文章编号:1000-6788(2005)09-0071-07

投影寻踪在机会约束规划中的应用研究

李寿德[1] ,安　凯[2]
(1. 上海交通大学管理学院, 上海 200052; 2. 山东航天电子技术研究所,山东 烟台 264000)

摘要: 投影寻踪方法应用于机会约束规划中,将随机变量约束下的集合 $X(\alpha)$、$X_i(\alpha_i)$ 转化为确定函数约束下的集合,从而将机会约束规划转变成一般的非线性规划问题,并对 $(b(w), A(w))$ 和 $(b(W), A(W))$ 的某一行向量的几种分布给出了 $X(\alpha)$、$X_i(\alpha_i)$ 凸性的结论.

关键词: 机会约束规划; 投影寻踪; 概率密度
中图分类号: O221　　**文献标识码**: A

Applications Studies of Projection Pursuit in the Chance Constrained Programming

LI Shou-de[1] , AN Kai[2]
(1. Management School, Shanghai Jiaotong University, Shanghai 200052, China; 2. Shandong Institute of Aerospace Electronic Technology, Yantai 264000, China)

Abstract The application of projection pursuit to the chance constrained programming, this paper convert the set of constrained stochastic variable $X(\alpha)$、$X_i(\alpha_i)$ to the set of constrained determined function, so as to convert the chance constrained programming to common nonlinear programming, and give the $X(\alpha)$、$X_i(\alpha_i)$ convex conclusion for some row vector geometry distribution of the $(b(w), A(w))$ and $(b(W), A(W))$.

Keywords chance constrained programming; projection pursuit; probability density

1 引言

机会约束规划问题主要有如下两种不同的形式,即或者是

$$\begin{aligned} &\min \phi(x) \\ &\text{s.t. } p(\{w \mid A(W)x \geq= b(W)\}) \geq \alpha \\ &\qquad x \in X_i \end{aligned} \tag{1.1}$$

或者

$$\begin{aligned} &\min \phi(x) \\ &\text{s.t. } p_w(\{w \mid A_i(w)x \geq b_i(w)\}) \geq \alpha_i \quad i = 1,2,\cdots,m, \\ &\qquad x \in X_0 \end{aligned} \tag{1.2}$$

其中 $A_i(w)$ 表示 $A(w)$ 的第 i 行, $b_i(w)$ 为 $b(w)$ 的第 i 个分量. 给定 ϕ 为一凸函数, X_0 为一凸集,则主要问题是确定集合

$$X(\alpha) = \{x \mid p_w(\{w \mid A(w)x \geq b(w)\}) \geq \alpha\} \tag{1.3}$$

和

$$X_i(\alpha_i) = \{x \mid p_w(\{w \mid A_i(w)x \geq b_i(w)\}) \geq \alpha_i\} \tag{1.4}$$

是否为凸集. 文献[1]中已对几种特殊的情形给出了 $X(\alpha)$、$X_i(\alpha_i)$ 凸性的结论,但并没有解决机会约束规

收稿日期:1999-11-16
资助项目:国家自然科学基金(70273021)

作者简介:李寿德(1964 -),男,青海人,博士,副教授,研究方向:环境管理;安凯(1957 -),男,山西人,博士,副研究员,研究方向:智能控制.

划向非线性规划的转化问题.本文将使用投影寻踪的方法来解决这一问题,并对随机系数矩阵 $A(w)$ 和它的某一行随机向量 $A_i(w)$ 的几种分布讨论 $X(\alpha)$、$X_i(\alpha_i)$ 的凸性.

2 主要结论及证明

在给出主要结论之前我们需要以下的一个关于投影寻踪方法的定义.

定义 1 设 ξ 是 P 维随机向量,其概率密度为 $f(x)$, $a_1, \cdots, a_k (k \le p)$ 是 R^p 中的线性独立的 k 个单位向量,$(a_1^{\mathrm T}\xi, \cdots, a_k^{\mathrm T}\xi)$ 的联合概率密度称为 $f(x)$ 在 $a_1, \cdots, a_k$ 方向上的投影密度,记作 $f^{a_1,\cdots,a_k}(t_1,\cdots,t_k)$.

若 A 是前 k 行分别为 $a_1^{\mathrm T}, \cdots a_k^{\mathrm T}$ 是非奇异矩阵,则容易求得:

$$f^{a_1,\cdots,a_k}(t_1,\cdots,t_k) = \int_{R^{p-K}} \frac{1}{|\det A|} f(A^{-1}T)\,\mathrm dt_{k+1}\cdots\mathrm dt_p. \tag{2.1}$$

有了这一定义,我们就可以将(1.3),(1.4)中的集合 $X(\alpha)$、$X_i(\alpha_i)$ 化为确定的函数约束下的集合.

