1 Выпуклые множества

1. Покажем выпуклость множества $\mathcal{C}$ по определению. Требуется доказать, что для любых двух точек $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ и любого $\theta \in [0;1]$ их выпуклая комбинация также принадлежит множеству $\mathcal{C}$: $\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2 \in \mathcal{C}$.

   Рассмотрим квадратичную функцию $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} + \mathbf{b}^\mathsf{T}\mathbf{x} + \mathbf{c}$, при этом $\mathbf{A} \succ 0$. Множество $\mathcal{C}$ задано неравенством $f(\mathbf{x}) \leqslant 0$, поэтому требуется показать, что $\forall \mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ $\forall \theta \in [0;1]$ выполнено $f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leqslant 0$.

   Функция $f(x)$ является выпуклой по критерию второго порядка: матрица Гессе функции положительно определен: $\mathbf{H}_f(\mathbf{x}) = \mathbf{A} + \mathbf{A}^\mathsf{T} \succ 0$.

   Для выпуклых функций выполняется: $\forall \theta \in [0;1]$ $f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leqslant \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2)$. Так как $f(\mathbf{x}_1) \leqslant 0$ и $f(\mathbf{x}_2) \leqslant 0$, получаем: $f(\theta\mathbf{x}_1 + (1-\theta)\mathbf{x}_2) \leqslant \theta f(\mathbf{x}_1) + (1-\theta) f(\mathbf{x}_2) \leqslant 0$. Это доказывает выпуклость множества $\mathcal{C}$.

2. Рассматривается множество $\mathcal{X} = \{\mathbf{X}\mathbf{X}^\mathsf{T} \mid \mathbf{X} \in \mathbb{R}^{n\times k}, \operatorname{rank}(\mathbf{X}) = k\}$. Необходимо описать его коническую оболочку $\mathbf{coni}(\mathcal{X}) = \{\sum_{i=1}^m \alpha_i x_i \mid x_i \in \mathcal{X}, \alpha_i \in \mathbb{R}_+, m \in \mathbb{N}\}$.

   Матрицы $\mathbf{X}\mathbf{X}^\mathsf{T} \in \mathbb{R}^{n\times n}$ являются неотрицательно определенными ранга k ($k \leq n$), оба этих факта легко можно показать, например, используя (полное) сингулярное разложение $\mathbf{X}$: $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}$, где $\mathbf{U} \in \mathbb{R}^{n\times n}$, $\mathbf{V} \in \mathbb{R}^{k\times k}$ — ортогональные матрицы левых и правых сингулярных векторов, а $\mathbf{\Sigma} \in \mathbb{R}^{n\times k}$ — прямоугольная диагональная матрица с положительными сингулярными значениями $\sigma_1,...,\sigma_k$ на главной диагонали (так как матрица имеет ранг k). Тогда $\mathbf{X}\mathbf{X}^\mathsf{T} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\mathsf{T}\mathbf{V}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T} = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^\mathsf{T}\mathbf{U}^\mathsf{T}$, здесь $\mathbf{\Sigma}\mathbf{\Sigma}^\mathsf{T} \in \mathbb{R}^{n\times n} = \operatorname{diag}(\sigma_1^2, ..., \sigma_k^2, 0, ..., 0)$. Так как ортогональные преобразования сохраняют ранг, а диагональная матрица имеет ранг ровно k, то и $\mathbf{X}\mathbf{X}^\mathsf{T}$ имеет ранг k. Таким образом доказано, что $\mathbf{X}\mathbf{X}^\mathsf{T} \succeq 0$ и $\operatorname{rank}(\mathbf{X}\mathbf{X}^\mathsf{T}) = k$.

   Линейная комбинация описанных выше матриц с нулевыми коэффициентами естественным образом дает нулевую матрицу. Линейная комбинация таких матриц с хотя бы одним положительным коэффициентом имеет ранг не меньше k, вплоть до n. Докажем это. Рассмотрим $\mathbf{A}_1,...,\mathbf{A}_m \in \mathcal{X}$, пусть их коническая комбинация $\mathbf{A} = \sum_{i=1}^m \alpha_i \mathbf{A}_i \in \mathbf{coni}(\mathcal{X})$ имеет ранг $0 \leqslant r \leqslant n$ (очевидные неравенства следуют из размерности). Для определенности пусть все коэффициенты в комбинации больше нуля. Возьмем матрицу $\mathbf{N} \in \mathbb{R}^{n\times(n-r)}$, принадлежащую ядру $\mathbf{A}$: $\mathbf{N}^\mathsf{T}\mathbf{A}\mathbf{N} = \sum_{i=1}^m \alpha_i \mathbf{N}^\mathsf{T}\mathbf{A}_i\mathbf{N} = 0$. Из неотрицательной определенности $\mathbf{A}_i$ следует, что $\mathbf{N}^\mathsf{T}\mathbf{A}_i\mathbf{N} = 0$, $i = \overline{1,m}$, отсюда получаем $k = \operatorname{rank}(\mathbf{A}_i) \leqslant r$, $i = \overline{1,m}$. Таким образом $\operatorname{rank}(\mathbf{A}) \geqslant k$.

