ВЫПУКЛЫЙ АНАЛИЗ

ЛЕКЦИЯ 27

10. ТЕОРИЯ ДВОЙСТВЕННОСТИ

10. ТЕОРИЯ ДВОЙСТВЕННОСТИ

10.1. Постановка двойственной задачи. Рассмотрим задачу выпуклого программирования

$$I(u) \to \inf, \ u \in U,$$

$$U = \left\{ u \in U_0 \,\middle|\, g_i(u) \le 0, i = 1, \ldots, m; \ g_i(u) = 0, i = m+1, \ldots, s \right\},$$

$$\Lambda^0 = \left\{ \lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_s \end{pmatrix} \in R^s \,\middle|\, \begin{array}{l} \lambda_i \ge 0, \\ i = 1, \ldots, m \end{array} \right\}$$

Пусть $L : U_0 \times \Lambda^0 \to R^1$ – ее регулярная функция Лагранжа. Рассмотрим функцию $\chi : U_0 \to R^1$, определенную формулой

$$\chi(u) = \sup_{\lambda \in \Lambda^0} L(u, \lambda) = \sup_{\lambda \in \Lambda^0} \left[I(u) + \sum_{i=1}^{s} \lambda_i g_i(u) \right], u \in U_0.$$

Нетрудно видеть, что

$$u \in U \Rightarrow \sum_{i=1}^{s} \overset{\begin{array}{ll} \ge 0, & i=1,\ldots,m \\ \forall \text{знак} & i=m+1,\ldots,s \end{array}}{\lambda_i} \cdot \overset{\begin{array}{ll} \le 0, & i=1,\ldots,m, \\ 0, & i=m+1,\ldots,s \end{array}}{g_i(u)} \le 0, \forall \lambda \in \Lambda^0$$

и при $\lambda = 0 \in \Lambda^0$ неравенство переходит в равенство.

Пусть $$u \in U_0 \setminus U \Rightarrow \begin{array}{l} \text{либо } \exists\, i \in \{1, \ldots, m\} : g_i(u) > 0, \\ \text{либо } \exists\, i \in \{m+1, \ldots, s\} : g_i(u) \ne 0. \end{array}$$

Тогда величина $\sum_{i=1}^{s} \lambda_i g_i(u)$ выбором вектора $\lambda \in \Lambda^0$ может быть сделана сколь угодно большой. Отсюда выводим равенство

$$\chi(u) = \sup_{\lambda\in\Lambda^0} L(u,\lambda) = \begin{cases} \sup\limits_{\lambda\in\Lambda^0}\left(I(u) + \overbrace{\sum_{i=1}^{s} \lambda_i g_i(u)}^{\leq 0} \right), & u \in U, \\ +\infty, & u \in U_0 \setminus U. \end{cases} =$$

$$= \begin{cases} I(u), & u\in U, \\ +\infty, & u\in U_0\setminus U \end{cases} \Rightarrow \quad \inf_{u\in U_0} \chi(u) = \inf_{u\in U} I(u) = I_*.$$

Тогда исходную задачу можно переписать в виде.

Задача 1. $$\chi(u) \to \inf, \quad u \in U_0.$$

Наряду с функцией χ рассмотрим функцию $\psi : \Lambda^0 \to R^1$, определенную формулой

$$\psi(\lambda) = \inf_{u\in U_0} L(u,\lambda), \;\; \lambda\in\Lambda^0$$

и сконструируем задачу.

Задача 2. $$\psi(\lambda) \to \sup, \quad \lambda\in\Lambda^0.$$

Определение 1. **Задача 2** ***называется двойственной к*** **задаче 1** ***(основной).*** ***Переменные*** $u^1,\ldots,u^n$ ***называются основными, а переменные*** $\lambda_1,\ldots,\lambda_s$ – ***двойственными.***

Обозначим через

$$\Lambda^* = \left\{ \lambda^* \in \Lambda^0 \middle| \psi\left(\lambda^*\right) = \psi^* \right\}, \quad \psi^* = \sup_{\lambda \in \Lambda^0} \psi\left(\lambda\right)$$

множество всех решений двойственной задачи и оптимальное значение ее целевой функции, соответственно.

