第32卷第1期 上海师范大学学报(自然科学版) Vol.32, No.1
2003年3月 Journal of Shanghai Teachers University(Natural Sciences) 2003, Mar.

两类熵密度规划的对偶规划和最优性条件

张乐瑛，朱德通

(上海师范大学 数理信息学院，上海 200234)

摘　要：基于广义的 Fenchel 对偶定理及其相应的 Kuhn-Tucker 条件，给出了带有二次约束和熵密度约束的二次规划问题和熵密度问题的对偶规划，强对偶定理以及 Kuhn-Tucker 条件.

关键词：广义 Fenchel 对偶定理；Kuhn－Tucker 条件；熵密度问题

中图分类号：O221.2　　**文献标识码**：A　　**文章编号**：1000-5137(2003)01-0021-06

1 介　绍

在[1]中，朱德通使用导出的广义 Fenchel 定理，推导了一类仅带二次约束的熵密度规划和 Kuhn-Tucker 条件. 最近，更一般的带有二次约束和熵密度约束的两类熵密度规划和二次规划在许多领域中有着新的应用，特别是在经济、金融领域，所以引起了关注. 其中一类是带二次约束和熵密度约束的二次规划问题，称之为问题 (Q)：

$$(Q)\quad \min\quad Q_0(z) = \frac{1}{2}z^TH_0z + h_0^Tz + c_0 \tag{1.1}$$

$$\text{s t}\quad P_j(z) = \sum_{k=1}^{n} z_k\cdot\ln\Big(\frac{z_k}{ed_{jk}}\Big) \leqslant 0,\quad j = 1,\dots,l; \tag{1.2}$$

$$Q_i(z) = \frac{1}{2}z^TH_iz + h_i^Tz + c_i \leqslant 0,\quad i = 1,\dots,r; \tag{1.3}$$

$$z = (z_1,\dots,z_n)^T \geqslant 0; \tag{1.4}$$

这里 $H_i(i = 0,\dots,r)$ 是对称半正定的 n 矩阵，列向量 $h_i, z\in\mathscr{R}^n$，$c_i(i = 0,1,\dots,r)$ 和 $d_{jk}\geqslant 0\ (k = 1\dots, n; j = 1,\dots,l)$ 均为实数，并且规定 $0\times\ln 0 = 0$，这样问题 (Q) 在 $z\geqslant 0$ 是定义明确的. e 是自然对数的基，其目的只是为了简便其对偶的表示形式.

另一类是带熵信息和二次约束的最小区别信息量问题(简称 MDI 问题)，记为问题 (P)：

$$(P)\quad \min\quad P_0(z) = \sum_{k=1}^{n} z_k\cdot\ln\Big(\frac{z_k}{ed_{0k}}\Big), \tag{1.5}$$

$$\text{s t}\quad P_j(z) = \sum_{k=1}^{n} z_k\cdot\ln\Big(\frac{z_k}{ed_{jk}}\Big) \leqslant 0,\quad j = 1,\dots,l; \tag{1.6}$$

$$Q_i(z) = \frac{1}{2}z^TH_iz + h_i^Tz + c_i \leqslant 0,\quad i = 1,\dots,r; \tag{1.7}$$

$$z = (z_1,\dots,z_n)^T \geqslant 0.$$

收稿日期：2002-09-05

基金项目：上海市教委基金资助(020K06)；上海市科委自然科学基金资助(02ZA14070)

作者简介：张乐瑛(1979-)，女，上海师范大学数理信息学院硕士研究生. 朱德通(1954-)，男，上海师范大学数理信息学院教授.

这两类问题在许多领域都有着广泛地应用,特别是它们的性质、求解技巧性有着重要的应用背景,由于其规划特殊性,其对偶形式求解显得十分重要,许多方法的推导需要的技巧性都很强,而且不容易用显式表示其对偶,本文使用凸分析的知识以及广义 Fenchel 对偶定理[1] 来导出两类问题的对偶规划,强对偶定理以及相应的 Kuhn-Tucker 条件.

