, \nabla f(x^*) \ge 0$，

$$(\nabla f(x^*))^T x^* = \sum_{i=1}^{n} \lambda_i x_i = 0 \text{。}$$

附加证明题

1. 对于任意给定的一个函数 $f(x)$,用最速下降法求解其最小值,若一维搜索是精确的,证明两个相邻的搜索方向一定正交。

证明：根据 $x_{k+1} = x_k + \alpha_k p_k$，一维搜索是精确的，

因此 α_k 是 $\min f(x_k + \alpha p_k)$ 的最优解。

若令 $\varphi(\alpha) = f(x_k + \alpha p_k)$，则 $\varphi'(\alpha_k) = 0$。

$\varphi'(\alpha) = (\nabla(x_k + \alpha p_k))^T p_k$，所以 $(\nabla(x_k + \alpha_k p_k))^T p_k = 0$，即 $g_{k+1}^T p_k = 0$

最速下降法求解时，每次选取的方向 $p_k = -g_k$，

所以 $p_{k+1}^T p_k = 0$，得证。

2. 设 H_k 是对称半正定但奇异的矩阵，从而存在某个向量 $u \ne 0$ 使得 $H_k u = 0$，证明 DFP 修正公式得到的 H_{k+1} 也是奇异的。

（提示：DFP 修正公式为：$H_{k+1} = H_k - \dfrac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \dfrac{s_k s_k^T}{y_k^T s_k}$.）

证明：对于奇异矩阵 H_k，存在向量 $u \ne 0$ 使得 $H_k u = 0$，

$$H_{k+1} u = H_k u - \frac{H_k y_k y_k^T H_k u}{y_k^T H_k y_k} + \frac{s_k s_k^T u}{y_k^T s_k} = \frac{s_k s_k^T u}{y_k^T s_k},$$

由于 $s_k = x_{k+1} - x_k = \alpha_k p_k = -\alpha_k H_k g_k$，

则 $s_k^T u = -\alpha_k g_k^T H_k^T u = -\alpha_k g_k^T H_k u = 0$，

从而 $H_{k+1} u = 0$，

因此，H_{k+1} 也是奇异矩阵。