10.2. Теорема двойственности. Установим связь между оптимальными значениями целевых функций основной и двойственной задач.

Теорема 1. ***Имеет место неравенство***

$$\psi^* = \sup_{\lambda \in \Lambda^0} \psi\left(\lambda\right) \le \inf_{u \in U} I(u) = I_*.$$

Доказательство. Для всех $u \in U_0, \lambda \in \Lambda^0$ справедливо

$$\psi\left(\lambda\right) = \inf_{u \in U_0} L\left(u, \lambda\right) \le L\left(u, \lambda\right). \qquad (1)$$

Возьмем точную верхнюю грань по переменной $\lambda \in \Lambda^0$ от обеих частей (1)

$$\psi^* = \sup_{\lambda \in \Lambda^0} \psi\left(\lambda\right) \le \sup_{\lambda \in \Lambda^0} L\left(u, \lambda\right) = \chi(u), \quad \forall u \in U_0. \qquad (2)$$

Переходя в неравенстве (2) $\psi^* \le \chi(u) \quad (2)$ к точной нижней грани по переменной $\forall u \in U_0,$ получим

$$\psi^* \le \inf_{u \in U_0} \chi(u) = I_* \Rightarrow \psi^* \le I_*.$$

Теорема доказана.

Теорема 2 (двойственности). ***Для того чтобы было выполнено***

$$U_* \ne \varnothing, \Lambda^* \ne \varnothing, I_* = \psi^* \qquad (3)$$

необходимо и достаточно, чтобы функция Лагранжа имела седловую точку на множестве $U_0 \times \Lambda^0$. ***Множество седловых точек функции Лагранжа совпадает с множеством*** $U_* \times \Lambda^* = \left\{ \left(u_*, \lambda^* \right) \middle| u_* \in U_*, \lambda^* \in \Lambda^* \right\}.$

Доказательство. Необходимость. Пусть выполнены соотношения (3). Покажем, что для любых $u_* \in U_*, \lambda^* \in \Lambda^*$ пара $\left(u_*, \lambda^* \right)$ является седловой точкой для функции Лагранжа. Имеем

$$\psi^* = \psi\left(\lambda^*\right) \overset{\psi(\lambda) = \inf\limits_{u \in U_0} L(u,\lambda)}{=} \inf_{u \in U_0} L(u, \lambda^*) \le L(u_*, \lambda^*) \le$$

$$\le \sup_{\lambda \in \Lambda^0} L(u_*, \lambda) \overset{\sup\limits_{\lambda \in \Lambda^0} L(u,\lambda) = \chi(u)}{=} \chi\left(u_*\right) = I_*. \quad (4)$$

По условию необходимости $I_* = \psi^*$. Тогда из неравенства (4)

$$\psi^* = \inf_{u \in U_0} L(u, \lambda^*) \le L(u_*, \lambda^*) \le \sup_{\lambda \in \Lambda^0} L(u_*, \lambda) = I_* \quad (4)$$

следует, что

$$\sup_{\lambda \in \Lambda^0} L(u_*, \lambda) = L(u_*, \lambda^*) = \inf_{u \in U_0} L(u, \lambda^*). \qquad (5)$$

Последнее равенство означает, что

$$L(u_*, \lambda) \le L(u_*, \lambda^*) \le L(u, \lambda^*), \ \forall u \in U_0, \forall \lambda \in \Lambda^0.$$

Таким образом, пара $\left(u_*, \lambda^*\right)$ – седловая точка. Отсюда также следует, что множество $U_* \times \Lambda^*$ вложено в множество седловых точек функции Лагранжа. Необходимость доказана.