定理 2.1 设 mp 维随机向量 $\xi(w) = (b_1(w), A_1(w), \cdots, b_m(w), A_m(w))^{\mathrm T} = (b_1(w), a_{11}(w), \cdots, a_{1,p-1}(w), \cdots, b_m(w), a_{m1}(w), \cdots, a_{m,p-1)}(w))^{\mathrm T}$ 的联合概率密度为 $f(x)$,则

$$\begin{aligned} X(\alpha) &= \{x \mid p_w(\{w \mid A(w)x \ge b(w)\}) \ge \alpha\} \\ &= \left\{ x \,\middle|\, \underbrace{\int_{-\infty}^{0}\cdots\int_{-\infty}^{0}}_{m重} f^{a_1,\cdots,a_m}(t_1,\cdots,t_m)\,\mathrm dt_1\cdots\mathrm dt_m \ge \alpha, \right. \\ &\left. a_i = (\underbrace{0\cdots0}_{(i-1)p个0} \frac{1}{\sqrt{1+\|x\|^2}}(1,-x^{\mathrm T}), \underbrace{0\cdots0}_{(m-i)p个0})^{\mathrm T},\ i = 1,2,\cdots,m \right\} \triangleq Y(\alpha), \end{aligned} \tag{2.2}$$

其中

$$A(w) = \begin{pmatrix} A_1(w) \\ \cdots \\ A_m(w) \end{pmatrix}, \quad b(w) = \begin{pmatrix} b_1(w) \\ \cdots \\ b_m(w) \end{pmatrix}.$$

证明 任取 $x \in X(\alpha)$,设 $x = \begin{pmatrix} x_1 \\ \cdots \\ x_{p-1} \end{pmatrix}$,则

$$\begin{aligned} \alpha &\le p_w(\{w \mid A(w)x \ge b(w)\}) \\ &= p_w(\{w \mid A_1(w)x \ge b_1(w), \cdots, A_m(w)x \ge b_m(w)\}) \\ &= p_w(\{w \mid b_1 - x_1 a_{11}(w) - \cdots - x_{p-1}a_{1,p-1}(w) \le 0, \cdots, \\ &\qquad b_m(w) - x_1 a_{m1}(w) - \cdots - x_{p-1}a_{m,p-1}(w) \le 0\}) \\ &= p_w\left(\left\{ w \,\middle|\, \frac{1}{\sqrt{1+\|x\|^2}} b_1(w) + \left(\frac{-x_1}{\sqrt{1+\|x\|^2}}\right) a_{11}(w) + \cdots + \right.\right. \\ &\qquad \left(\frac{-x_{p-1}}{\sqrt{1+\|x\|^2}}\right) a_{1,p-1}(w) \le 0, \cdots, \frac{1}{\sqrt{1+\|x\|^2}} b_m(w) + \\ &\qquad \left.\left(\frac{-x_1}{\sqrt{1+\|x\|^2}}\right) a_{m1}(w) + \cdots + \left(\frac{-x_{p-1}}{\sqrt{1+\|x\|^2}}\right) a_{m,p-1}(w) \le 0 \right\} \\ &= p_w(\{w \mid a_1^{\mathrm T}\xi(w) \le 0, \cdots, a_m^{\mathrm T}\xi(w) \le 0\}) \\ &= \underbrace{\int_{-\infty}^{0}\cdots\int_{-\infty}^{0}}_{m重} f^{a_1,\cdots,a_m}(t_1,\cdots,t_m)\,\mathrm dt_1\cdots\mathrm dt_m. \end{aligned} \tag{2.3}$$

这里 $f^{a_1,\cdots,a_m}$ 是 $f(x)$ 在 $a_1, \cdots, a_m$ 方向上的投影密度.由(2.3)知 $x \in Y(\alpha)$,再由 $x \in X(\alpha)$ 的任意性即知 $X(\alpha) \subset Y(\alpha)$.又由于(2.3)的每一等式都是可逆的,因此同法可证 $X(\alpha) \supset Y(\alpha)$.结合起来便得出要证明的结论.

定理 2.2　设 $a_1, \cdots, a_m$ 如(2.2)所示,若 $f^{a_1, \cdots, a_m}(t_1, \cdots, t_m)$ 是 x 的凹函数,则 $X(\alpha)$ 是凸集.

证明　由定理 2.1 只需证 $Y(\alpha)$ 是凸集即可.

