Теория оптимизации 2023-2024

# ККТ и двойственность

Забара Илья

## Задача 1

**1)** Найдём множество допустимых значений ф-ии.

$(x-2)(x-4) \leq 0 \Rightarrow x \in [2;4] \Rightarrow f(x) \in [2^2+1;4^2+1] \Rightarrow \boxed{f(x) \in [5;17]}$

$L = x^2+1+\lambda(x-2)(x-4) = x^2+1+\lambda(x^2+6x+8)$

Задача выпуклая, условия ККТ справедливы в данной задаче.

$\cdot \nabla_x L = 0 \Rightarrow 2x+2\lambda x - 6\lambda = 0 \Rightarrow \lambda = \dfrac{x}{3-x}$

$\cdot (x-2)(x-4) \leq 0$

$$\cdot \lambda(x-2)(x-4) = 0 \Rightarrow \left[\begin{array}{c} \lambda = 0 : x = 0 \notin S; \\ x = 2 : \lambda = 2; \\ x = 4 : \lambda = -4 < 0 \end{array}\right.$$

Итого имеем, что $\boxed{x^* = 2, f(x^*) = 5}$.

**2)** Во втором пункте строим требуемый график (Рис.1).

По графику видим, что действительно $p^* \geq \inf_x L(x,\lambda)$.

$g(\lambda) = \min\limits_x L(x,\lambda) = \min\limits_x (x^2+1+\lambda(x^2-6x+8))$

$\nabla_x L = x + \lambda x - 3\lambda = 0 \Rightarrow \tilde{x} = \dfrac{3\lambda}{1+\lambda}$

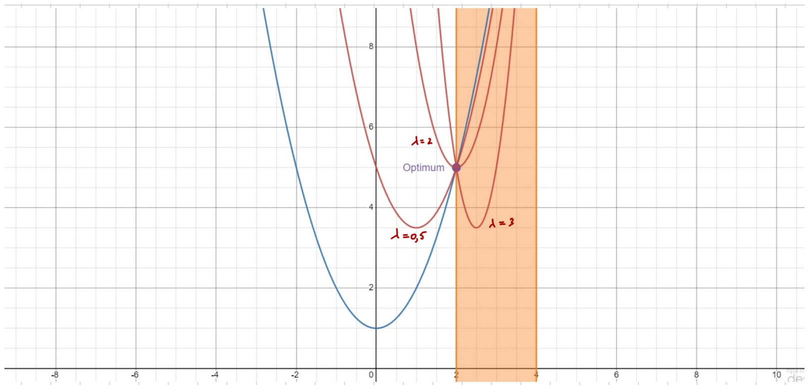


Рис. 1

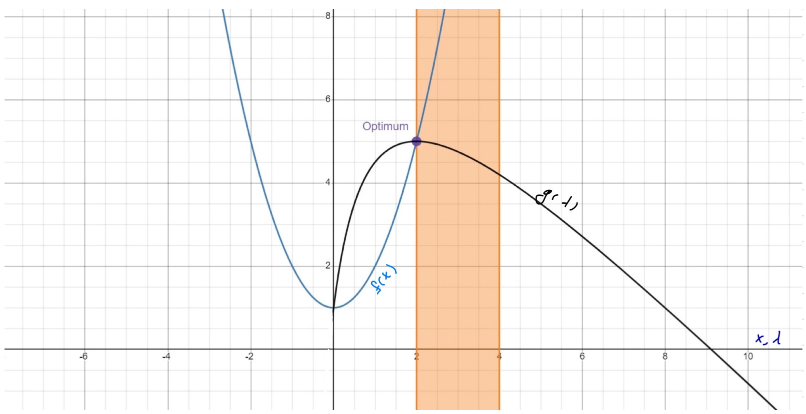


Рис. 2

Тогда $g(\lambda) = L(\tilde{x}, \lambda) = \dfrac{-\lambda^2 + 9\lambda + 1}{\lambda + 1}$. График этой функции изображен на Рис.2.