Достаточность. Пусть $\left(u_*, \lambda^*\right) \in U_0 \times \Lambda^0$ – седловая точка функции Лагранжа. Тогда

$$L(u_*, \lambda) \le L(u_*, \lambda^*), \quad \forall \lambda \in \Lambda^0.$$

Отсюда следует

$$\sup_{\lambda \in \Lambda^0} L(u_*, \lambda) = \chi\left(u_*\right) \le L(u_*, \lambda^*). \qquad (6)$$

Аналогично из неравенства

$$L(u_*,\lambda^*)\le L(u,\lambda^*),\ \forall\, u\in U_0$$

выводим

$$\inf_{u\in U_0} L(u,\lambda^*)=\psi\left(\lambda^*\right)\ge L(u_*,\lambda^*). \qquad (7)$$

Из (7),(6) $\sup\limits_{\lambda\in\Lambda^0} L(u_*,\lambda)=\chi\left(u_*\right)\le L(u_*,\lambda^*) \quad (6)$ в силу **теоремы 1** получим

$$L(u_*,\lambda^*)\overset{(7)}{\le}\psi\left(\lambda^*\right)\overset{\psi^*=\sup\limits_{\lambda\in\Lambda^0}\psi(\lambda)}{\le}\psi^*\overset{\substack{\psi^*\le I_*\\ \textit{теорема } 1}}{\le} I_*\overset{I_*=\inf\limits_{u\in U_0}\chi(u)}{\le}\chi(u_*)\overset{(6)}{\le} L(u_*,\lambda^*)$$

Тогда

$$\psi\left(\lambda^*\right)=\psi^*=I_*=\chi\left(u_*\right)\Rightarrow$$

$$\psi^*=I_*,\quad u_*\in U_*,\lambda^*\in\Lambda^*$$

Отсюда выводим, что множество седловых точек функции Лагранжа вложено в множество $U_*\times\Lambda^*$. Теорема доказана.

Отсюда, в частности следует, что если пары

$$(u_*^{(1)},\lambda^{*(1)}),(u_*^{(2)},\lambda^{*(2)})\in U_0\times\Lambda^0$$

образуют седловые точки функции Лагранжа, то и пары

$$(u_*^{(2)}, \lambda^{*(1)}), (u_*^{(1)}, \lambda^{*(2)}) \in U_0 \times \Lambda^0$$

также являются седловыми точками. Тогда множители Лагранжа в теореме Куна – Таккера можно выбирать одними и теми же для всех $u_* \in U_*$.

Замечание. В доказательстве теоремы двойственности нигде не использовался тот факт, что исходная задача является задачей выпуклого программирования. Таким образом, теорема двойственности верна и для произвольной задачи математического программирования.

Очевидно, что если пара $(u_*, \lambda^*) \in U_0 \times \Lambda^0$ образует седловую точку функции Лагранжа, то пара $(\hat{u}, \hat{\lambda}) \in U_0 \times \Lambda^0$ определенная из условия

$$L(\hat{u}, \hat{\lambda}) = L(u_*, \lambda^*)$$

не обязана быть седловой точкой. Более того пусть $(u_*, \lambda^*) \in U_0 \times \Lambda^0 -$ седловая точка функции Лагранжа. Рассмотрим множества

$$U^*\left(\lambda^*\right) = \left\{u \in U_0 \,\middle|\, L(u, \lambda^*) = L(u_*, \lambda^*)\right\},$$

$$\Lambda\left(u_*\right)=\left\{\lambda\in\Lambda^0 \,\middle|\, L(u_*,\lambda)=L(u_*,\lambda^*)\right\}.$$

В общем случае выполняется лишь

$$U_*\subset U^*\left(\lambda^*\right),\quad \Lambda^*\subset\Lambda^*\left(u_*\right).$$

Пример 1. Пусть

$$L\left(u,\lambda\right)=\lambda u,\quad U_0=\Lambda^0=R^1.$$

Данная функция имеет единственную седловую точку $(u_*,\lambda^*)=\left(0,0\right),$ а

$$U^*\left(\lambda^*\right)=\Lambda^*\left(u_*\right)=R^1.$$

Существуют задачи выпуклого программирования, для которых $\psi^*<I_*,$ даже если