   Алгоритм получения любой неотрицательно определенной матрицы $\mathbf{A}$: $\operatorname{rank}(\mathbf{A}) \in \{0\} \cup [k;n]$ также можно показать через сингулярное разложение: выберем из $\mathcal{X}$ такие матрицы $\mathbf{A}_i$, что их матрицы левых и правых сингулярных векторов одинаковы и могут быть дополнены (в полном разложении) до матриц левых и правых сингулярных векторов $\mathbf{A}$, а матрица сингулярных чисел $\mathbf{A}$ «составляется» как коническая оболочка матриц сингулярных значений $\mathbf{A}_i$. То есть $\mathbf{A} = \mathbf{U}_0\mathbf{\Sigma}_0\mathbf{V}_0^\mathsf{T}$, $\mathbf{A}_i = \mathbf{U}_0\mathbf{\Sigma}_i\mathbf{V}_0^\mathsf{T}$, $\mathbf{\Sigma}_0 = \sum_i \alpha_i \mathbf{\Sigma}_i$.

   Было доказано, что $\mathbf{coni}(\mathcal{X})$ — это множество неотрицательно определенных матриц ранга от k до n и нулевая матрица.

3. Перспективное отображение множества: $P(C) = \{\mathbf{x}/t \mid (\mathbf{x}, t) \in C, t > 0\}$. Рассмотрим два множества C и опишем их перспективное отображение по определению, в максимально простой форме:

(a) гиперплоскость $C = \{(\mathbf{x}, t) \,|\, \mathbf{a}^\top \mathbf{x} + ct = \gamma\}$, $\mathbf{a}$ и c не равны нулю одновременно.

$$P(C) = \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c = \gamma/t, t > 0\} = \begin{cases} \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c > 0\}, & \gamma > 0; \\ \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c < 0\}, & \gamma < 0; \\ \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c = 0\}, & \gamma = 0. \end{cases}$$

(b) полупространство $C = \{(\mathbf{x}, t) \,|\, \mathbf{a}^\top \mathbf{x} + ct \leqslant \gamma\}$, $\mathbf{a}$ и c не равны нулю одновременно.

$$P(C) = \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c \leqslant \gamma/t, t > 0\} = \begin{cases} \mathbb{R}^n, & \gamma > 0; \\ \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c < 0\}, & \gamma < 0; \\ \{\mathbf{y} \,|\, \mathbf{a}^\top \mathbf{y} + c \leqslant 0\}, & \gamma = 0. \end{cases}$$

4. Докажем, что множество $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \,|\, \mathbf{x}^\top \mathbf{A} \mathbf{x} \leqslant (\mathbf{c}^\top \mathbf{x})^2, \mathbf{c}^\top \mathbf{x} > 0\}$, где $\mathbf{A} \in \mathbf{S}^n_+$, выпукло.

   Для этого воспользуемся определением выпуклости. Рассмотрим точки $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$ и их выпуклую комбинацию $\mathbf{x}_\theta = \theta \mathbf{x}_1 + (1-\theta)\mathbf{x}_2$, $\theta \in [0; 1]$. Требуется доказать, что $\mathbf{x}_\theta \in \mathcal{C}$, то есть $\mathbf{x}_\theta^\top \mathbf{A} \mathbf{x}_\theta \leqslant (\mathbf{c}^\top \mathbf{x}_\theta)^2$ и $\mathbf{c}^\top \mathbf{x}_\theta > 0$.