2 两类问题的对偶

在此节中,先引入[1]中朱德通推导的广义 Fenchel 对偶定理及其相关推论:

定理2.1 设f是$\mathscr{R}$上的正常凸函数, $g_j(j = 1, \dots, r)$ 是$\mathscr{R}^{m_j}$ $(j = 1, \dots, r)$ 上的正常凹函数, 且A_j $(j = 1, \dots, r)$ 是$\mathscr{R}^n \to \mathscr{R}^{m_j}$ $(j = 1, \dots, r)$ 的线性变换. 定义双函数

$$(Fu)(x) = f(x) - \sum_{j=1}^{r} g_j(A_j x + u_j) \quad \forall\ x \in \mathscr{R}^n,\ \forall\ u \in \mathscr{R}^m$$

其中 $u = (u_1, \dots, u_r)$, $u_j \in \mathscr{R}^{m_j}$, $j = 1, \dots, r$, $m = \sum_{j=1}^{r} m_j$, 则 F 是关于 u 的正常凹函数, 若 f 和 g_j $(j = 1, \dots, r)$ 是闭的,那么 F 也是闭的,与 F 相应的凸规划(Q)最优值为

$$\inf_x \{f(x) - \sum_{j=1}^{r} g_j(A_j x)\} = \inf(F0)$$

且(Q)是闭相容的当且仅当存在 $x \in \mathrm{ri}(\mathrm{dom} f)$ 有 $A_j x \in \mathrm{ri}(\ \mathrm{dom}\ g_j)$. $j = 1, \dots, r$. F 的伴随函数是

$$(F^* x^*)(u) = \sum_{j=1}^{r} g_j^*(u_j^*) - f^*(x^* + \sum_{j=1}^{r} A_j^* u_j^*) \quad \forall\ u^* \in \mathscr{R}^m, \quad \forall\ x^* \in \mathscr{R}^n$$

和 F^* 相对应的对偶问题 (Q^*) 的最优值为 $\sup_{u^*}\{\sum_{j=1}^{r} g_j^*(u_j^*) - f^*(\sum_{j=1}^{r} A_j^* u_j^*)\} = \sup(F^* 0)$, 其中 $u^* = (u_1^*, \dots, u_r^*)$. 并且 (Q^*) 是强相容的当且仅当存在 $u^* = (u_1^*, \dots, u_r^*)$ 满足 $u_j^* \in \mathrm{ri}(\mathrm{dom}\ g_j^*), j = 1, \dots, r$, 有 $\sum_{j=1}^{r} A_j^* u_j^* \in \mathrm{ri}(\mathrm{dom} f^*)$.

进一步,当下列条件之一成立时,

(a) 存在 $x \in \mathrm{ri}(\mathrm{dom} f)$ 有 $A_j x \in (g_j)$ $(j = 1, \dots, r)$; 或者

(b) 存在 $u^* = (u_1^*, \dots, u_r^*)$, 满足 $u_j^* \in \mathrm{ri}(\mathrm{dom}\ g_j^*)$ $(j = 1, \dots, r)$ 有 $\sum_{j=1}^{r} A_j^* u_j^* \in \mathrm{ri}(\mathrm{dom} f^*)$; 都有 $\inf_x\{f(x) - \sum_{j=1}^{r} g_j(A_j x)\} = \sup_{u^*}\{\sum_{j=1}^{r} g^*(u_j^*) - f^*(\sum_{j=1}^{r} A_j^* u_j^*)\}$. 并且当(b)成立时,有上确界在 u^* 点达到; 当(a)成立有下确界在 x 点达到.