$U_*\neq\varnothing, \Lambda^*\neq\varnothing.$

Пример 2. Пусть

$$I\left(u\right)=e^{-u}, U_0=R^1,\ m=0, s=1, g_1\left(u\right)=g\left(u\right)=ue^{-u}$$

Тогда

$$U=\left\{u\in U_0 \,\middle|\, g\left(u\right)=0\right\}=\left\{0\right\}\Rightarrow\ U_*=\left\{0\right\}, I_*=I\left(0\right)=1,$$

и

$$\Lambda^0=\left(-\infty,+\infty\right).$$

Выпишем функцию Лагранжа для данной задачи $I(u)=e^{-u}, g(u)=ue^{-u}$

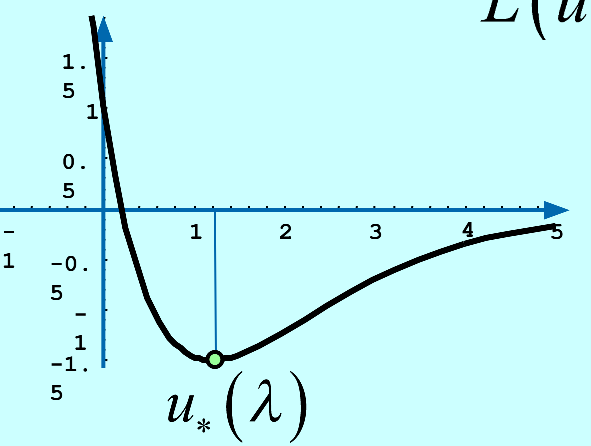
$$L(u,\lambda)=e^{-u}+\lambda ue^{-u}=e^{-u}(1+\lambda u).$$

Вычисляем

$$\psi(\lambda)=\inf_{u\in R^1}L(u,\lambda)=\begin{cases}0, & \lambda=0;\\ -\infty, & \lambda>0;\\ \lambda e^{-1+\frac{1}{\lambda}}, & \lambda<0\end{cases}\qquad(8)$$

Пояснения требует третья строчка в равенстве (8). Приведем график функции

$$L(u,\lambda)\Big|_{\lambda=const<0}=e^{-u}(1-\lambda u)\Big|_{\lambda=const<0}$$



при $\lambda=-5.$ Найдем минимизирующую точку

$u_*(\lambda)\in R^1$ для $\lambda<0.$ Имеем

$$\frac{\partial L(u,\lambda)}{\partial u}=-e^{-u}-\lambda ue^{-u}+\lambda e^{-u}=0\Rightarrow$$

$$-1-\lambda u+\lambda=0\Rightarrow\qquad u_*(\lambda)=\frac{\lambda-1}{\lambda}.$$

Таким образом, для $\lambda < 0$ имеем

$$\psi(\lambda) = L(u_*(\lambda), \lambda)\Big|_{u_*(\lambda)=\frac{\lambda-1}{\lambda}.} = e^{-u_*(\lambda)}\left(1 - u_*(\lambda)\cdot\lambda\right)\Big|_{u_*(\lambda)=\frac{\lambda-1}{\lambda}} =$$

$$= e^{-\frac{\lambda-1}{\lambda}}\left(1 + \lambda\cdot\frac{\lambda-1}{\lambda}\right) = \lambda e^{-1+\frac{1}{\lambda}}$$

и соотношение (8) доказано. Отсюда

$$\psi^* = \sup_{\lambda\in R^1}\psi(\lambda) = \sup_{\lambda\in R^1}\left(\begin{cases} 0, & \lambda = 0; \\ -\infty, & \lambda > 0; \\ \lambda e^{-1+\frac{1}{\lambda}} < 0, & \lambda < 0 \end{cases}\right) = \psi(0) = 0 \Rightarrow$$

$$\psi^* = 0, \Lambda^* = \{0\} \neq \varnothing,$$

Таким образом,

$$\Lambda^* = \{0\} \neq \varnothing, U_* = \{0\} \neq \varnothing, \quad \text{но} \quad \psi^* = 0 < 1 = I_*.$$