   В пункте 1 было показано, что $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ — выпуклая функция для $\mathbf{A} \succ 0$, то есть $\mathbf{x}_\theta^\top \mathbf{A} \mathbf{x}_\theta \leqslant \theta \mathbf{x}_1^\top \mathbf{A} \mathbf{x}_1 + (1-\theta)\mathbf{x}_2^\top \mathbf{A} \mathbf{x}_2$. Так как $\mathbf{c}^\top \mathbf{x}$ — линейное преобразование, оно является выпуклым. Легко видеть, что второе неравенство выполняется: $\mathbf{c}^\top \mathbf{x}_\theta = \theta \mathbf{c}^\top \mathbf{x}_1 + (1-\theta)\mathbf{c}^\top \mathbf{x}_2 > 0$. Напрямую выпуклость функций для доказательства первого неравенства использовать не получится, так как правая часть переносится налево со знаком минус.

   Распишем первое неравенство после подстановки $\mathbf{x}_\theta$. Его левая часть: $\mathbf{x}_\theta^\top \mathbf{A} \mathbf{x}_\theta = \theta^2 \mathbf{x}_1^\top \mathbf{A} \mathbf{x}_1 + 2\theta(1-\theta)\mathbf{x}_1^\top \mathbf{A} \mathbf{x}_2 + (1-\theta)^2 \mathbf{x}_2^\top \mathbf{A} \mathbf{x}_2$. Правая часть: $(\mathbf{c}^\top \mathbf{x}_\theta)^2 = \theta^2 (\mathbf{c}^\top \mathbf{x}_1)^2 + 2\theta(1-\theta)\mathbf{c}^\top \mathbf{x}_1 \mathbf{c}^\top \mathbf{x}_2 + (1-\theta)^2 (\mathbf{c}^\top \mathbf{x}_2)^2$. Так как $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{C}$, выполняется $\mathbf{x}_1^\top \mathbf{A} \mathbf{x}_1 \leqslant (\mathbf{c}^\top \mathbf{x}_1)^2$ и $\mathbf{x}_2^\top \mathbf{A} \mathbf{x}_2 \leqslant (\mathbf{c}^\top \mathbf{x}_2)^2$. Остается доказать неравенство для вторых слагаемых.

   Докажем $\mathbf{x}_1^\top \mathbf{A} \mathbf{x}_2 \leqslant \mathbf{c}^\top \mathbf{x}_1 \mathbf{c}^\top \mathbf{x}_2$, используя неравенство Коши-Буняковского (рассмотрим левую часть как скалярное произведение). Тогда $\mathbf{x}_1^\top \mathbf{A} \mathbf{x}_2 \leqslant \sqrt{\mathbf{x}_1^\top \mathbf{A} \mathbf{x}_1}\sqrt{\mathbf{x}_2^\top \mathbf{A} \mathbf{x}_2} < \mathbf{c}^\top \mathbf{x}_1 \mathbf{c}^\top \mathbf{x}_2$, здесь также используется $\mathbf{c}^\top \mathbf{x} > 0$. Таким образом доказано что $\mathbf{x}_\theta^\top \mathbf{A} \mathbf{x}_\theta \leqslant (\mathbf{c}^\top \mathbf{x}_\theta)^2$ и $\mathbf{c}^\top \mathbf{x}_\theta > 0$, то есть $\mathbf{x}_\theta \in \mathcal{C}$. Значит, $\mathcal{C}$ выпукло.

2 Двойственные конусы

1. Покажем, что экспоненциальный конус $\mathcal{K} = \{(x, y, z) \in \mathbb{R} \times \mathbb{R}_{++} \times \mathbb{R}_+ \,|\, y e^{x/y} \leqslant z\}$ является выпуклым конусом, найдем его замыкание $\mathbf{cl}(\mathcal{K})$ и двойственный к нему конус $(\mathbf{cl}(\mathcal{K}))^*$.

   (a) Докажем, что замыканием конуса $\mathbf{cl}(\mathcal{K})$ является множество:

   $$\mathcal{K}_c = \mathcal{K} \cup \mathcal{B}, \text{ где } \mathcal{B} = ((-\mathbb{R}_+) \times \{0\} \times \mathbb{R}_+)$$

   Для доказательства данного факта покажем, что $\mathcal{K}_c \subseteq \mathbf{cl}(\mathcal{K})$ и $\mathcal{K}_c \supseteq \mathbf{cl}(\mathcal{K})$.