先将问题 (Q) 作适当变形以便运用上述定理,由于矩阵 H_i 半正定 $(i = 1, \dots, r)$, 则可表示为 $H_i = A_i^T A_i$ 的形式,其中 A_i 为行满秩的 $m_i \times n$ 矩阵, $m_i = \mathrm{rank}(A_i)$. 那么二次约束即可写成

$$Q_i = \frac{1}{2} z^T A_i^T A_i z + h_i (A_i^T A_i)^+ A_i^T A_i z + c_i \leqslant 0, \quad i = 1, \dots, r. \tag{2.1}$$

这里的 $(A_i^T A_i)^+$ 为 $A_i^T A_i$ 的广义逆. 从而, 问题(Q) 转化为更易于求解的形式:

$$(PQ) \quad \min \quad Q_0 = \frac{1}{2} z^T H_0 z + h_0^T z + c_0$$

$$\text{s t} \quad R_s(x_s) = \sum_{k=1}^{n} x_{sk} \cdot \ln(\frac{x_{sk}}{ed_{sk}}) \leqslant 0, \quad s = 1, 2, \dots, l; \tag{2.2}$$

$$R_s(x_s) = \frac{1}{2} \| x_s \|^2 + b_s^T x_s + c_s \leqslant 0, \quad s = l+1, \dots, l+r; \tag{2.3}$$

$$x_s = (x_{s1}, \dots, x_{sn})^T \geqslant 0, \quad s = 1, \dots, l, \dots, l+r;$$

其中 $x_s = B_s z$, 用

$$B_s = \begin{cases} I_n & 若 s = 1, \ldots, l, \\ A_{s-l} & 若 s = l+1, \ldots, l+r; \end{cases} \tag{2.4}$$

这里 I_n 是 n 维空间的单位矩阵, $b_{s-l}^T \overset{\text{def}}{=} h_{s-l}^T (A_{s-l}^T A_{s-l})^+ A_{s-l}^T$, $(s = l+1, \ldots, l+r)$, 即

$$((x_1)^T, \ldots, (x_l)^T, \ldots, (x_{l+r})^T)^T = (\underbrace{I_n, \ldots, I_n}_{l}, A_1^T, \ldots, A_r^T)^T z; \tag{2.5}$$

记集合 $\Omega_s \overset{\text{def}}{=} \{x_s \mid R_s(x_s) \leqslant 0\}$, $s = 1, \ldots, l, \ldots, l+r$; 则集合 Ω_s 的指示函数为

$$\delta(x_s \mid \Omega_s) = \begin{cases} 0 & 若 x_s \in \Omega_s; \\ +\infty & 否则; \end{cases} \quad s = 1, \ldots, l, \ldots, l+r; \tag{2.6}$$

令 $f(z) = Q_0(z)$, $g_s(x_s) = -\delta(x_s \mid \Omega_s)$; $s = 1, \ldots, l, \ldots, l+r$; 那么,由定理2.1问题 (Q) 转化为

$$\inf_z \{Q_0(z) + \sum_{s=1}^{l+r} \delta(B_s z \mid \Omega_s)\} = \inf_z \{f(z) - \sum_{s=1}^{l+r} g_s(B_s z)\}. \tag{2.7}$$

显然 $Q_0(z)$ 是正常凸函数,而 $g_s(x_s)$ 是正常凹函数, 那么由广义的 Fenchel 对偶定理,则(2.7) 的对偶形式为 $\sup_{x_s^*} \{ \sum_{s=1}^{l+r} g_s^*(x_s^*) - f^*(\sum_{s=1}^{l+r} B_s^* x_s^*) \}$; 现分别求 g_s, f 的对偶, 由共轭函数的定义有

$$g_s^*(x_s^*) = (-\delta)^*(x_s^* \mid \Omega_s) = -\delta^*(-x_s^* \mid \Omega_s), \quad s = 1, \ldots, l, \ldots, l+r; \tag{2.8}$$

对 s 来分情况讨论: 当 $s = 1, \ldots, l$ 时, 令 $r_k(x_{s,k}) \overset{\text{def}}{=} x_{s,k} \ln(\frac{x_{s,k}}{ed_{s,k}})$, 则 $R_s(x_s) = \sum_{k=1}^{n} x_{s,k} \cdot \ln(\frac{x_{s,k}}{ed_{s,k}}) \overset{\text{def}}{=} \sum_{k=1}^{n} r_k(x_{s,k})$; 因为 $x_{s,k} (k = 1, \ldots, n)$ 是单变量,依据共轭函数定义,有

$$R_s^*(x_s^*) = \sum_{k=1}^{n} r_k^*(x_{s,k}^*) = \sum_{k=1}^{n} \sup_{x_{s,k}} \{x_{s,k} \cdot x_{s,k}^* - x_{s,k} \cdot \ln \frac{x_{s,k}}{ed_{s,k}}\}$$