Действительно видно, что зазор двойственности нулевой и при $\lambda = 2$ имеем оптимум.

**3)** $g(\lambda) = \dfrac{-\lambda^2 + 9\lambda + 1}{\lambda + 1} \longrightarrow \max\limits_{\lambda}$ при неотрицательных $\lambda$ есть двойственная задача.

$$\nabla_\lambda g = \frac{9}{(1+\lambda)^2} - 1 = 0 \Rightarrow \left[\begin{array}{l} \lambda = 2 \\ \lambda = -4 \end{array}\right.$$

$\lambda = -4$ не подхожит из неотрицательности $\Rightarrow \lambda^* = 2, g(\lambda^*) = g^* = d^* = 5 = p^* \Rightarrow$ сильная двойственность сохраняется.

**4)** $\begin{cases} x^2+1 \longrightarrow \min\limits_x \\ (x-2)(x-4)-u \leq 0 \end{cases}$

$L = x^2 + 1 + \lambda(x^2 - 6x + 8) - \lambda u$

$\cdot L_x^{'} = 2x + 2\lambda x - 6\lambda = 0 \Rightarrow \lambda = \dfrac{x}{3-x}$

Условие дополняющей нежёсткости:

$\cdot \lambda(x^2 - 6x + 8 - u) = 0 \Rightarrow \left[ \begin{array}{c} \lambda = 0 \Rightarrow x = 0 \rightsquigarrow u \geq 8 \\ x^2 - 6x + 8 = u \Rightarrow x = 3 \pm \sqrt{u+1} \rightsquigarrow u \geq -1 \end{array} \right.$

Для обоих корней второго случая ограничения-неравенства выполняются. Проверим, дают ли они $\lambda \geq 0$.

Если $x = 3 + \sqrt{u+1}$, то $\lambda = \dfrac{3+\sqrt{1+u}}{-\sqrt{1+u}} < 0 \;\; \forall u$.

Если $x = 3 - \sqrt{u+1}$, то $\lambda = \dfrac{3-\sqrt{1+u}}{\sqrt{1+u}} \geq 0$ при $u \leq 8$.

Таким образом, при $u \in [-1; 8]$ имеем: $p^*(u) = (3 - \sqrt{u+1})^2 + 1 = u - 6\sqrt{u+1} + 11$.

При $u \geq 8$ имеем: $x = 0 \Rightarrow p^*(u) = 0^2 + 1 = 1$.

Итого, получаем: $\boxed{p^*(u) = \begin{cases} u - 6\sqrt{u+1} + 11, & u \in [-1; 8] \\ 1, & u > 8 \end{cases}}$.

В остальных случаях ($u < -1$) область ограничений-неравенств становится пустой.

График $p^*(u)$ изображен на Рис.3.

Осталось проверить последнее утверждение.

$\dfrac{dp}{du}(0) = \left(1 - \dfrac{6}{2\sqrt{u+1}}\right)\Bigg|_{\hat{u}=0} = 1 - 3 = -2 = -\lambda^*$.

Действительно, утверждение из условия верное.

**Ответ:** все пункты рассмотрены.

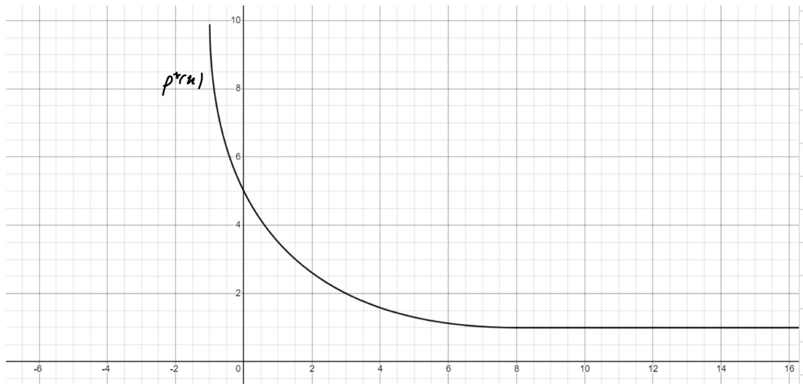


Рис. 3

# Задача 2

За небольшой штраф по баллам разрешено $\|\cdot\|$ заменить на $\|\cdot\|_2$. Далее под нормой подразумевается вторая норма. Нужно выписать двойственную задачу.