   Начнем с доказательства $\mathcal{K}_c \subseteq \mathbf{cl}(\mathcal{K})$. Пусть $\mathbf{u} \in \mathcal{K}_c$. Это значит, что $\mathbf{u}$ лежит в $\mathcal{K}$ или в $\mathcal{B}$. Но если $\mathbf{u} \in \mathcal{K}$, то автоматически $\mathbf{u} \in \mathbf{cl}(\mathcal{K})$. Тогда возьмем $\mathbf{u} \in \mathcal{B}$. Покажем, что $\mathbf{u} \in \mathbf{cl}(\mathcal{K})$. Для этого рассмотрим последовательность точек $\{\mathbf{u}^{(k)}\}_{k=1}^\infty \subset \mathcal{K}$ и покажем, что ее предельной точкой является $\mathbf{u}$:

   $$\mathbf{u}^{(k)} = \big(u_1 - 1/k, \quad 1/k^2, \quad u_3 + 1/k\big)$$

Подставим данную последовательность в определение множества $\mathbf{cl}(\mathcal{K})$ и проведем оценки:

$$u_2^{(k)} \exp\left(u_1^{(k)}/u_2^{(k)}\right) = u_2^{(k)} \exp\left(k^2 u_1 - k\right) \leqslant u_2^{(k)} \exp(-k) \leqslant u_2^{(k)}/e < u_2^{(k)} \leqslant 1/k + u_3 = u_3^{(k)}.$$

Таким образом было показано, что последовательность действительно лежит в конусе ($\{\mathbf{u}^{(k)}\}_{k=1}^{\infty} \subset \mathcal{K}$) и $\lim_{k\to\infty} \mathbf{u}^{(k)} = (u_1, 0, u_3) = \mathbf{u}$. Доказано $\mathcal{K}_c \subseteq \mathbf{cl}(\mathcal{K})$.

Теперь от противного докажем $\mathcal{K}_c \supseteq \mathbf{cl}(\mathcal{K})$. Пусть $\mathbf{u} \in \mathbf{cl}(\mathcal{K})$, но $\mathbf{u} \notin \mathcal{K}_c = \mathcal{K} \cup \mathcal{B}$. Это возможно только в случае $\mathbf{u} \in \mathbb{R}_{++} \times \{0\} \times \mathbb{R}_+$ (иначе точка точно будет лежать в $\mathcal{K}$ или в $\mathcal{B}$). Рассмотрим последовательность $\{\mathbf{u}^{(k)}\}_{k=1}^{\infty} \subset \mathcal{K}$ с предельной точкой $\mathbf{u}$, ее элементы должны удовлетворять неравенствам:

$$u_2^{(k)} \exp\left(u_1^{(k)}/u_2^{(k)}\right) \leqslant u_3^{(k)},\, u_2^{(k)} > 0,\, u_3^{(k)} \geqslant 0.$$

Мы предположили, что $\lim_{k\to\infty} u_1^{(k)} = u_1 > 0$, $\lim_{k\to\infty} u_2^{(k)} = 0$ и $\lim_{k\to\infty} u_3^{(k)} = u_3 \geqslant 0$. Но левая часть неравенства выше стремится к $+\infty$, тогда как правая часть стремится к конечному числу. Таким образом, предположение о том, что $\mathbf{u} \notin \mathcal{K}_c$ может являться предельной точкой множества $\mathcal{K}$ неверно, пришли к противоречию. Получаем $\mathcal{K}_c \supseteq \mathbf{cl}(\mathcal{K})$ и, окончательно, $\mathcal{K}_c = \mathbf{cl}(\mathcal{K})$, что и требовалось.

(b) То, что $\mathcal{K}$ — конус, очевидно из подстановки $\alpha(x, y, z)$, $\alpha \geqslant 0$ (сокращаются α в дроби в экспоненте, слева и справа в неравенстве). Для доказательства выпуклости конуса перепишем неравенство $ye^{x/y} \leqslant z$ в виде:

$$f(x) = ye^{x/y} - z \leqslant 0$$

Выпуклость $f(x)$ докажем при помощи критерия второго порядка. Матрица Гессе функции равна:

$$\mathbf{H}_f(x, y, z) = \begin{pmatrix} \frac{e^{x/y}}{y} & -\frac{xe^{x/y}}{y^2} & 0 \\ -\frac{xe^{x/y}}{y^2} & \frac{x^2e^{x/y}}{y^3} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Учитывая ограничения $y > 0, z \geqslant 0$, можно легко видеть, что все главные миноры матрицы неотрицательны (их 7, бóльшая часть равна 0, остальные положительные). Значит, на области определения $\mathbf{H}_f(x, y, z) \succeq 0$, из чего следует выпуклость f.

Таким образом, множество $\mathcal{K}$ является выпуклым как пересечение линии уровня выпуклой функции f (являющееся выпуклым множеством) с двумя выпуклыми множествами (неравенства на y и z, являющиеся полупространствами).