由微分法求极值计算 $r_k^*(x_{s,k}^*) = d_{sk} \cdot \exp\{x_{s,k}^*\}$. 从而 $R_s^*(x_s^*) = \sum_{k=1}^{n} d_{s,k} \cdot \exp\{x_{s,k}^*\} \overset{\text{def}}{=} d_s^T \cdot \exp\{x_s^*\}$. 为简便起见, 这里使用 $\exp\{x_s^*\}$ 表示向量 $(e^{x_{s,1}^*}, \ldots, e^{x_{s,n}^*})^T$, $d_s^T = (d_{s,1}, \ldots, d_{s,n})^T$ 则

$$\delta^*(-x_s^* \mid \Omega_s) = \inf_{\mu_s \geqslant 0} \mu_s \sum_{k=1}^{n} d_{s,k} \exp\{\frac{-x_{s,k}^*}{\mu_s}\} = \inf_{\mu_s \geqslant 0} \mu_s d_s^T \exp\{\frac{-x_s^*}{\mu_s}\} \quad s = 1, 2, \ldots, l; \tag{2.9}$$

而当 $s = l+1, \ldots, l+r$ 时, $R_s(x_s) = \frac{1}{2} \| x_s \|^2 + b_s^T x_s + c_s$, 由[1]可知 $R_s^*(x_s^*) = \frac{1}{2} \| x_s^* - b_s \|^2 - c_s$; 以及 $\delta^*(-x_s^* \mid \Omega_s) = \inf_{\lambda_s \geqslant 0} \{\frac{1}{2} \lambda_s \| \frac{-x_s^*}{\lambda_s} - b_s \|^2 - \lambda_s c_s\}$, $s = l+1, \ldots, l+r$. 现将 $\sum_{s=1}^{l+r} g_s^*(x_s^*)$ 分成 $\sum_{s=1}^{l} g_s^*(x_s^*)$ 和 $\sum_{s=l+1}^{l+r} g_s^*(x_s^*)$ 两类,可得

$$\sum_{s=1}^{l+r} g_s^*(x_s^*) = \sum_{s=1}^{l} -\delta^*(-x_s^* \mid \Omega_s) + \sum_{s=l+1}^{l+r} -\delta^*(-x_s^* \mid \Omega_s) =$$

$$-\sum_{s=1}^{l} \inf_{\mu_s \geqslant 0} \mu_s d_s^T \exp\{\frac{-x_s^*}{\mu_s}\} - \sum_{s=l+1}^{l+r} \inf_{\lambda_s \geqslant 0} \{\frac{1}{2} \lambda_s \| \frac{-x_s^*}{\lambda_s} - b_s \|^2 - \lambda_s c_s\} \tag{2.10}$$

令 $v_s = x_s^*$, $s = 1, \ldots, l$, $u_{s-l} = x_s^*$, $s = l+1, \ldots, l+r$; 则有

$$\sum_{j=1}^{l+r} g_s^*(x_s^*) = -\sum_{j=1}^{l} \inf_{\mu_j \geqslant 0} \mu_j d_j^T \exp\{\frac{-v_j}{\mu_j}\} - \sum_{i=1}^{r} \inf\{\frac{1}{2} \lambda_i \| \frac{-u_i}{\lambda_i} - b_i \|^2 - \lambda_i c_i\}. \tag{2.11}$$