$L = \dfrac{1}{2}\|y-b\|^2 + \dfrac{\lambda}{2}\|x\|^2 + \nu^T(Ax-y)$

$g(\nu) = \min\limits_{x,y}(\dfrac{1}{2}\|y-b\|^2 + \dfrac{\lambda}{2}\|x\|^2 + \nu^T(Ax-y))$

$\nabla_y L = y - b - \nu$

$\nabla_x L = \lambda x + A^T\nu$

$$\nabla L = 0 \Leftrightarrow \begin{cases} y - b - \nu = 0 \\ \lambda x + A^T\nu = 0 \end{cases} \Rightarrow \begin{cases} y^* = b + \nu \\ x^* = -\dfrac{1}{\lambda}A^T\nu \end{cases}$$

В итоге имеем $g(\nu) = L(x^*, y^*, \nu) = \dfrac{1}{2}\|\nu\|^2 + \dfrac{1}{2\lambda}\|A^T\nu\|^2 - \nu^T\left(\dfrac{1}{\lambda}AA^T\nu + b + \nu\right)$.

Двойственная задача:

$$\begin{cases} g(\nu) = L(x^*, y^*, \nu) = \dfrac{1}{2}\|\nu\|^2 + \dfrac{1}{2\lambda}\|A^T\nu\|^2 - \nu^T\left(\dfrac{1}{\lambda}AA^T\nu + b + \nu\right) \longrightarrow \max\limits_{\nu} \\ \nu \succcurlyeq 0 \end{cases}$$

**Ответ:** система выше.

# Задача 3

За небольшой штраф по баллам разрешено $\|\cdot\|$ заменить на $\|\cdot\|_2$. Далее под нормой подразумевается вторая норма. Тут, как и в прошлой задаче, нужно выписать двойственную задачу.

$L = \langle 1, t\rangle + \dfrac{\lambda}{2}\|x\|^2 + \nu_1^T(1 - t - Ax) + \nu_2^T(-t)$

$g(\nu_1, \nu_2) = \min\limits_{x,t}(\langle 1, t\rangle + \dfrac{\lambda}{2}\|x\|^2 + \nu_1^T(1 - t - Ax) + \nu_2^T(-t)) =$
$\min\limits_{x,t}(\langle 1 - \nu_1 - \nu_2, t\rangle + \dfrac{\lambda}{2}\|x\|^2 - \nu_1^T Ax + \nu_1^T 1)$

Первое слагаемое даёт конечный минимум всего выражение только тогда, когда равно нулю. Далее надо рассмотреть следующие два слагаемых, зависящих от $x$. Очевидно, это выпуклая функция.

$\nabla\left(\dfrac{\lambda}{2}\langle x, x\rangle - \nu_1^T Ax\right) = \lambda x - A^T\nu_1 = 0 \Rightarrow x^* = \dfrac{1}{\lambda}A^T\nu_1$

Тогда минимум той части, которая зависит от $x$ будет равен $\dfrac{1}{2\lambda}\langle A^T\nu_1, A^T\nu_1\rangle - \dfrac{1}{\lambda}\nu_1^T AA^T\nu_1$

Итого, имеем двойственную функцию:

$g(\nu_1, \nu_2) = \begin{cases} \dfrac{1}{2\lambda}\langle A^T\nu_1, A^T\nu_1\rangle - \dfrac{1}{\lambda}\nu_1^T AA^T\nu_1 + \nu_1^T 1, & \text{при } \langle 1 - \nu_1 - \nu_2, t\rangle = 0 \\ \infty, & \text{otherwise} \end{cases}$