(c) Теперь перейдем к самой трудной части: нахождению сопряженного конуса $\mathcal{K}^*$. Доказательство довольно объемное, оказалось легче и быстрее написать его от руки, оно приведено в приложении в конце документа. Здесь укажем ответ:

$$(\mathbf{cl}(\mathcal{K}))^* = \{(a, b, c) \in (-\mathbb{R}_{++}) \times \mathbb{R} \times \mathbb{R}_+ \,|\, e^{b/a} \leqslant -ec/a\} \cup (\{0\} \times \mathbb{R}_+ \times \mathbb{R}_+)$$

2. Покажем, что множество $\mathbf{COP}_n = \{\mathbf{X} \in \mathbb{R}^{n\times n} \,|\, \mathbf{X} = \mathbf{X}^\mathsf{T}, \mathbf{y}^\mathsf{T}\mathbf{X}\mathbf{y} \geqslant 0, \forall \mathbf{y} \geqslant 0\}$ является выпуклым замкнутым конусом и найдем его двойственный.

Множество $\mathbf{COP}_n$ является выпуклым конусом, так как $\forall \alpha, \beta \geqslant 0$, $\forall \mathbf{A}, \mathbf{B} \in \mathbf{COP}_n, \forall \mathbf{y} \geqslant 0$ выполнено $\mathbf{y}^\mathsf{T}(\alpha\mathbf{A} + \beta\mathbf{B})\mathbf{y} = \alpha\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} + \beta\mathbf{y}^\mathsf{T}\mathbf{B}\mathbf{y} \geqslant 0$. Кроме того, данный конус не содержит прямых, так как если $\mathbf{X}, -\mathbf{X} \in \mathbf{COP}_n$, это означает, что $\mathbf{y}^\mathsf{T}\mathbf{X}\mathbf{y} = 0, \forall \mathbf{y} \geqslant 0$, а это влечет $\mathbf{X} = \mathbf{0}$.

Замкнутость следует из нестрогих неравенств нулю в определении множества и непрерывности линейной функции относительно элементов $x_{i,j}$ матриц $\mathbf{X} \in \mathbf{COP}_n$, а именно: $\mathbf{y}^\mathsf{T}\mathbf{X}\mathbf{y} = \sum_{i,j} y_i y_j x_{i,j} \geqslant 0$.

Для нахождения двойственного конуса $\mathbf{COP}_n^*$ посмотрим на множество под другим углом. $\mathbf{COP}_n$ — выпуклый замкнутый конус, так как данное множество является пересечением замкнутых однородных полупространств, определяемых неравенствами: $\mathbf{y}^\mathsf{T}\mathbf{X}\mathbf{y} = \mathrm{tr}(\mathbf{y}\mathbf{y}^\mathsf{T}\mathbf{X}) = \langle \mathbf{y}\mathbf{y}^\mathsf{T}, \mathbf{X}\rangle \geqslant 0$, то есть $\mathbf{COP}_n = \{\mathbf{X} \in \mathbf{S}^n \mid \langle \mathbf{y}\mathbf{y}^\mathsf{T}, \mathbf{X}\rangle \geqslant 0, \forall \mathbf{y} \geqslant 0\}$.

По определению, двойственным конусом $\mathbf{COP}_n$ является $\mathbf{COP}_n^* = \{\mathbf{Y} \in \mathbb{R}^{n\times n} \mid \langle \mathbf{Y}, \mathbf{X}\rangle = \mathbf{tr}(\mathbf{Y}^\mathsf{T}\mathbf{X}) \geqslant 0, \forall \mathbf{X} \in \mathbf{COP}_n\}$, иными словами, это *множество нормальных векторов всех однородных полупространств, содержащих* $\mathbf{COP}_n$, *и нулевой вектор.* Заметим, что это не только полупространства, участвующие в определении множества ко-положительных матриц. Кроме того, известно, что любой двойственный конус является выпуклым и замкнутым (следует непосредственно из определения).

Таким образом, двойственным конусом является множество:

$$\begin{aligned}\mathbf{COP}_n^* =& \{\mathbf{Y} \in \mathbb{R}^{n\times n} \mid \langle \mathbf{Y}, \mathbf{X}\rangle = \mathbf{tr}(\mathbf{Y}^\mathsf{T}\mathbf{X}) \geqslant 0, \forall \mathbf{X} \in \mathbf{COP}_n\} = \\ & \mathbf{conv}(\{\mathbf{y}\mathbf{y}^T \in \mathbb{R}^{n\times n} \mid \langle \mathbf{y}\mathbf{y}^T, \mathbf{X}\rangle \geqslant 0, \forall \mathbf{X} \in \mathbf{COP}_n\}) = \\ & \mathbf{conv}(\{\mathbf{y}\mathbf{y}^T \in \mathbb{R}^{n\times n} \mid \mathbf{y} \geqslant 0\})\end{aligned}$$

3. Покажем, что множество $\mathcal{C} = \{\mathbf{x} \in \mathbb{R}^n \mid x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n \geqslant 0\}$ — выпуклый замкнутый конус и найдем его двойственный.