任取 $x, y \in Y(\alpha)$,设

$$a_i' = \Big(\underbrace{0 \cdots 0}_{(i-1)p\text{个}}, \frac{1}{\sqrt{1 + \|x\|^2}}(1, -x^{\mathrm{T}}), \underbrace{0 \cdots 0}_{(m-i)p\text{个}}\Big),$$

$$a_i'' = \Big(\underbrace{0 \cdots 0}_{(i-1)p\text{个}}, \frac{1}{\sqrt{1 + \|x\|^2}}(1, -y^{\mathrm{T}}), \underbrace{0 \cdots 0}_{(m-i)p\text{个}}\Big),$$

$$i = 1, 2, \cdots, m.$$

则对于 $0 < \lambda < 1$,令 $a_i''' = \Big(\underbrace{0 \cdots 0}_{(i-1)p\text{个}}, \frac{1}{\sqrt{1 + \|\lambda x + (1-\lambda) y\|^2}}(1, -\lambda x^{\mathrm{T}} - (1-\lambda) y^{\mathrm{T}}), \underbrace{0 \cdots 0}_{(m-i)p\text{个}}\Big)$

由于 $f^{a_1, \cdots, a_m}$ 是 x 的凹函数,故

$$\begin{aligned}
&\int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f^{a_1''', \cdots, a_m'''}(t_1, \cdots, t_m)\,\mathrm{d}t_1 \cdots \mathrm{d}t_m \\
&= \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} [\lambda f^{a_1', \cdots, a_m'}(t_1, \cdots, t_m) + (1-\lambda) f^{a_1''', \cdots, a_m'''}(t_1, \cdots, t_m)]\,\mathrm{d}t_1 \cdots \mathrm{d}t_m \\
&= \lambda \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f^{a_1', \cdots, a_m'}(t_1, \cdots, t_m)\,\mathrm{d}t_1 \cdots \mathrm{d}t_m \\
&\quad + (1-\lambda) \int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f^{a_1'', \cdots, a_m''}(t_1, \cdots, t_m)\,\mathrm{d}t_1 \cdots \mathrm{d}t_m \\
&\geq \lambda \alpha + (1-\lambda)\alpha = \alpha.
\end{aligned}$$

故 $\lambda x + (1-\lambda) y \in Y(\alpha)$,再由 $x, y \in Y(\alpha)$ 的任意性即知 $Y(\alpha)$ 是凸集,从而利用定理 2.1 知 $X(\alpha)$ 是凸集.

推论 2.1　设 $a_1, \cdots, a_m$ 如(2.2)所示,若积分 $\int_{-\infty}^{0} \cdots \int_{-\infty}^{0} f^{a_1, \cdots, a_m}(t_1, \cdots, t_m)\,\mathrm{d}t_1 \cdots \mathrm{d}t_m = \beta$ 是一与 x 无关的常数,则当 $\alpha \leq \beta$ 时 $X(\alpha) = R^{p-1}$ 是凸集;当 $\alpha > \beta$ 时 $X(\alpha) = \phi$.

这是定理 2.1 的直接推论,证明从略.

在以上的结论中若取 $m = 1$ 则可得到 $X_i(\alpha_i)$ 的相应的结论. 以下我们给出这些结论,它们均为本节前面几个结论的直接推论.

定理 2.3　设 p 维随机向量 $\xi(w) = (b_i(w), A_i(w))^{\mathrm{T}} = (b_i(w), a_{i1}(w), \cdots, a_{i,p-1}(w))^{\mathrm{T}}$ 的联合概率密为 $f(x)$,则

$$\begin{aligned}
X_i(\alpha_i) &= \{x \mid p_w(\{w \mid A_i(w) x \geq b_i(w)\}) \geq \alpha_i\} \\
&= \left\{ x \;\middle|\; \int_{-\infty}^{0} f^a(f^a(t)\,\mathrm{d}t) \geq \alpha;\ a = \frac{1}{\sqrt{1 + \|x\|^2}}\begin{pmatrix} 1 \\ -x \end{pmatrix} \right\},
\end{aligned} \tag{2.4}$$

定理 2.4　若 $f^a(t)\left(a = \frac{1}{\sqrt{1 + \|x\|^2}}\begin{pmatrix} 1 \\ -x \end{pmatrix}\right)$ 是 x 的凹函数,则 $X_i(\alpha_i)$ 是凸集.

推论 2.2　若 $\int_{-\infty}^{0} f^a(t)\,\mathrm{d}t \left(a = \frac{1}{\sqrt{1 + \|x\|^2}}\begin{pmatrix} 1 \\ -x \end{pmatrix}\right) = \beta$ 是一与 x 无关的常数,则当 $\alpha_i \leq \beta$ 时 $X_i(\alpha_i) = R^{p-1}$ 是凸集;当 $\alpha_i > \beta$ 时,$X_i(\alpha_i) = \phi$.