Двойственная задача всегда является задачей выпуклого программирования независимо от того, какой была основная (исходная) задача. Действительно, двойственная задача

$$\psi(\lambda) \to \sup, \quad \lambda \in \Lambda^0$$

эквивалентна задаче

$$\hat{\psi}(\lambda) \to \inf, \quad \lambda \in \Lambda^0, \hat{\psi}(\lambda) = -\psi(\lambda), \quad \lambda \in \Lambda^0.$$

Достаточно установить выпуклость функции $\hat{\psi}$ на множестве Λ^0. Имеем

$$\hat{\psi}(\lambda) = -\psi(\lambda) = -\inf_{u \in U_0} L(u,\lambda) = \sup_{u \in U_0}\left[-L(u,\lambda)\right], \ \lambda \in \Lambda^0. \quad (9)$$

Для всех $\lambda^{(1)}, \lambda^{(2)} \in \Lambda^0, \quad \alpha \in [0,1]$ справедливо

$$\hat{\psi}\left(\alpha\lambda^{(1)} + (1-\alpha)\lambda^{(2)}\right) = \sup_{u \in U_0}\left[-L\left(u, \alpha\lambda^{(1)} + (1-\alpha)\lambda^{(2)}\right)\right] =$$

$$= \sup_{u \in U_0}\left[-\left(\alpha + (1-\alpha)\right)I(u) - \sum_{i=1}^{s}\left(\alpha\lambda_i^{(1)} + (1-\alpha)\lambda_i^{(2)}\right)g_i(u)\right] =$$

$$= \sup_{u \in U_0}\left[-\alpha I(u) - (1-\alpha)I(u) - \alpha\sum_{i=1}^{s}\lambda_i^{(1)}g_i(u) - (1-\alpha)\sum_{i=1}^{s}\lambda_i^{(2)}g_i(u)\right] =$$

$$= \sup_{u\in U_0}\left[-\alpha I(u)-(1-\alpha)I(u)-\alpha\sum_{i=1}^{s}\lambda_i^{(1)}g_i(u)-(1-\alpha)\sum_{i=1}^{s}\lambda_i^{(2)}g_i(u)\right]=$$

$$=\sup_{u\in U_0}\left\{-\alpha\left[I(u)+\sum_{i=1}^{s}\lambda_i^{(1)}g_i(u)\right]-(1-\alpha)\left[I(u)+\sum_{i=1}^{s}\lambda_i^{(2)}g_i(u)\right]\right\}=$$

$$=\sup_{u\in U_0}\left[-\alpha L\left(u,\lambda^{(1)}\right)-(1-\alpha)L\left(u,\lambda^{(2)}\right)\right]\le$$

$$\le\sup_{u\in U_0}\left[-\alpha L\left(u,\lambda^{(1)}\right)\right]+\sup_{u\in U_0}\left[-(1-\alpha)L\left(u,\lambda^{(2)}\right)\right]=$$

$$\overset{(9)}{=}\alpha\overbrace{\sup_{u\in U_0}\left[-L\left(u,\lambda^{(1)}\right)\right]}^{\hat\psi\left(\lambda^{(1)}\right)}+(1-\alpha)\overbrace{\sup_{u\in U_0}\left[-L\left(u,\lambda^{(2)}\right)\right]}^{\hat\psi\left(\lambda^{(2)}\right)}=$$

$$=\alpha\hat\psi\left(\lambda^{(1)}\right)+(1-\alpha)\hat\psi\left(\lambda^{(2)}\right).$$

Выпуклость функции $\hat\psi$ на множестве Λ^0 доказана, т.е. двойственная задача является задачей выпуклого программирования. Благодаря этому обстоятельству в задачах математического программирования, регулярная функция Лагранжа которых имеет седловую точку, удобнее сначала исследовать двойственную к ней задачу.

Двойственная задача к двойственной не обязана совпадать с исходной задачей.

Например, если исходная задача не является задачей выпуклого программирования.