   Пусть $\mathbf{a}, \mathbf{b} \in \mathcal{C}$; $\alpha, \beta \geqslant 0$. Тогда $\alpha\mathbf{a} + \beta\mathbf{b} \in \mathcal{C}$, ведь $\alpha a_1 \geqslant \alpha a_2 \geqslant \ldots \geqslant \alpha a_n \geqslant 0$ и $\beta b_1 \geqslant \beta b_2 \geqslant \ldots \geqslant \beta b_n \geqslant 0$, а значит $\alpha a_1 + \beta b_1 \geqslant \alpha a_2 + \beta b_2 \geqslant \ldots \geqslant \alpha a_n + \beta b_n \geqslant 0$. Таким образом, $\mathcal{C}$ является выпуклым конусом. Его замкнутость следует из того факта, что данный конус определен как пересечение n однородных замкнутых полупространств (n нестрогих неравенств на координаты).

   Найдем сопряженный конус $\mathcal{C}^* = \{\mathbf{y} \in \mathbb{R}^n \mid \mathbf{y}^\mathsf{T}\mathbf{x} \geqslant 0, \forall \mathbf{x} \in \mathcal{C}\}$. Заметим, что скалярное произведение можно переписать в виде:

$$\mathbf{y}^\mathsf{T}\mathbf{x} = \sum_{i=1}^{n} x_i y_i = \sum_{i=1}^{n-1}\left[(x_i - x_{i+1})\sum_{j=1}^{i} y_j\right] + x_n \sum_{j=1}^{n} y_j.$$

   Таким образом, можно видеть, что скалярное произведение неотрицательно на , когда суммы $\sum_{j=1}^{m} y_j$ неотрицательны для $m = \overline{1, n}$. Получаем, что сопряженный конус имеет вид:

$$\mathcal{C}^* = \left\{\mathbf{y} \in \mathbb{R}^n \mid \sum_{j=1}^{m} y_j \geqslant 0, m = \overline{1, n}\right\}.$$

3 Выпуклые функции

1. Проверим на выпуклость/вогнутость следующие функции:

   (a) $f(\mathbf{x}) = \prod_{i=1}^{n}(1 - e^{-x_i})^{\lambda_i}$, $\mathrm{dom}(f) = \{\mathbf{x} \in \mathbb{R}^n_{++} \mid \sum_{i=1}^{n} \lambda_i e^{-x_i} \leqslant 1\}$, $\lambda_i > 0, i = \overline{1,n}$.
   Воспользуемся критерием второго порядка: найдем $\mathbf{H}_f(\mathbf{x})$. Докажем, что $\mathbf{H}_f(\mathbf{x}) \succeq 0$. Для этого надо проверить, что все главные миноры Δ_l, $l = \overline{1,n}$ положительны. Элементы матрицы Гессе равны:

$$h_{dd}(\mathbf{x}) = \frac{\partial^2 f(\mathbf{x})}{\partial x_d^2} = \lambda_d e^{-x_d(\lambda_d+1)}(\lambda_d e^{-x_d} - 1)(e^{x_d} - 1)^{\lambda_d - 2} \prod_{i \neq d}(1 - e^{-x_i})^{\lambda_i}$$
$$h_{km}(\mathbf{x}) = \frac{\partial^2 f(\mathbf{x})}{\partial x_k \partial x_m} = \lambda_k \lambda_m e^{-\lambda_k x_k - \lambda_m x_m}(e^{x_k} - 1)^{\lambda_k - 1}(e^{x_m} - 1)^{\lambda_m - 1} \prod_{i \neq k,m}(1 - e^{-x_i})^{\lambda_i},\ k \neq m$$

   Введем следующие обозначения: $a_i = \lambda_i e^{-x_i} > 0$, $b_i = 1 - e^{-x_i} > 0$. Заметим, что по условию $\sum_{i=1}^{n} a_i \leqslant 1$. Для компактности записи всюду опускается обозначение зависимости от x_i. Заметим, что элементы матрицы Гессе в этих обозначениях имеют вид:

$$h_{ii} = \left(\frac{a_i^2}{b_i^2} - \frac{a_i}{b_i^2}\right) f; \quad h_{ij} = \frac{a_i a_j}{b_i b_j} f,\ i \neq j.$$