3　定理的具体应用

在本节我们将使用上节的结论,对 $(b(w), A(w))$ 和它的某一行随机向量 $(b_i(w), A_i(w))$ 的几种分布探讨 $X(\alpha)$ 和 $X_i(\alpha_i)$ 的凸集.

3.1　独立柯西分布的情况.

结论 3.1.1　设 p 维随机向量 $\xi(w) = (b_i(w), A_i(w))^{\mathrm{T}} = (b_i(w), a_{i1}(w), \cdots, a_{i,p-1}(w))^{\mathrm{T}}$ 的联合概率密度为 $f(y) = \prod_{j=1}^{p} \frac{\lambda_j}{\pi} \frac{1}{\lambda_j^2 + (y_i - \mu_j)^2}$, $-\infty < \mu_j < \infty$, $\lambda_j > 0$,则当 $\alpha_i \geq \frac{1}{2}$ 时 $X_i(\alpha_i)$ 为凸集;当 $\alpha_i =$

0 时 $X_i(\alpha_i) = R^{p-1}$. 若 $p > 2$,则当 $0 < \alpha_i < \frac{1}{2}$ 时 $X_i(\alpha_i) = R^{p-1}$,或者 $X_i(\alpha_i)$ 为非空非凸集.

证明 对任一 p 维单位向量 $a = (a_1, \cdots, a_p)^{\mathrm{T}}$ 可求得 $a^{\mathrm{T}}\xi(\mu)$ 的特征函数为 $\exp\left\{ i\sum_{j=1}^{p} a_j\mu_j t - \sum_{j=1}^{p}|a_j\lambda_j||t|\right\}$,因此,$f^a(t) = \frac{\sum_{j=1}^{p}|a_j\lambda_j|}{\pi}\left[\left(\sum_{j=1}^{p}|a_j\lambda_j|\right)^2 + \left(t - \sum_{j=1}^{p}a_j\mu_j\right)^2\right]^{-1}$.

$$\begin{aligned}\int_{-\infty}^{0} f^a(t)\,\mathrm{d}t &= \int_{-\infty}^{0}\frac{\sum_{j=1}^{p}|a_j\lambda_j|}{\pi}\left[\left(\sum_{j=1}^{p}|a_j\lambda_j|\right)^2 + \left(t - \sum_{j=1}^{p}a_j\mu_j\right)^2\right]^{-1}\mathrm{d}t \\ &= \int_{-\infty}^{-\sum_{j=1}^{p}a_j\mu_j}\frac{\sum_{j=1}^{p}|a_j\lambda_j|}{\pi}\left[\left(\sum_{j=1}^{p}|a_j\lambda_j|\right)^2 + t^2\right]^{-1}\mathrm{d}t \\ &= \int_{-\infty}^{-\sum_{j=1}^{p}a_j\mu_j/\sum_{j=1}^{p}|a_j\lambda_j|}\frac{1}{\pi}\frac{1}{1+t^2}\mathrm{d}t \\ &\triangleq C_i\left(-\frac{\sum_{j=1}^{p}a_j\mu_j}{\sum_{j=1}^{p}|a_j\lambda_j|}\right).\end{aligned} \tag{3.1}$$

故 $\int_{-\infty}^{0} f^a(t)\,\mathrm{d}t \geq \alpha_i$ 等价于 $C_i^{-1}(\alpha_i)\sum_{j=1}^{p}|a_i\lambda_i| + \sum_{j=1}^{p}a_j\mu_j \leq 0$,将 $a = \frac{1}{\sqrt{1+\|x\|^2}}\begin{pmatrix}1\\ -x\end{pmatrix}$ 代入不等式并在两端同乘 $\sqrt{1+\|x\|^2}$ 得:

$$C_i^{-1}(\alpha_i)\left(|\lambda_1| + \sum_{j=2}^{p}|\lambda_j x_{j-1}|\right) + \mu_1 - (\mu_2, \cdots, \mu_p)x \leq 0. \tag{3.2}$$

由于 $|\lambda_1| + \sum_{j=2}^{p}|\lambda_j x_{j-1}|$ 为 x 的凸函数,而 $(\mu_1, \cdots, \mu_p)x$ 也是 x 的凸函数,故当 $C_i^{-1}(\alpha_i) \geq 0$ 时 (3.2) 的左端为 x 的凸函数,而 $\int_{-\infty}^{0}\frac{1}{\pi}\frac{1}{1+t^2}\mathrm{d}t = \frac{1}{2}$,故 $C_i^{-1}(\alpha_i) \geq 0$ 等价于 $\alpha_i \geq \frac{1}{2}$,于是当 $\alpha_i \geq \frac{1}{2}$ 时 (3.2) 的左端是 x 的凸函数,即 $X_i(\alpha_i)$ 是凸集. $X_i(0) = R^{p-1}$ 是显然的.