Двойственная задача:

$\begin{cases} g(\nu_1, \nu_2) \longrightarrow \max\limits_{\nu_1, \nu_2} \\ \nu_1 \succcurlyeq 0 \\ \nu_2 \succcurlyeq 0 \end{cases}$

**Ответ:** выше.

# Задача 4

Предъявим $x_i = \begin{cases} 1, & i = arg\min\limits_{k}(c_k) \\ 0, & else \end{cases}$. Покажем, что это то, что нужно.

Пусть минимальный элемент вектора $c$ есть $c_i$. Рассмотрим $c^T x$ (целевую функцию) и оценим её снизу:

$c^T x = \sum\limits_{j=1}^{n} c_j x_j \geq \sum\limits_{j=1}^{n} c_i x_j = c_i \sum\limits_{j=1}^{n} x_j = c_i \cdot 1 = c_i$

Таким образом показано, что любой другой вектор даёт большее значение целевой функции.

**Ответ:** вектор размерности $n$ из $(n-1)$ нулей и одной единицы, стоящей на позиции, которая совпадает с позицией минимального элемента вектора $c$. Если у вектора $c$ таких элементов несколько, то выбираем любой из них.

# Задача 5

Целевая функция выпуклая, ограничения аффинные $\Rightarrow$ имеем дело с выпуклой задачей $\Rightarrow$ существует единственное решение.

Условия Слейтера выполняются, ограничения линейные $\Rightarrow$ используем ККТ.

$L = \langle C^{-1}, X \rangle - lndetX + \lambda(\langle Xa, a \rangle - 1)$

$\nabla_X L = \nabla_X(\langle C^{-1}, X \rangle - lndetX) + \lambda \nabla_X(\langle Xa, a \rangle) = C^{-1} - X^{-T} + \lambda aa^T = 0 \Rightarrow X^{-T} = C^{-1} + \lambda aa^T$

По формуле $(A + uv^T) = A^{-1} - \dfrac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1} u}$ имеем

$X^T = X_{opt} = C + \dfrac{\lambda C aa^T C}{1 + \lambda a^T C a}$, т.к. $C \in \mathbb{S}^n_{++}$ и $aa^T \in \mathbb{S}^n_{++}$.

Условие дополняющей нежёсткости:

$\lambda(\langle Xa, a \rangle - 1) = 0 \Rightarrow \left[ \begin{array}{c} \lambda = 0 \\ \langle Xa, a \rangle = 1, \quad \text{при } \lambda > 0 \end{array} \right.$

**Рассмотрим сначала второй случай.** Здесь и далее под $X$ подразумевается $X_{opt}$.

$a^TXa = a^TCa + \dfrac{\lambda a^TCaa^TCa}{1+\lambda a^TCa} = 1 \Rightarrow 1 + \dfrac{\lambda a^TCa}{1+\lambda a^TCa} = \dfrac{1}{a^TCa} \Rightarrow$
$\boxed{\lambda = 1 - \dfrac{1}{a^TCa}}$

При этом должно выполняться $\lambda \geq 0 \Rightarrow \boxed{a^TCa \geq 1}$.

В этом случае решение принимает вид:

$$X_{opt} = C + \frac{\left(1 - \dfrac{1}{a^TCa}\right)Caa^TC}{1 + \left(1 - \dfrac{1}{a^TCa}\right)a^TCa}.$$

Итого, после преобразований:

$$\boxed{X_{opt} = C + \left(\frac{1}{a^TCa} - \frac{1}{(a^TCa)^2}\right)Caa^TC, \quad a^TCa \geq 1}$$

**Рассмотрим теперь первый случай.** $\lambda = 0, \quad a^TXa < 1 \Rightarrow X_{opt} = C$.

Тогда при этом должно выполняться $a^TCa < 1$.

