Conjugate functions

Ковалев Алексей

1. $f(x) = |2x|$.
$$f^*(y) = \sup_{x\in\mathbb{R}} \big(\langle x,\, y\rangle - f(x)\big) = \sup_{x\in\mathbb{R}} \big(xy - 2|x|\big)$$
При $y > 2$ имеем $\sup\limits_{x\in\mathbb{R}} \big(xy - 2|x|\big) = \sup\limits_{x\in\mathbb{R}_+} x(y-2) = +\infty$.
При $y < 2$ имеем $\sup\limits_{x\in R} \big(xy - 2|x|\big) = \sup\limits_{x\in\mathbb{R}_-} x(y+2)) = +\infty$.
Пусть при $|y| \leqslant 2$ выполняется $xy - 2|x| > 0$. Тогда $xy > 2|x|$, то есть $y \operatorname{sign} x > 2$, что противоречит условию $|y| \leqslant 2$. Значит при $|y| \leqslant 2$ имеем $\sup\limits_{x\in\mathbb{R}} \big(xy - 2|x|\big) \leqslant 0$, причем 0 достигается при любом y на $x = 0$. Значит $f^*(y) = 0$, $\operatorname{dom} f^* = [-2; 2]$.
Ответ: $f^*(y) = 0$, $\operatorname{dom} f^* = [-2; 2]$.

2. $f(x) = \inf\limits_{u+v=x} \big(g(u) + h(v)\big)$.
$$-\inf_{u+v=x} \big(g(u) + h(v)\big) = \sup_{u+v=x} \big(-g(u) - h(v)\big)$$
$$\begin{aligned}
f^*(y) &= \sup_{x\in\operatorname{dom} f} \big(\langle x,\, y\rangle - f(x)\big) = \sup_{x\in\operatorname{dom} f} \left(\langle x,\, y\rangle - \inf_{u+v=x} \big(g(u) + h(v)\big)\right) \\
&= \sup_{x\in\operatorname{dom} f} \left(\langle x,\, y\rangle + \sup_{u+v=x} \big(-g(u) - h(v)\big)\right) = \sup_{x\in\operatorname{dom} f}\ \sup_{u+v=x} \big(\langle x,\, y\rangle - g(u) - h(v)\big) \\
&\overset{(0)}{=} \sup_{\substack{u\in\operatorname{dom} g\\ v\in\operatorname{dom} h}} \big(\langle u,\, y\rangle - g(u) + \langle v,\, y\rangle - h(v)\big) \overset{(1)}{=} \sup_{u\in\operatorname{dom} g} \big(\langle u,\, y\rangle - g(u)\big) + \sup_{v\in\operatorname{dom} h} \big(\langle v,\, y\rangle - h(v)\big) \\
&= g^*(y) + h^*(y)
\end{aligned}$$
Переход (0) объясняется тем, что в левой части равенства супремум берется по всем $x \in \operatorname{dom} f$, а в правой части – по всем $(u, v) \in \operatorname{dom} g \times \operatorname{dom} h$, причем при всех таких (u, v) определены $g(u)$, $h(v)$, а значит и $f(u + v) = f(x)$, и наоборот, если определена $f(x)$, то найдется некоторая $(u, v) \in \operatorname{dom} g \times \operatorname{dom} h$.
Переход (1) объясняется тем, что $\sup\limits_{a\in A,\, b\in B} \big(\varphi(a) + \psi(b)\big) = \sup\limits_{a\in A} \varphi(a) + \sup\limits_{b\in B} \psi(b)$, так как φ не зависит от b, ψ не зависит от a. □

3.
$$f(x) = \log \sum_{k=1}^{n} e^{x_k}$$
$$f^*(y) = \sup_{x\in\mathbb{R}^n} \big(\langle x,\, y\rangle - f(x)\big) = \sup_{x\in\mathbb{R}^n} g(x, y)$$
Покажем, что если $\langle \mathbf{1},\, y\rangle \neq 1$, то функция $g(x, y)$ неограниченна как функция x. Пусть $x_1 = x_2 = \ldots = x_n = a$.
$$g(x, y) = \sum_{k=1}^{n} x_k y_k - \log \sum_{k=1}^{n} e^{x_k} = a \sum_{k=1}^{n} y_k - \log(ne^a) = a \left(\sum_{k=1}^{n} y_k - 1\right) - \log n = a\big(\langle \mathbf{1},\, y\rangle - 1\big) - \log n$$
Если $\langle \mathbf{1},\, y\rangle - 1 > 0$, то можно взять $a > 0$ сколь угодно большим и получить $g(x, y)$ сколь угодно большим. Если $\langle \mathbf{1},\, y\rangle - 1 < 0$, то можно брать $a < 0$ сколь угодно маленьким и также получить $g(x, y)$ сколь угодно