   Из матрицы $\mathbf{H}_f$ можно вынести f, при этом заметим, что функция положительна на области определения, этот множитель не влияет на определенность матрицы. Рассмотрим матрицу $\tilde{\mathbf{H}} = f\mathbf{H}_f$ и ее главные миноры $\det(\tilde{\mathbf{H}}_I) = \tilde{\Delta}_I = f^k \Delta_I$, $I \subseteq \{1, .. n\}, |I| = k$, где I — набор индексов размера $k = \overline{1,n}$ строк/столбцов матрицы $\tilde{\mathbf{H}}$. Найдем $\tilde{\Delta}_I$, разложив $\tilde{\mathbf{H}}_I$ — симметричные подматрицы $\tilde{\mathbf{H}} \in \mathbb{R}^{k \times k}$ в сумму диагональных $\mathbf{D}_I$ и матриц $\mathbf{v}_I \mathbf{v}_I^\top$ ранга 1:

$$\tilde{\mathbf{H}}_I = \mathbf{D}_I + \mathbf{v}_I \mathbf{v}_I^\top, \text{ где } \mathbf{D}_I = -\mathrm{diag}((a_i/b_i^2)_{i \in I}),\ \mathbf{v}_I = \left((a_i/b_i)_{i \in I}\right).$$

   Теперь воспользуемся леммой об определителе и упростим выражение:

$$\tilde{\Delta}_I = \det(\mathbf{D}_I + \mathbf{v}_I \mathbf{v}_I^\top) = \det(\mathbf{D}_I)(1 + \mathbf{v}_I^\top \mathbf{D}_I^{-1} \mathbf{v}_I)$$
$$= (-1)^k \prod_{i \in I} \frac{a_i}{b_i^2} \left(1 + (-1)^k \sum_{i \in I} a_i\right),\ I \subseteq \{1, .. n\}, |I| = k.$$

   Из формулы выше можно видеть, что $\tilde{\Delta}_I \geqslant 0$ для главных миноров подматриц четного размера и $\tilde{\Delta}_I \leqslant 0$ — для нечетного размера. А это значит, что главные миноры матрицы Гессе $\tilde{\Delta}_I$ имеют те же знаки (так как связывающий их коэффициент f^k положителен). Таким образом, матрица Гессе $\mathbf{H}_f(\mathbf{x})$ рассматриваемой функции $f(\mathbf{x})$ является отрицательно полуопределенной. Значит, функция $f(\mathbf{x})$ вогнутая.

   (b) $f(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 + \lambda\|\mathbf{x}\|_\infty$, $\lambda > 0$.
   Докажем, что любая векторная норма в $\mathbb{R}^n$ выпукла. Действительно, по определению нормы (используются однородность и неравенство треугольника), для $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, $\forall \theta \in [0; 1]$:

$$\|\theta\mathbf{x} + (1 - \theta)\mathbf{y}\| \leqslant \|\theta\mathbf{x}\| + \|(1 - \theta)\mathbf{y}\| = \theta\|\mathbf{x}\| + (1 - \theta)\|\mathbf{y}\|.$$

Известно, что композиция выпуклой функции $g(\mathbf{x})$ и аффинной функции — выпуклая (на соответствующих выпуклых множествах определения). Для доказательства этого факта снова воспользуемся определением $\forall \mathbf{x}, \mathbf{y}$, $\forall \theta \in [0; 1]$:

$$g(\mathbf{A}(\theta\mathbf{x} + (1-\theta)\mathbf{y}) + \mathbf{b}) = g(\theta(\mathbf{A}\mathbf{x} - \mathbf{b}) + (1-\theta)(\mathbf{A}\mathbf{y} - \mathbf{b})) \leqslant \theta g(\mathbf{A}\mathbf{x} - \mathbf{b}) + (1-\theta) g(\mathbf{A}\mathbf{y} - \mathbf{b}).$$

Функция $f(\mathbf{x})$ является положительной линейной комбинацией двух выпуклых функций: $\|\mathbf{x}\|_\infty$ (как векторной нормы) и $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2$ (как композиции выпуклой и аффинной функции). Таким образом, $f(\mathbf{x})$ выпукла.