若 $p > 2$,当 $0 < \alpha_i < \frac{1}{2}$ 时 $C_i^{-1}(\alpha_i) < 0$. 由于 $x \in X_i(\alpha_i)$ 当且仅当 (3.2) 成立,假设 $X_i(\alpha_i) \neq R^{p-1}$,则有 $x = \begin{pmatrix}x_1\\ \cdots\\ x_{p-1}\end{pmatrix} \notin X_i(\alpha_i)$. 因为 $p > 2$,故必有 $y = \begin{pmatrix}y_1\\ \cdots\\ y_{p-1}\end{pmatrix} \neq 0$,使 $(\mu_2, \cdots, \mu_p)y = 0$,取 λ_0 充分大使

$$\begin{aligned}|C_i^{-1}(\alpha_i)|\left(|\lambda_1| + \sum_{j=2}^{p}|\lambda_j(x_{j-1} \pm \lambda_0 y_{j-1})|\right) &> |\mu_1 - (\mu_2, \cdots, \mu_p)x| \\ &= |\mu_1 - (\mu_2, \cdots, \mu_p)(x \pm \lambda_0 y)|.\end{aligned}$$

由此知 $x + \lambda_0 y$, $x - \lambda_0 y$ 均满足 (3.2) 式. 故

$$x + \lambda_0 y \in X_i(\alpha_i), \qquad x - \lambda_0 y \in X_i(\alpha_i).$$

而
$$x = \frac{1}{2}(x + \lambda_0 y) + \frac{1}{2}(x - \lambda y) \notin X_i(\alpha_i).$$

因此 $X_i(\alpha_i)$ 是非空非凸集.

结论 3.1.2 设 mp 维随机向量 $\xi(w) = (b_1(w), A_1(w), \cdots, b_m(w), A_m(w))^{\mathrm{T}} = (b_1(w), a_{11}(w), \cdots, a_{1,p-1}(w), \cdots, b_m(w), a_{m1}(w), \cdots, a_{m,p-1}(w))^{\mathrm{T}}$ 的联合概率密度为 $f(y) = \prod_{j=1}^{mp}\frac{1}{\pi}\frac{1}{1+y_j^2}$. 则当 $\alpha \leq \frac{1}{2^m}$ 时

$X(\alpha) = R^{p-1}$是凸集;当 $\alpha > \frac{1}{2^m}$时 $X(\alpha) = \phi$.

证明　由结论 3.1 的证明过程可知

$$f^{a_1, \cdots, a_m}(t_1, \cdots, t_m) = \prod_{j=1}^{m}\left(\pi\sum_{j=1}^{p}|\beta|\right)^{-1}\left(1+\left(\sum_{j=1}^{p}|\beta|\right)^{-2}t_j^2\right)^{-1}$$

这里的 $a_1, \cdots, a_m$ 如(2.2)所示,$\begin{pmatrix}\beta_1\\ \cdots\\ \beta_p\end{pmatrix} = \frac{1}{\sqrt{1+\|x\|^2}}\begin{pmatrix}1\\ \cdots\\ -x\end{pmatrix}$,故

$$\int_{-\infty}^{0}\cdots\int_{-\infty}^{0} f^{a_1, \cdots, a_m}(t_1, \cdots, t_m)\,\mathrm{d}t_1\cdots\mathrm{d}t_m$$
$$= \prod_{j=1}^{m}\int_{-\infty}^{0}\left(\pi\sum_{j=1}^{p}|\beta|\right)^{-1}\left(1+\left(\sum_{j=1}^{p}|\beta|\right)^{-2}t_j^2\right)^{-1}$$
$$= \prod_{j=1}^{m}\int_{-\infty}^{0}\frac{1}{\pi(1+t^2)}\mathrm{d}t$$
$$= \frac{1}{2^m}.$$

由推论 2.1 知,当 $\alpha \leq \frac{1}{2^m}$时 $X(\alpha) = R^{p-1}$是凸;当 $\alpha > \frac{1}{2^m}$时 $X(\alpha) = \varnothing$.