большим. Значит $g(x, y)$ неограниченна как функция x и f^* не определена при $\langle \mathbf{1},\, y\rangle \neq 1$.
Пусть теперь $\langle \mathbf{1},\, y\rangle = 1$. Покажем, что $f(x)$ – выпуклая функция.

$$\nabla f_i(x) = \frac{e^{x_i}}{\sum\limits_{k=1}^{n} e^{x_k}}$$

При $i \neq j$ имеем

$$\nabla^2 f_{ij}(x) = -\frac{e^{x_i + x_j}}{\left(\sum\limits_{k=1}^{n} e^{x_k}\right)^2}$$

При $i = j$ имеем

$$\nabla^2 f_{ij}(x) = \nabla^2 f_{ii}(x) = \frac{\sum\limits_{k=1}^{n} e^{x_i + x_k} - e^{2x_i}}{\left(\sum\limits_{k=1}^{n} e^{x_k}\right)^2}$$

Тогда для любого $0 \neq z \in \mathbb{R}^n$ получаем

$$z^\top \nabla^2 f(x) z = \sum_{i \neq j} (z_i^2 + z_j^2 - 2z_i z_j) e^{x_i + x_j} = \sum_{i \neq j} (z_i - z_j)^2 e^{x_i + x_j} \geqslant 0$$

Значит $f(x)$ действительно является выпуклой. Тогда $\nabla_x g(x, y) = y - \nabla f(x) = 0$, то есть $y_i \sum\limits_{k=1}^{n} e^{x_k} = e^{x_i}$ или же $\log y_i + \log \sum\limits_{k=1}^{n} e^{x_k} = x_i$. Получаем

$$f^*(y) = \langle x,\, y\rangle - \sum_{k=1}^{n} e^{x_k} = \sum_{k=1}^{n} y_k \left(\log y_k + \log \sum_{k=1}^{n} e^{x_k}\right) - \log \sum_{k=1}^{n} e^{x_k} = \sum_{k=1}^{n} y_k \log y_k$$

Ответ: $f^*(y) = \sum\limits_{k=1}^{n} y_k \log y_k$, $\operatorname{dom} f^* = \{y :\ y \in \mathbb{R}^n,\, \langle \mathbf{1},\, y\rangle = 1\}$.

4. $f(x) = g(Ax)$. Пусть $t = Ax$. Тогда

$$f^*(y) = \sup_{x \in \operatorname{dom} f} \big(\langle x,\, y\rangle - g(Ax)\big) = \sup_{t \in \operatorname{dom} g} \big(\langle A^{-1}t,\, y\rangle - g(t)\big) = \sup_{t \in \operatorname{dom} g} \big(\langle t,\, A^{-\top}y\rangle - g(t)\big) = g^*\big(A^{-\top}y\big)$$

$\square$

5. $f(X) = -\ln\det X$, $X \in \mathbb{S}^n_{++}$. $f(X)$ – выпуклая функция (конспект к занятию 4), причем $\nabla f(X) = -X^{-\top}$ (задание 1, Automatic differentiation).

$$f^*(Y) = \sup_{X \in \mathbb{S}^n_{++}} \big(\langle X,\, Y\rangle - f(X)\big) = \sup_{X \in \mathbb{S}^n_{++}} g(X, Y)$$

$$\nabla_X g(X, Y) = \nabla\langle X,\, Y\rangle - \nabla f(X) = Y + X^{-\top} = 0$$
$$X = -Y^{-\top}$$

Отсюда получаем $\operatorname{dom} f^* = \{Y :\ -Y^{-\top} \in \mathbb{S}^n_{++}\} = \mathbb{S}^n_{--}$.

$$f^*(Y) = \langle X,\, Y\rangle - f(X) = \langle -Y^{-\top},\, Y\rangle - f(-Y^{-\top}) = -\langle I,\, I\rangle + \ln\frac{(-1)^n}{\det Y} = -n + \ln\frac{(-1)^n}{\det Y}$$

Ответ: $f^*(Y) = -n + \ln\frac{(-1)^n}{\det Y}$, $\operatorname{dom} f^* = \mathbb{S}^n_{--}$.