Итого, $\boxed{X_{opt} = C, \quad a^TCa < 1}$

**Ответ:** $X_{opt} = \begin{cases} X_{opt} = C + \left(\dfrac{1}{a^TCa} - \dfrac{1}{(a^TCa)^2}\right)Caa^TC, & a^TCa \geq 1 \\ X_{opt} = C, \quad a^TCa < 1 \end{cases}.$

# Задача 6

Задача выпуклая $\Rightarrow$ используем ККТ.

$L = c^Tx + \lambda((x-x_c)^TA(x-x_c)-1)$

$\nabla L = c + \lambda\nabla(\langle x - x_c, Ax - Ax_x\rangle) = c + 2\lambda A(x-x_c) = 0.$

Условие дополняющей нежёсткости:

$\lambda((x-x_c)^TA(x-x_c)-1) = 0 \Rightarrow \left[\begin{array}{c} \lambda = 0 \\ (x-x_c)^TA(x-x_c) = 1, \quad \lambda > 0 \end{array}\right.$

**Рассмотрим первый случай**. $\lambda = 0$ даёт $c + 0 = 0 \Rightarrow c = 0$, что противоречит тому, что $c \neq 0$.

**Рассмотрим второй случай**. $(x - x_c)^T A(x - x_c) = 1, \quad \lambda > 0$.

$c + 2\lambda A(x - x_c) = 0 \Rightarrow A(x - x_c) = -\dfrac{c}{2\lambda} \Rightarrow x^* = x_c - A^{-1}\dfrac{c}{2\lambda}$

$\left(A^{-1}\dfrac{c}{2\lambda}\right)^T A \left(A^{-1}\dfrac{c}{2\lambda}\right) = 1$

Т.к. $A^{-1} \in (S)^n_{++}$, то уравнение выше принимает вид:

$\dfrac{1}{4\lambda^2} c^T A^{-1} c = 1 \Rightarrow \lambda = \dfrac{1}{2}\sqrt{c^T A^{-1} c}$ (взяли только тот корень, который $> 0$).

Итого, имеем решение:

$$x^* = x_c - A^{-1}\frac{c}{2 \cdot \frac{1}{2}\sqrt{c^T A^{-1} c}} \Rightarrow \boxed{x^* = x_c - A^{-1}\frac{c}{\sqrt{c^T A^{-1} c}}}$$

Матрица $A^{-1}$ существует, т.к. исходная матрица симметричная и положительно определённая.

**Ответ:** в рамке.

# Задача 7

Нужно выписать ККТ для исходной и двойственной задач.

**1) Primal**

Исходная ф-я выпукла, ограничения аффинны. Используем ККТ:

$L = \|Ax - b\|^2 + \nu^T(Cx - d)$

$\nabla L = 2A^T(Ax - b) + C^T\nu = 0$

$Cx = d$

Итого, $\boxed{\begin{cases} 2A^T(Ax^* - b) + C^T\nu = 0 \\ Cx^* = d \end{cases}}$ – система ККТ для нахождения $x^*$.

Решение этой системы завсит от размера матриц (соотношение кол-ва строк и столбцов).

**2) Dual**

$g(\nu) = \min_x(\|Ax-b\|^2 + \nu^T(Cx-d))$

Найдём $x$, задающий минимум.

$\nabla(\|Ax-b\|^2 + \nu^T(Cx-d)) = 2A^T(Ax-b) + C^T\nu = 0 \Leftrightarrow x = \frac{1}{2}(A^TA)^{-1}(2A^Tb - C^T\nu)$

Тогда:

$$g(\nu) = \|1/2A(A^TA)^{-1}(2A^Tb - C^T\nu) - b\|^2 + \nu^T(C(A^TA)^{-1}(2A^Tb - C^T\nu) - d)$$

Двойственная задача может быть поставлена как $-g(\nu) \longrightarrow \min_\nu$ при положительных аргументах.