再令 $-w_i = \frac{u_i}{\lambda_i} + b_i$, $i = 1, \ldots, r$; 即 $u_i = -\lambda_i(w_i + b_i)$, $i = 1, \ldots, r$; 那么

$$\sum_{s=1}^{l+r} g_s^*(x_s^*) = -\sum_{j=1}^{l} \inf_{\mu_j \geqslant 0} \mu_j d_j^T \exp\{\frac{-v_j}{\mu_j}\} - \sum_{i=1}^{r} \inf_{\lambda_i \geqslant 0} \{\frac{1}{2} \lambda_i w_i^T w_i - \lambda_i c_i\}. \tag{2.12}$$

对于 $Q_0^*(z^*)$，可令 $p(z) = \frac{1}{2}z^T H_0 z$, $q(z) = h_0^T z + c_0$. 据[2]中定理 12.3, 可得

$$p^*(z^*) = \begin{cases} \frac{1}{2}(z^*)^T H_0^+ z^*, & 若 z^* \in L \\ +\infty & 否则; \end{cases}$$

其中 L 为 H_0 的零空间的正交补, H_0^+ 为 H_0 的广义逆. 另外, 由[1]中(3.9)式可知

$$q^*(z^*) = \begin{cases} -c_0, & 若 z^* = h_0 \\ +\infty & 否则; \end{cases}$$

从而由[2]中的定理 16.4, 可得 $Q_0^*(z^*) = \frac{1}{2}(z^* - h_0)^T H_0^+ (z^* - h_0) - c_0$. 因为 B_s 是实数矩阵, 那么 $B_s^* = B_s^T$, 由(2.4)式类似于(2.12)的变量代换，可得

$$Q_0^*\Big(\sum_{j=1}^{l} v_j + \sum_{i=1}^{r} A_i^T u_i\Big) = Q_0^*\Big(\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i)\Big) = \tag{2.13}$$
$$\frac{1}{2}\Big(\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i) - h_0\Big)^T H_0^+\Big(\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i) - h_0\Big) - c_0$$

由(2.16)和(2.17), 问题 (Q) 的对偶形式 $\sup\limits_{x_s^*}\{\sum_{s=1}^{l+r} g_s^*(x_s^*) - f^*(\sum_{s=1}^{l+r} B_s^* x_s^*)\}$ 可推得

$$(DQ) \qquad \sup\Big\{\sum_{j=1}^{l} g_j^*(v_j) + \sum_{i=1}^{r} g_i^*(u_i) - Q_0^*\Big(\sum_{j=1}^{l} v_j + \sum_{i=1}^{r} A_i^T u_i\Big)\Big\} =$$
$$\sup_{w_i, v_j, \lambda_i, \mu_j \geq 0}\Big\{\sum_{j=1}^{l}\Big(-\mu_j d_j^T \exp\Big\{\frac{-v_j}{\mu_j}\Big\}\Big) + \sum_{i=1}^{r}\Big(-\frac{1}{2}\lambda_i w_i^T w_i + \lambda_i c_i\Big) - \tag{2.14}$$
$$\frac{1}{2}\Big(\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i) - h_0\Big)^T H_0^+\Big(\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i) - h_0\Big) + c_0\Big\}.$$

基于问题 (Q) 的对偶推导, 类似地, 将问题 (P) 可转化为:

$$(PP) \qquad \min \quad P_0 = \sum_{k=1}^{n} z_k \cdot \ln \frac{z_k}{ed_{0k}}. \tag{2.15}$$

$$\text{s t} \quad R_s(x_s) = \sum_{k=1}^{n} x_{sk} \cdot \ln \frac{x_{sk}}{ed_{sk}} \leqslant 0, \quad s = 1, 2, \ldots, l; \tag{2.16}$$

$$R_s(x_s) = \frac{1}{2}\|x_s\|^2 + b_s^T x_s + c_s \leqslant 0, \quad s = l+1, \ldots, l+r; \tag{2.17}$$

$$x_s = (x_{s1}, \ldots, x_{sn})^T \geqslant 0, \quad s = 1, \ldots, l, \ldots, l+r;$$