6.

$$f_{\mathrm{cshub}}(x) = f_{\mathrm{hub}}(\|x\|_2) = \begin{cases} \frac{1}{2}\|x\|_2^2 & \|x\|_2 \leqslant 1 \\ \|x\|_2 - \frac{1}{2} & \|x\|_2 > 1 \end{cases}$$

$$\nabla f_{\mathrm{cshub}}(x) = \begin{cases} x & \|x\|_2 \leqslant 1 \\ \frac{x}{\|x\|_2} & \|x\|_2 > 1 \end{cases}$$

Покажем, что $f_{\mathrm{cshub}}(x)$ – выпуклая. Для этого воспользуемся дифференциальным критерием выпуклости: $\forall x,\, y \in \mathbb{R}^n$

$$f(x+y) \geqslant f(x) + \langle \nabla f(x),\, y\rangle$$

Рассмотрим 4 случая:

- Пусть $\|x+y\|_2 \leqslant 1$, $\|x\|_2 \leqslant 1$.

$$f_{\mathrm{cshub}}(x+y) = \frac{1}{2}\langle x+y,\, x+y\rangle = \frac{1}{2}\|x\|_2^2 + \frac{1}{2}\|y\|_2^2 + \langle x,\, y\rangle \geqslant \frac{1}{2}\|x\|_2^2 + \langle x,\, y\rangle = f_{\mathrm{cshub}}(x) + \langle \nabla f_{\mathrm{cshub}}(x),\, y\rangle$$

- Пусть $\|x+y\|_2 \leqslant 1$, $\|x\|_2 > 1$. Тогда $\langle x,\, y\rangle < 0$

$$\forall t \in \mathbb{R}\ \frac{1}{2}t^2 \geqslant t - \frac{1}{2}$$

$$f_{\mathrm{cshub}}(x+y) = \frac{1}{2}\langle x+y,\, x+y\rangle = \frac{1}{2}\|x\|_2^2 + \frac{1}{2}\|y\|_2^2 + \langle x,\, y\rangle \geqslant \|x\|_2 - \frac{1}{2} + \left\langle \frac{x}{\|x\|_2},\, y\right\rangle = f_{\mathrm{cshub}}(x) + \langle \nabla f_{\mathrm{cshub}}(x),\, y\rangle$$

- $\|x+y\|_2 > 1$, $\|x\|_2 \leqslant 1$

$$f_{\mathrm{cshub}}(x+y) = \|x+y\|_2 - \frac{1}{2} \geqslant \frac{1}{2}\|x\|_2^2 + \langle x,\, y\rangle = f_{\mathrm{cshub}}(x) + \langle \nabla f_{\mathrm{cshub}}(x),\, y\rangle$$

- $\|x+y\|_2 > 1$, $\|x\|_2 > 1$

$$f_{\mathrm{cshub}}(x+y) = \|x+y\|_2 - \frac{1}{2} \geqslant \|x\|_2 - \frac{1}{2} + \left\langle \frac{x}{\|x\|_2},\, y\right\rangle = f_{\mathrm{cshub}}(x) + \langle \nabla f_{\mathrm{cshub}}(x),\, y\rangle$$

Значит f_{cshub} – выпуклая функция. $\square$
Найдем теперь сопряженную функцию

$$f^*_{\mathrm{cshub}}(y) = \sup_{x \in \mathbb{R}^n} \big(\langle x,\, y\rangle - f_{\mathrm{cshub}}(x)\big) = \sup_{x \in \mathbb{R}^n} g(x, y)$$

$$\nabla_x g(x, y) = y - \nabla f_{\mathrm{cshub}}(x) = 0$$

$$y = \begin{cases} x & \|x\|_2 \leqslant 1 \\ \frac{x}{\|x\|_2} & \|x\|_2 > 1 \end{cases}$$

Отсюда получаем, что $\|y\|_2 \leqslant 1$ при любом x, так как $\left\|\frac{x}{\|x\|_2}\right\|_2 = 1$. Значит $\mathrm{dom}\, f^*(y) = \{y :\, y \in \mathbb{R}^n,\, \|y\|_2 \leqslant 1\}$.
При этом сама $f^*(y) = \max\left\{\langle y,\, y\rangle - \frac{1}{2}\|y\|_2^2,\, \langle y,\, y\rangle \cdot \|x\|_2 - \|x\|_2 + \frac{1}{2}\right\} = \max\left\{\frac{1}{2}\|y\|_2^2,\, \frac{1}{2}\right\} = \frac{1}{2}\|y\|_2^2$.
Ответ: $f^*(y) = \frac{1}{2}\|y\|_2^2$, $\mathrm{dom}\, f^* = \{y :\, y \in \mathbb{R}^n,\, \|y\|_2 \leqslant 1\}$.