3.2　椭球等高分布的情况

结论 3.2　设 p 维随机向量 $\xi(w) = (b_i(w), A_i(w))^{\mathrm{T}} = (b_i(w), a_{i1}(w), \cdots, a_{i,p-1}(w))^{\mathrm{T}}$ 的联合概率密度为 $f(y) = (\det A)^{-\frac{1}{2}} g[(y-\mu)^{\mathrm{T}}A^{-1}(y-\mu)]$,其中 $A > 0$, $g(\cdot) \geq 0$,令 $G_1(t) = \int_{-\infty}^{t}\left[\int_{R^{p-1}} g(\|w\|^2)\,\mathrm{d}w_2\cdots\mathrm{d}w_p\right]\mathrm{d}w_1$,则当 $G_1^{-1}(\alpha_i) \geq 0$时 $X_i(\alpha_i)$ 是凸集 . 当 $\alpha_i = 0$ 时, $X_i(\alpha_i)$ R^{p-1};若 $p > 2$,则当 $-\infty < G_1^{-1}(\alpha_i) < 0$ 时 $X_i(\alpha_i) = R^{p-1}$或为非空非凸集.

证明　由于 $A > 0$,不防设 $A^{-1} = H'H$, 作代换 $Z = H(y-\mu)$,由 Z 的密度为 $g(\|z\|^2)$. 对任何单位向量 $a \in R^p$,以 a^T 为第一行构造正交阵 B,则 $W = BZ$ 的概率密度为 $g(\|w\|^2)$. 因此 a^TZ 的概率密度为

$$\int_{R^{p-1}} g(\|w\|^2)\,\mathrm{d}w_2 - \mathrm{d}w_0 \triangleq g_1(w_1^2),$$

而 $a^{\mathrm{T}}Y = a^{\mathrm{T}}H^{-1}Z + a^{\mathrm{T}}\mu = \|a^{\mathrm{T}}H^{-1}\|\frac{a^{\mathrm{T}}H^{-1}}{\|a^{\mathrm{T}}H^{-1}\|}Z + a^{\mathrm{T}}\mu$,由前面的推导$\frac{a^{\mathrm{T}}H^{-1}}{\|a^{\mathrm{T}}H^{-1}\|}Z$ 的概率密度为 $g_1(t^2)$,故 $\|a^{\mathrm{T}}H^{-1}\|\frac{a^{\mathrm{T}}H^{-1}}{\|a^{\mathrm{T}}H^{-1}\|}Z$ 的概率密度为 $\|a^{\mathrm{T}}H^{-1}\|g_1\left[\left(\frac{a}{\|a^{\mathrm{T}}H^{-1}\|}\right)^2\right]$. 从而 $a^{\mathrm{T}}Y$ 的概率密度为 $\|a^{\mathrm{T}}H^{-1}\|g_1\left[\left(\frac{t-a^{\mathrm{T}}\mu}{\|a^{\mathrm{T}}H^{-1}\|}\right)^2\right]$,而 $\|a^{\mathrm{T}}H^{-1}\|^2 = (a^{\mathrm{T}}H^{-1})(a^{\mathrm{T}}H^{-1})' = a^{\mathrm{T}}Aa$. 故

$$f^a(t) = \|a^{\mathrm{T}}H^{-1}\|g_1\left[\left(\frac{t-a^{\mathrm{T}}\mu}{\|a^{\mathrm{T}}H^{-1}\|}\right)^2\right]$$
$$= \frac{1}{\sqrt{a^{\mathrm{T}}Aa}}g_1\left[\left(\frac{t-a^{\mathrm{T}}\mu}{\sqrt{a^{\mathrm{T}}Aa}}\right)^2\right],$$
$$\int_{-\infty}^{0} f^a(t) = \frac{1}{\sqrt{a^{\mathrm{T}}Aa}}\int_{-\infty}^{0} g_1\left[\left(\frac{t-a^{\mathrm{T}}\mu}{\sqrt{a^{\mathrm{T}}Aa}}\right)^2\right]\mathrm{d}t$$
$$= \int_{-\infty}^{-\frac{a^{\mathrm{T}}\mu}{\sqrt{a^{\mathrm{T}}Aa}}} g_1(t^2)\,\mathrm{d}t$$
$$\triangleq G_1\left(-\frac{a^{\mathrm{T}}\mu}{\sqrt{a^{\mathrm{T}}A\mu}}\right),$$

故 $\int_{-\infty}^{0} f^{a}(t)\,\mathrm{d}t \geq \alpha_i$ 等价于 $G_1\left(-\frac{a^{\mathrm T}\mu}{\sqrt{a^{\mathrm T}A\mu}}\right) \geq \alpha_i$ 或 $-\frac{a^{\mathrm T}\mu}{\sqrt{a^{\mathrm T}A\mu}} \geq G_1^{-1}(\alpha_i)$，即 $G_1^{-1}(\alpha_i)\sqrt{a^{\mathrm T}Aa} + a^{\mathrm T}\mu \leq 0$.