其他的条件同问题 (PQ) 一样, 只因目标函数不同, 故只要对目标函数进行推导即可. 比较问题(PQ), 令 $f(z) = P_0(z)$, $g_s(x_s) = -\delta(B_s \mid \Omega_s)$; $s = 1, 2, \ldots, l, \ldots, l+r$, 那么带约束的问题 (P) 转化为求解 $\inf\limits_z\{P_0(z) + \sum_{s=1}^{l+r}\delta(B_s z \mid \Omega_s)\} = \inf\limits_z\{f(z) - \sum_{s=1}^{l+r} g_s(B_s z)\}$; 显然此问题满足定理 2.1 的条件, 则其对偶为 $\sup\limits_{x_s^*}\{\sum_{s=1}^{l+r} g_s^*(x_s^*) - f^*(\sum_{s=1}^{l+r} B_s^* x_s^*)\}$. 从(2.11), 并由 (DQ) 中的变量代换, 可得

$$P_0^*\Big(\sum_{j=1}^{l} v_j + \sum_{i=1}^{r} A_i^T u_i\Big) = d_0^T \exp\Big\{\sum_{j=1}^{l} v_j - \sum_{i=1}^{r} \lambda_i A_i^T(w_i + b_i)\Big\}. \tag{2.18}$$

联合(2.12)和(2.18), 问题 (P) 的对偶为

$$(DP) \qquad \sup\Big\{\sum_{j=1}^{l} g_j^*(v_j) + \sum_{i=1}^{r} g_i^*(u_i) - P_0^*\Big(\sum_{j=1}^{l} v_j + \sum_{i=1}^{r} A_i^T u_i\Big)\Big\} =$$
$$\sup_{w_i, v_j, \lambda_i, \mu_j \geq 0}\Big\{\sum_{j=1}^{l}\Big(-\mu_j d_j^T \exp\Big\{\frac{-v_j}{\mu_j}\Big\}\Big) + \tag{2.19}$$

$$\sum_{i=1}^{r}(-\frac{1}{2}\lambda_i w_i^T w_i - \lambda_i c_i) - d_0^T \exp\{\sum_{j=1}^{l} v_j - \sum_{i=1}^{r}\lambda_i A_i^T(w_i + b_i)\}\Big\}$$

从上可见，利用广义的 Fenchel 对偶定理能较简单、快速地导出问题 (P) 和 (Q) 的对偶形式且对偶形式只含有变量的非负约束，以后遇到类似的问题可以考虑用此种方法来解决.

3　强对偶定理及 Kuhn-Tucker 条件

基于广义的 Fenchel 对偶定理及其相关定理，本节推出这两类问题的广义 Kuhn-Tucker 条件和无对偶间隙的强对偶定理. 设 φ_1 为问题 (Q) 的最优值，ψ_1 为其对偶问题(DQ)的上确界；设 φ_2 为问题 (P) 的最优值，ψ_2 为其对偶问题(DP) 的上确界.

定理 3.1　(强对偶定理) 若问题 (Q) 可行，则 $\varphi_1 = \psi_1$，因而不存在对偶间隙(证明见[2]).

定理 3.2　(强对偶定理) 若问题 (P) 可行，则 $\varphi_2 = \psi_2$，因而不存在对偶间隙(证明见[2]).

定理 3.3　假设问题 (Q) 有一个 $z \geqslant 0$ 的可行解，那么问题 (Q) 的基于广义 Fenchel 对偶定理的广义 Kuhn-Tucker 条件为常义下的 Kuhn-Tucker 条件.

证明　由[3]中引理3，在最优解 $\hat{z}$ 处，$Q_0(z)$ 是可导的，又据[1]中定理2.2 的(2.8)式，则存在 $\hat{u}_i, \hat{v}_j$，使得 $\sum_{j=1}^{l}\hat{v}_j + \sum_{i=1}^{r}A_i^T\hat{u}_i \in \partial f(\hat{z})$，使用导数分量的表示形式有

$$(H_0\hat{z} + h_0)_s - (\sum_{j=1}^{l}\hat{v}_j + \sum_{i=1}^{r}A_i^T\hat{u}_i)_s = 0, \quad s = 1, 2, \ldots, n \tag{3.1}$$

