Методы оптимизации. Сдача, задача 3

Сергей Володин, 374 гр.

10 мая 2016 г.

Задача 3

Пусть $\{a_i\}_{i=1}^m \subset \mathbb{R}^n$. $\rho > 0$. Функция $f\colon \mathbb{R}^n \to \mathbb{R}$:

$$f(y) = \left(\sum |(a_i, y)|^\rho\right)^{\frac{1}{\rho}}$$

Задача (1):

$$\sup_{||y||_2 \leqslant 1} f(y)$$

Требуется построить двойственную задачу.
Перепишем (1):

$$\min_{||y|| \leqslant 1} -f(y)$$

Функция Лагранжа:

$$L(y, \lambda) = -f(y) + \lambda(||y|| - 1)$$

Двойственная задача:

$$\inf_{y \in \mathbb{R}^n} L(y, \lambda) \to \max_{\lambda \geqslant 0}$$

Перепишем и получим двойственную задачу:

$$-\inf_{y \in \mathbb{R}^n} L(y, \lambda) = \sup_{y \in \mathbb{R}^n} -L(y, \lambda) = \sup_{y \in \mathbb{R}^n} (f(y) - \lambda||y|| + \lambda) = \underbrace{\sup_{y \in \mathbb{R}^n} (f(y) - \lambda||y||_2) + \lambda}_{g(\lambda)} \to \min_{\lambda \geqslant 0}$$

Осталось найти $g(\lambda)$. Обозначим

$$v(\lambda) = \sup_{y \in \mathbb{R}^n} \underbrace{f(y) - \lambda||y||_2}_{M(y,\lambda)}.$$

Тогда $g(\lambda) = \lambda + v(\lambda)$. Найдем v.

Пусть $\alpha \in \mathbb{R}$. Заметим, что $f(\alpha y) = \left(\sum\limits_{i=1}^m |(a_i, \alpha y)|^\rho\right)^{\frac{1}{\rho}} = \left(|\alpha|^\rho \sum\limits_{i=1}^m |(a_i, y)|^\rho\right)^{\frac{1}{\rho}} = |\alpha| f(y)$. Также $||\alpha y||_2 = |\alpha| \cdot ||y||_2$, откуда $M(\alpha y, \lambda) = |\alpha|(f(y) - \lambda||y||_2)$.

Фиксируем $\lambda \geqslant 0$. Пусть $\exists y \in \mathbb{R}^n\colon\ f(y) - \lambda||y||_2 > 0$. Тогда возьмем $0 < \alpha_k \to \infty$ и получим $M(\alpha_k y, \lambda) = \alpha_k \underbrace{M(y, \lambda)}_{>0} \to +\infty$.

То есть, $v(\lambda) = +\infty$

Пусть верно обратное, то есть, $\forall y \in \mathbb{R}^n \hookrightarrow f(y) - \lambda||y||_2 \leqslant 0$. Но это значит, что $v(\lambda) \leqslant 0$. Но $v(\lambda) \geqslant M(0, \lambda) = 0$, значит, $v(\lambda) = 0$.

Перепишем условие $\forall y \in \mathbb{R}^n f(y) - \lambda||y||_2 \leqslant 0 \Leftrightarrow \lambda \geqslant \sup\limits_{y \in \mathbb{R}^n} \frac{f(y)}{||y||_2} = \lambda^*$.

Получаем, что $g(\lambda) = \begin{cases} \lambda & \lambda \geqslant \lambda^* \\ +\infty & \lambda \in [0, \lambda^*] \end{cases}$

Заметим, что решением двойственной задачи $\inf\limits_{\lambda \geqslant 0} g(\lambda)$ является число λ^*, так как $\inf g(\mathbb{R}_+) = \inf[\lambda^*, +\infty] = \lambda^*$.

Также

$$\lambda^* = \sup_{y \in \mathbb{R}^n} \frac{f(y)}{||y||} = \sup_{||y||=1, t \in \mathbb{R}_+} \frac{f(ty)}{||ty||} = \sup_{||y||=1} f(y) = \sup_{||y||=1, t \in [0,1]} tf(y) = \sup_{||y|| \leqslant 1} f(y)$$

То есть, число λ^* является решением исходной задачи (1).
Вопрос: этого достаточно, или нужно найти λ^* (т.е. решить исходную задачу)?
Поиск λ^*
Рассмотрим $\lambda^* = \sup\limits_{y \in \mathbb{R}^n} \frac{f(y)}{||y||}$. Найдем

$$\frac{\partial}{\partial y_j} \frac{f(y)}{||y||_2} = \frac{1}{||y||^2}(f_j||y|| - f\frac{y_j}{||y||}), \text{ где } f_j = \frac{\partial f}{\partial y_j}$$

. Приравняем нулю, получим

$$f_j||y||^2 = y_j f$$

Найдем

$$f_j = \frac{1}{\rho}\left(\sum_{i=1}^{m}|(a_i,y)|^{\rho}\right)^{\frac{1}{\rho}-1}\sum_{i=1}^{m}\rho|(a_i,y)|^{\rho-1}\frac{\partial|(a_i,y)|}{\partial y_j} = f^{1-\rho}\sum_{i=1}^{m}|(a_i,y)|^{\rho-1}\frac{\partial|(a_i,y)|}{\partial y_j}$$

Подставим, получим

$$f^{1-\rho}\sum_{i=1}^{m}|(a_i,y)|^{\rho-1}\frac{\partial|(a_i,y)|}{\partial y_j}||y||^2 = y_j f$$

$$\sum_{i=1}^{m}|(a_i,y)|^{\rho-1}\frac{\partial|(a_i,y)|}{\partial y_j}||y||^2 = y_j f^{\rho}$$

То есть, $\lambda^* = f(y)$, где $||y|| = 1$ и

$$y_j\sum_{i=1}^{m}|(a_i,y)|^{\rho} = \sum_{i=1}^{m}|(a_i,y)|^{\rho-1}\frac{\partial|(a_i,y)|}{\partial y_j}$$

Заметим, что в этой точке также выполнено $f^2 = f_1^2 + ... + f_n^2$

Случай $m = 2$

Пусть $m = 2$. Найдем λ^*.