将 $a = \frac{1}{\sqrt{1+\|x\|^2}}\begin{pmatrix}1\\ -x\end{pmatrix}$ 代入并在不等式的两端同乘 $\sqrt{1+\|x\|^2}$ 得：

$$G_1^{-1}(\alpha_i)\sqrt{(1,-x)^{\mathrm T}A\begin{pmatrix}1\\ -x\end{pmatrix}} + (1,-x)^{\mathrm T}\mu \leq 0. \tag{3.3}$$

容易求得，$(1,-x)^{\mathrm T}A\begin{pmatrix}1\\ -x\end{pmatrix}$ 的海赛矩阵为 A 除去第一行后的矩阵，而 A 正定. 故此海赛矩阵的行列式大于零，从而 $(1,-x)^{\mathrm T}A\begin{pmatrix}1\\ -x\end{pmatrix}$ 为凸函数. 因此 $\sqrt{(1,-x)^{\mathrm T}A\begin{pmatrix}1\\ -x\end{pmatrix}}$ 为 x 的凸函数，而 $(1,x)^{\mathrm T}\mu$ 也是 x 的凸函数(也是 x 的凹函数). 故当 $G_1^{-1}(\alpha_i) \geq 0$ 时满足(3.3)的 x 构成一凸集，即 $X_i(\alpha_i)$ 为凸集，而 $\alpha_i = 0$ 时 $X_i(0) = R^{p-1}$ 显然成立.

若 $p > 2$，且 $-\infty < G_1^{-1}(\alpha_i) < 0$，由于 $x \in X_i(\alpha)$，当且仅当(3.3)成立. 假设 $X_i(\alpha_i) \neq R^{p-1}$，则有 $x = \begin{pmatrix}1\\ \cdots\\ x_{p-1}\end{pmatrix} \bar{\in} X_i(\alpha_i)$，因为 $p > 2$，故必有 $y = \begin{pmatrix}1\\ \cdots\\ x_{p-1}\end{pmatrix} \neq 0$. 使 $(\mu_2, \cdots, \mu_p)y = 0$，取 λ_0 充分大使

$$-G_1^{-1}(\alpha_i)\sqrt{(1,-(x\pm\lambda_0 y))^{\mathrm T}A\begin{pmatrix}1\\ -(x\pm\lambda_0 y)\end{pmatrix}} > |(1,-x)^{\mathrm T}\mu|$$
$$= |(1,-(x\pm\lambda_0 y))^{\mathrm T}\mu|.$$

由此知 $x+\lambda_0 y$，$x-\lambda_0 y$ 均满足(3.3)式. 故

$$x+\lambda_0 y \in X_i(\alpha_i), \qquad x-\lambda_0 y \in X_i(\alpha_i),$$

而

$$x = \frac{1}{2}(x+\lambda_0 y) + \frac{1}{2}(x-\lambda_0 y) \;\bar{\in}\; X_i(\alpha_i).$$

因此 $X_i(\alpha_i)$ 是非空非凸集.

若在结论 3.2 中取 $g(x) = (2\pi)^{-\frac{p}{2}}\mathrm{e}^{-\frac{x}{2}}$ 则得到文献[1]中的一个定理：

推论 3.2.1　设 p 维随机向量 $\xi(w) = (b_i(w), A_i(w))^{\mathrm T} = (b_i(w), a_{i1}(w), \cdots, a_{i,p-1}(w))^{\mathrm T}$ 的联合概率密度为 $f(y) = (\det A)^{-\frac{1}{2}}(2\pi)^{-\frac{p}{2}}\mathrm{e}^{-\frac{(y-\mu)^{\mathrm T}A^{-1}(y-\mu)}{2}}$，则对于 $\alpha_i \geq \frac{1}{2}$，$X_i(\alpha_i)$ 为凸集；$\alpha_i = 0$ 时 $X_i(0) = R^{p-1}$；若 $p > 2$，则当 $0 < \alpha_i < \frac{1}{2}$ 时或者 $X_i(\alpha_i) = R^{p-1}$，或者 $X_i(\alpha_i)$ 为非凸集.

证明　在结论 3.2 中取 $g(x) = (2\pi)^{-\frac{p}{2}}\mathrm{e}^{-\frac{x}{2}}$，则

$$G_1(t) = \int_{-\infty}^{t}\left[\int_{R^{p-1}}(2\pi)^{-\frac{p}{2}}\mathrm{e}^{-\frac{t_1^2+\cdots+t_p^2}{2}}\mathrm{d}t_2\cdots\mathrm{d}t_p\right]\mathrm{d}t_1$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}\mathrm{e}^{-\frac{t_1^2}{2}}\mathrm{d}t_1$$

.