至于[1]中定理2.2 的(2.8)式的另外一个表达式，由 δ^* 可导，则从上节的变量代换，得

$$\partial g_j^*(\hat{v}_j) = -\nabla\delta^*(-\hat{v}_j \mid \Omega_j),\ j = 1, \ldots, l; \quad \partial g_i^*(\hat{u}_i) = -\nabla\delta^*(-\hat{u}_i \mid \Omega_i),\ i = 1, \ldots, r;$$

由(2.15)知

$$\delta^*(-\hat{v}_j \mid \Omega_j) = \inf_{\mu_j \geqslant 0}\mu_j \cdot \sum_{k=1}^{n} d_{jk} \cdot \exp\{\frac{-\hat{v}_{jk}}{\mu_j}\},\ j = 1, 2, \ldots, l; \tag{3.2}$$

$$\delta^*(-\hat{u}_i \mid \Omega_i) = \inf_{\lambda_i \geqslant 0}\{\frac{1}{2}\lambda_i \| \frac{-\hat{u}_i}{\lambda_i} - b_i \|^2 - \lambda_i c_i\},\ i = 1, 2, \ldots, r; \tag{3.3}$$

据[3]中定理2.2 的(2.8)式的 $A_j\hat{x} \in \partial g_j^*(\hat{u}_j^*)$ 即存在 $\hat{u}_j^*, \lambda_i^*$ 使得

$$A_i\hat{z} = -(\frac{\hat{u}_i}{\hat{\lambda}_i} + b_i),\ i = 1, \ldots, r; \tag{3.4}$$

$$\hat{z} = (d_{j1} \cdot \exp\{\frac{-\hat{v}_{j1}}{\hat{\mu}_j}\}, \ldots, d_{jn} \cdot \exp\{\frac{-\hat{v}_{jn}}{\hat{\mu}_j}\})^T,\ j = 1, \ldots, l; \tag{3.5}$$

将(3.2)代入到(3.1)中，得到

$$(H_0\hat{z} + h_0)_s + \sum_{j=1}^{l}\hat{\mu}_j\ln(\frac{\hat{z}_s}{d_{js}}) + \sum_{i=1}^{r}\hat{\lambda}_i(A_i^TA_i\hat{z} + A_i^Tb_i)_s = 0.$$

由 $b_i^T = h_i^T(A_i^TA_i)^+A_i^T$，则

$$(H_0\hat{z} + h_0)_s + \sum_{j=1}^{l}\hat{\mu}_j\ln(\frac{\hat{z}_s}{d_{js}}) + \sum_{i=1}^{r}\hat{\lambda}_i(H_i\hat{z} + h_i)_s = 0,\ s = 1, \ldots, n \tag{3.6}$$

令

$$\theta_i(\hat{\lambda}_i) \overset{\text{def}}{=} \frac{1}{2}\hat{\lambda}_i \| -\frac{\hat{u}_i}{\hat{\lambda}_i} - b_i \|^2 - \hat{\lambda}_i c_i = \frac{1}{2}\hat{\lambda}_i\hat{w}_i^T\hat{w}_i - \hat{\lambda}_i c_i;\ i = 1, \ldots, r; \tag{3.7}$$

$$\theta_j(\hat{\mu}_j) \overset{\text{def}}{=} \hat{\mu}_j\sum_{k=1}^{n} d_{jk}\exp\{\frac{-\hat{v}_{jk}}{\hat{\mu}_j}\};\ j = 1, \ldots, l; \tag{3.8}$$