1. Пусть $X = Lin\{a_i\}_{i=1}^m$. Пусть $y \in \mathbb{R}^n$, причем $y = y_{\parallel} + y_{\perp}$, $y_{\parallel} \in X$, $y_{\perp} \in X^{\perp}$. Тогда $f(y) = f(y_{\parallel})$. Действительно, $f(y)$ зависит только от скалярных произведений с a_i, а $(y_{\perp}, a_i) = 0$ ■

   Значит, искать точку максимума можно только среди векторов $y \in X$

2. Заметим, что решение задачи не меняется при замене $a_i \to -a_i$. Действительно, f зависит только от модулей скалярных произведений с a_i. Поэтому, без ограничения общности, угол между a_1 и a_2 — острый: $(a_1, a_2) > 0$.

3. Заметим, что как скалярное произведение, так и норма $||\cdot||_2$ не зависят от выбора ортогонального базиса. Выберем базис таким образом: $e_1 \upuparrows a_1$, e_2: a_2 лежит в плоскости e_1, e_2 в первом ортанте. Остальные $n-2$ векторов базиса выбираем произвольно.

   Параметризуем
   $$y = ||\cos\varphi \ \sin\varphi||^T$$
   (ранее доказано, что максимум можно искать по сфере $||y||_2 = 1$)

   Представим
   $$\vec{a_1} = a_1 \cdot ||1\,0||^T$$
   a_1 — вторая норма $\vec{a_1}$ Представим
   $$\vec{a_2} = a_2 \cdot ||\cos\theta \ \sin\theta||^T$$
   Тогда получим, что $(a_1, y) = a_1\cos\varphi$, $(a_2, y) = a_2(\cos\varphi\cos\theta + \sin\varphi\sin\theta) = a_2(\cos(\varphi - \theta))$.

4. Задача свелась к задаче БМ
   $$f^{\rho}(\varphi) = a_1^{\rho}\cos^{\rho}\varphi + a_2^{\rho}\cos^{\rho}(\varphi-\theta) \to \max_{\varphi\in[0,\frac{\pi}{2}]}$$

5. Обозначим
   $$\alpha = \frac{a_1^{\rho}}{a_2^{\rho}}$$

6. Задача эквивалентна
   $$z(\varphi) = \alpha\cos^{\rho}\varphi + \cos^{\rho}(\varphi-\theta) \to \max_{\varphi\in[0,\frac{\pi}{2}]}$$

7. Сделаем замену $t = \cos\theta$, получим выражение
   $$z(t) = \alpha t^{\rho} + (t\cos\theta + \sqrt{1-t^2}\sin\theta)^{\rho} \to \max_{t\in[0,1]}$$
   $$z(t)'/\rho = \alpha t^{\rho-1} + (t\cos\theta + \sqrt{1-t^2}\sin\theta)^{\rho-1}(\cos\theta - \sin\theta\frac{t}{\sqrt{1-t^2}}) = 0$$
   Задача свелась к поиску нуля выражения выше, получаем

8. Рассмотрим случай $\rho = \frac{1}{3}$.

$$z'(\varphi)/\rho = \alpha t^{-2/3} + (t\cos\theta + \sqrt{1-t^2}\sin\theta)^{-2/3}(\cos\theta - \frac{t}{\sqrt{1-t^2}}\sin\theta) = 0$$

$$\alpha t^{-2/3} = (t\cos\theta + \sqrt{1-t^2}\sin\theta)^{-2/3}(\frac{t}{\sqrt{1-t^2}}\sin\theta - \cos\theta)$$

$$\alpha^3 t^{-2} = (\frac{t\sin\theta}{\sqrt{1-t^2}} - \cos\theta)^3(t\cos\theta + \sqrt{1-t^2}\sin\theta)^{-2}$$

$$\alpha^3(t\cos\theta + \sqrt{1-t^2}\sin\theta)^2 = (\frac{t\sin\theta}{\sqrt{1-t^2}} - \cos\theta)^3 t^2$$

$$t^2\alpha^3(\cos\theta + \frac{\sqrt{1-t^2}}{t}\sin\theta)^2 = (\frac{t\sin\theta}{\sqrt{1-t^2}} - \cos\theta)^3 t^2$$

Обозначим

$$\gamma = \frac{t}{\sqrt{1-t^2}}$$

$$\alpha^3(\cos\theta + \frac{\sin\theta}{\gamma})^2 = (\gamma\sin\theta - \cos\theta)^3$$

$$\alpha^3(\gamma\cos\theta + \sin\theta)^2 = \gamma^2(\gamma\sin\theta - \cos\theta)^3$$

Это уравнение пятой степени относительно γ.