故 $G_1^{-1}(\alpha_i) \geq 0$ 等价于 $\alpha_i \geq \frac{1}{2}$，$0 < G_1^{-1}(\alpha_i) < 0$ 等价于 $0 < \alpha_i < \frac{1}{2}$，利用结论 3.2 即得所要证的结论.

若在结论 3.2 中取 $\mu = 0$，A 为单位阵，$g(t) = C_p(1+t)^{-k}$，$k > \frac{p}{2}$，C_p 是使 $C_p(1+\|y\|^2)^{-k}$ 构成 p 维概率密度的常数，则可得如下结论：

推论 3.2.2　设 p 维随机向量 $\xi(w) = (b_i(w), A_i(w))^T = (b_i(w), a_{i1}(w), \cdots, a_{i,p-1}(w))^T$ 的联合概率密度为 $f(y) = C_p(1+\|y\|^2)^{-k}, k > \frac{p}{2}$,则当 $\alpha_i \geq \frac{1}{2}$ 时 $X_i(\alpha_i)$ 是凸集;$\alpha_i = 0$ 时 $X_i(0) = R^{p-1}$;若 $p > 2$,则当 $0 < \alpha_i < \frac{1}{2}$ 时,或者 $X_i(\alpha_i) = R^{p-1}$,或者 $X_i(\alpha_i)$ 为非空非凸集.

证明　在结论 3.2 中取 $\mu = 0$,A 为单位阵,$g(t) = C_p(1+t)^{-k}$,$\left(k > \frac{p}{2}\right)$,$C_p$ 是使 $C_p(1+\|y\|^2)^{-k}$ 构成 p 维概率密度的常数.则

$$
\begin{aligned}
G_1(t) &= \int_{-\infty}^{t}\left[\int_{R^{p-1}} g(\|y\|^2)\,\mathrm{d}y_2, \cdots, \mathrm{d}y_p\right]\mathrm{d}y_1 = \int_{-\infty}^{t}\left[\int_{R^{p-1}} C_p(1+y_1^2+\cdots+y_p^2)^{-k}\mathrm{d}y_2\cdots\mathrm{d}y_p\right]\mathrm{d}y_1 \\
&= C_p\int_{-\infty}^{t}\int_{R^{p-1}}(1+y_1^2+r^2)^{-k}r^{p-2}\mathrm{d}r\cdot C_0 \\
&= C_pC_0^*\int_{-\infty}^{t}\int_0^{\infty}(1+y_1^2+Z)^{-k}Z^{\frac{p-2}{2}-\frac{1}{2}}\mathrm{d}Z \\
&= C_pC_0^*\int_{-\infty}^{t}(1+y_1^2)^{-k+\frac{p-1}{2}}\mathrm{d}y_1\int_0^{\infty}Z^{\frac{p-1}{2}-1}(1+Z)^{-k}\mathrm{d}Z \\
&= C_p\int_{-\infty}^{t}(1+y_1^2)^{-\frac{2k+1-p}{2}}\mathrm{d}y_1,
\end{aligned}
\tag{3.4}
$$

其中 $C_p = \left(\int_R(1+y_1^2)^{-\frac{2k+1-p}{2}}\mathrm{d}y_1\right)^{-1}$.

由(3.4)式右端被积函数关于 0 的对称性可知 $G_1(0) = \frac{1}{2}$,即 $G_1^{-1}\left(\frac{1}{2}\right) = 0$.故 $G_1^{-1}(\alpha_i) \geq 0$ 等价于 $\alpha_i \geq \frac{1}{2}$,$-\infty < G_1^{-1}(\alpha_i) < 0$ 等价于 $0 < \alpha_i < \frac{1}{2}$,利用结论 3.2 即得所要证的结论.

参考文献:

[1]　P·卡尔.随机线性规划[M].王金德译.上海:上海科学技术出版社,1988.
P·Karer. Stochastic Linear Programming[M]. Wang Jin-de,translator. Shanghai:Shanghai Science and Technology Press,1988.

[2]　Cheng Ping. Optimum initial condition of projection pursuit approximation[J]. Acta Mathematicae Applicatae Sinica,Beijing,1985,2(3):269-280.

[3]　成平.投影寻踪经验分布的极限分布[J].应用概率统计,1987,3(1):8-20.
Cheng Ping. The limiting distribution of projection pursuit empirical distribution[J]. Applied Probability Statistics,1987,3(1):8-20.