因为(3.3)是极值形式,则对(3.7)用微分法求其极值有 $\frac{1}{2}\hat{\lambda}_i\hat{w}_i^T\hat{w}_i - \hat{\lambda}_i c_i + \hat{w}_i{}^T\hat{u}_i = 0$, 将 $\hat{u}_i = -\hat{\lambda}_i(\hat{w}_i + b_i)$, $b_i^T = h_i^T(A_i^TA_i)^+A_i^T$ 代入可得 $\hat{w}_i^T\hat{u}_i = -\hat{\lambda}_i\hat{w}_i^T\hat{w}_i - \hat{\lambda}_i h_i^T\hat{z}$. 将此式和(3.4)式代入上式微分求极值中整理,可得 $\hat{\lambda}_i(\frac{1}{2}\hat{z}^TA_i^TA_i\hat{z} + h_i^T\hat{z} + c_i) = 0$. $i = 1, \ldots, r$, 即互补松弛性条件成立,

$$\hat{\lambda}_i Q_i(\hat{z}) = 0. \quad i = 1, \ldots, r. \tag{3.9}$$

又由(3.2)式是极值形式,则对(3.8)用微分法求其极值可得:

$$\hat{\mu}_j \cdot \sum_{k=1}^{n} d_{jk} \cdot \exp\left(\frac{-\hat{v}_{jk}}{\hat{\mu}_j}\right) + \left(\sum_{k=1}^{n} \hat{v}_{jk}\right) d_{jk} \cdot \exp\left(\frac{-\hat{v}_{jk}}{\hat{\mu}_j}\right) = 0. \tag{3.10}$$

将(3.5)代入整理,可得 $\hat{\mu}_j \cdot \sum_{k=1}^{n}(\hat{z}_k \ln \frac{\hat{z}_k}{ed_{jk}}) = 0$. 则互补松弛性条件成立,

$$\hat{\mu}_j P_j(\hat{z}) = 0. \tag{3.11}$$

由(3.6), (3.9)以及(3.11)式,可以清楚地看到问题 (Q) 基于广义的 Fenchel 对偶定理的 Kuhn-Tucker 条件与[3]中常用的 Kuhn-Tucker 条件是一致的. □

定理3.4 假设问题 (P) 有一个 $z \geqslant 0$ 的可行解,那么问题 (P) 的基于广义 Fenchel 对偶定理的 Kuhn-Tucker 条件为常义下的 Kuhn-Tucker 条件.

证明 类似于定理3.3的证明,可得

$$\ln \frac{\hat{z}_s}{d_{js}} + \sum_{j=1}^{l}\hat{\mu}_j \ln\left(\frac{\hat{z}_s}{d_{js}}\right) + \sum_{i=1}^{r}\hat{\lambda}_i(H_i\hat{z} + h_i)_s = 0, \quad s = 1, \ldots, n. \tag{3.12}$$

由于问题 (P) 与问题 (Q) 仅是目标函数不同,那么通过(3.12), (3.9)和(3.11)式,可清楚地看到问题 (P) 基于广义的 Fenchel 对偶定理的 Kuhn-Tucker 条件与[3]中常用的 Kuhn-Tucker 条件是一致的.

□

参考文献:

[1] 朱德通. 广义的 Fenchel 对偶定理与 MDI 问题[J]. 上海师范大学学报, 1989, 18(4): 1-8.

[2] ROCKAFELLER R T. Convex Analysis[M]. Princeton University Press, Princeton, N Y, 1970.

[3] ZHANG J, BROCKETT P. Quadratically Constrained Information Theoretic Analysis[J]. SIAM J Applied Mathematics, 1987, 47: 871-886.

[4] 王国荣. 矩阵与算子广义逆[M]. 北京: 科学出版社, 1998.

Duality Theory of Two Classes of Entropy of Density Entropy Programmings

ZHANG Le-ying, ZHU De-tong

(College of Mathematics and Sciences, Shanghai Teachers University, Shanghai 200234, China)

Abstract: Consider the quadratic program problem and the entropy of the density problem with a set of quadratical inequality constraints and entropy inequality constraints of the density. The dual programmings of the two problems are derived by employing generalized Fenchel' s duality theorem and the knowledge of convex analysis. Furthermore, the duality theorems and related Kuhn-Tucker conditions for the two pairs of the dual programs are also established through the duality theory.

Key words: generalized Fenchel' s duality theorem; Kuhn-Tucker condition; density entropy problem