Выпишем для неё ККТ.

$\nabla(-\|1/2A(A^TA)^{-1}(2A^Tb - C^T\nu) - b\|^2 + \nu^T(C(A^TA)^{-1}(2A^Tb - C^T\nu) - d)) =$

$= (A(A^TA)^{-1}C^T)^T\left(\frac{1}{2}A(A^TA)^{-1}(b - 2A^Tb - C^T\nu)\right) = 0$

Итого,

$\boxed{\begin{cases} (A(A^TA)^{-1}C^T)^T\left(\frac{1}{2}A(A^TA)^{-1}(b - 2A^Tb - C^T\nu)\right) = 0 \\ \nu \succcurlyeq 0 \end{cases}}$ – система ККТ для нахождения $\nu^*$.

**Ответ:** в рамках.

# Задача 8

**1)** Во первых, $X \succ 0$, иначе логарифм не был бы определён.

$L = trX - lndetX + \nu^T(Xs - y)$

$\nabla_X L = I - X^{-1} + \nabla_X(\nu^T X s) = I - X^{-1} + \frac{1}{2}(\nu s^T + s\nu^T) = 0 \Rightarrow X^{-1} = I + \frac{1}{2}(\nu s^T + s\nu^T)$

$s = X^{-1}y = y + \frac{1}{2}(\nu + (\nu^T y)s)$

Если взять от обоих частей внешнее произведение и учесть $s^T y = 1$, то получим $\nu^T y = 1 - y^T y$.

Тогда, из приведённых выше соотношений имеем $\nu = (1+y^T y)s - 2y \Rightarrow X^{-1} = I + \frac{1}{2}(((1+y^T y)s - 2y)s^T + s((1+y^T y)s - 2y)^T)$.

Итого, для данной задачи условия из ККТ принимают вид:

$$\boxed{\begin{cases} X^{-1} = I + \frac{1}{2}(((1+y^T y)s - 2y)s^T + s((1+y^T y)s - 2y)^T) \\ X \succ 0 \\ Xs = y \end{cases}}.$$

**2)** Осталось убедиться, что $X^{-1}X^* = I$.

$X^{-1}X^* = X^{-1} = (I + \frac{1}{2}(((1+y^T y)s - 2y)s^T + s((1+y^T y)s - 2y)^T)) \cdot (I + yy^T - \frac{1}{s^T s}ss^T) = (I + (1+y^T y)ss^T - ys^T - sy^T) \cdot (I + yy^T - \frac{1}{s^T s}ss^T) = |$после обширных арифметических преоразований на бумаге$| = I$.

**Ответ:** в рамке, выражение для $X^*$ удовлетворяет $X^{-1}X^* = I$.

# Задача 10

Сопряжённая ф-я: $f^*(y) = \sup\limits_x (x^T y - f(x)) = -\inf\limits_x (f(x) - x^T y)$

$L = c^T x + \lambda f(x)$

$\inf\limits_x L = \inf\limits_x (c^T x + \lambda f(x)) = \lambda \inf\limits_x (f(x) - \langle -\frac{c}{\lambda}, x \rangle) = -\lambda f^*(-\frac{c}{\lambda})$.

Это означает $g(\lambda) = -\lambda f^*(-\frac{c}{\lambda})$– двойственная ф-я, где $f^*(-\frac{c}{\lambda})$– сопряжённая к функции в ограничении-неравенстве.

Тогда двойственная задача:

$$\begin{cases} g(\lambda) = -\lambda f^*(-\frac{c}{\lambda}) \longrightarrow \max\limits_\lambda \\ \lambda \geq 0 \end{cases}.$$

$f^*(-\dfrac{c}{\lambda})$ всегда выпуклая, так как сопряжённая (неважен характер выпуклости исхоной ф-ии) $\Rightarrow$ целевая ф-я двойственной задачи вогнутая. Ограничения-неравенства аффинны. Итого получаем, что двойственная задача в терминах $f^*$ выпуклая.

**Ответ:** доказано.

# Задача 11

**1)** Функция $\|Ax-b\|_2^2$ дифф-ма и градиент $\nabla\|Ax-b\|_2^2 = 2A^T(Ax-b)$.

Решение $\tilde{x}$ удовлетворяет уравнению:

$$\nabla\phi(\tilde{x}) = \nabla f_0(x) + \alpha\cdot 2A^T(A\tilde{x}-b) = 0$$

Лагранжиан исходной задачи:

$$L = f_0(x) + \nu^T(Ax-b) \Rightarrow \nabla L = \nabla f_0(x) + A^T\nu^* = 0.$$

Из последних двух соотношений имеем:

$$2\alpha(A\tilde{x}-b) = \nu^* \Rightarrow \boxed{\nu^* = 2\alpha(A\tilde{x}-b)}.$$

**2)** Двойственная функция у исходной задачи:

$$g(\nu) = \inf_x(f_0(x) + \nu^T(Ax-b)).$$

Нижняя граница:

$g(\nu^*) = \inf_x(f_0(x) + (\nu^*)^T(Ax-b)) = f_0(\tilde{x}) + (\nu^*)^T(A\tilde{x}-b) = f_0(\tilde{x}) + 2\alpha(A\tilde{x}-b)^T(A\tilde{x}-b) = f_0(\tilde{x}) + 2\alpha\|A\tilde{x}-b\|_2^2)$.

Итого, нижняя оценка:

$$\boxed{g(\nu^*) = f_0(\tilde{x}) + 2\alpha\|A\tilde{x}-b\|_2^2)}.$$

**Ответ:** в рамках.

## Задача 12

Нужно выписать двойственную задачу. По указанию делаем замену $y_i = b_i - a_i^T x$. Тогда задача приобретает вид:

$$\begin{cases} -\sum\limits_{i=1}^{m} ln y_i \longrightarrow min \\ y = b - Ax \end{cases}.$$

Строки матрицы A есть векторы $a_i$, её размер $mxn$. Лагранжиан:

$$L = -\sum_{i=1}^{m} ln y_i + \nu^T (y - b + Ax).$$

Тогда двойственная функция:

$$g(\nu) = \inf_{x,y}(-\sum_{i=1}^{m} ln y_i + \nu^T (y - b + Ax))$$

$$= \inf_{x,y}(-\sum_{i=1}^{m} ln y_i + \nu^T y - \nu^T b + \nu^T Ax)$$

Слагаемое $\nu^T Ax = \langle A^T \nu, x \rangle$ линейна по $x \Rightarrow$ двойственная ф-ия будет конечна только при $A^T \nu = 0$.

Градиент той части, которая зависит от $y$ будет равен:

$$\nabla \left( \sum_{i=1}^{m} ln y_i + \nu^T y \right) = \frac{1}{y} + \nu,$$

где пераое слагаемое есть вектор, все элементы которого обратны элементам вектора $y$.

Приравняв к нулю этот градиент получим условие на переменные, дающие минимум ф-ии: $\nu_i = \dfrac{1}{y_i}$. Тогда

$$\left( -\sum_{i=1}^{m} ln y_i + \nu^T y \right) = \sum_{i=1}^{m} ln \nu_i + \sum_{i=1}^{m} 1 = \sum_{i=1}^{m} ln \nu_i + m$$

Итого, двойственная функция:

$$g(\nu) = \begin{cases} \sum\limits_{i=1}^{m} ln\nu_i + m - b^T\nu, & A^T\nu = 0 \\ -\infty, & \text{else} \end{cases}$$

И, окончательно, двойственная задача:

$$\boxed{\begin{cases} g(\nu) = \sum\limits_{i=1}^{m} ln\nu_i + m - b^T\nu \longrightarrow \max\limits_{\nu} \\ A^T\nu = 0 \end{cases}}$$

**Ответ:** в рамке.