(c) $g(x) = \frac{1}{x}\int_0^x f(t)dt$, $x > 0$, где $f : \mathbb{R} \to \mathbb{R}$ — выпуклая дифференцируемая функция.
Заметим, что из дифференцируемости $f(x)$ следует дифференцируемость дважды $g(x)$ как интеграла. Воспользуемся дифференциальным критерием выпуклости второго порядка:

$$g''(x) = \frac{f'(x)}{x} - \frac{2f(x)}{x^2} + \frac{2}{x^3}\int\limits_0^x f(t)dt = \frac{2}{x^3}\int\limits_0^x \left[f(t) - (f(x) + f'(x)(t-x))\right] dt.$$

Так как $f(t)$ выпуклая, что по критерию первого порядка для всех x, t на области определения выполнено: $f(t) \geqslant f(x) + f'(x)(t-x)$. Получаем $g''(x) \geqslant 0$, это доказывает выпуклость $g(x)$.

(d) $f(\mathbf{X}) = \sum_{i=1}^k \lambda_i(\mathbf{X})$, где $\mathbf{X} \in \mathbf{S}^n$ и $\lambda_1(\mathbf{X}) \geqslant ... \geqslant \lambda_n(\mathbf{X})$ — собственные значения $\mathbf{X}$.
Для доказательства воспользуемся вариационным определением суммы k максимальных собственных векторов (определение как оптимизационная задача, аналогично доказательству выпуклости для $\lambda_{\max}(\mathbf{X})$):

$$\sum_{i=1}^k \lambda_i(\mathbf{X}) = \sup\Big\{\mathrm{tr}\left(\mathbf{Y}^\intercal\mathbf{X}\mathbf{Y}\right) \mid \mathbf{Y} \in \mathbb{R}^{n\times k},\, \mathbf{Y}^\intercal\mathbf{Y} = \mathbf{I}\Big\}.$$

Заметим, что след можно переписать в виде: $\mathrm{tr}\left(\mathbf{Y}^\intercal\mathbf{X}\mathbf{Y}\right) = \mathrm{tr}\left(\mathbf{X}\mathbf{Y}\mathbf{Y}^\intercal\right)$. Получаем, что оптимизируемая функция линейная (находим поточечный супремум), значит, $f(\mathbf{X})$ выпуклая.

(e) $f(\mathbf{X}) = (\det(\mathbf{X}))^{1/n}$, $\mathbf{X} \in \mathbf{S}^n_{++}$.
Для доказательства воспользуемся критерием выпуклости:

$$\begin{aligned}&[f(\mathbf{X}) \text{ — выпуклая}] \iff \\ &[g(t) = f(\mathbf{A} + t\mathbf{B}),\, \mathrm{dom}(g) = \{t \in \mathbb{R} \mid \mathbf{A} + t\mathbf{B} \in \mathrm{dom}(f)\},\, g \text{ — выпуклая}]\end{aligned}$$

Выберем $\mathbf{A}, \mathbf{B} \in \mathbf{S}^n_{++}$, $t \in \{t \in \mathbb{R} \mid \mathbf{A} + t\mathbf{B} \in \mathbf{S}^n_{++}\}$. Распишем $g(t)$:

$$\begin{aligned}(\det(\mathbf{A} + t\mathbf{B}))^{1/n} &= (\det(\mathbf{A}^{1/2}\mathbf{A}^{1/2} + t\mathbf{B}))^{1/n} = (\det(\mathbf{A}^{1/2})\det(\mathbf{I} + t\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})\det(\mathbf{A}^{1/2}))^{1/n} \\ &= \det(\mathbf{A})^{1/n}\left(\prod_{i=1}^n (1 + t\lambda_i)\right)^{1/n},\end{aligned}$$

здесь λ_i - собственные значения матрицы $\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$. Так как $\det(\mathbf{A})^{1/n} > 0$ и геометрическое среднее $\mathrm{gm}(\mathbf{x})$ — вогнутая функция на $\mathbb{R}^n_{++}$, а в геометрическом среднем стоит линейная функция (их композиция — вогнутая). Поэтому $f(\mathbf{X})$ вогнута.

Дополнение. Докажем вогнутость геометрического среднего $\mathrm{gm}(\mathbf{x})$ на $\mathbb{R}^n_{++}$, по определению. Воспользуемся критерием второго порядка:

$$\frac{\partial^2 \mathrm{gm}(\mathbf{x})}{\partial x_j \partial x_i} = \frac{\mathrm{gm}(\mathbf{x})}{n^2}\left(\frac{1}{x_i x_j} - \delta_{ij}\frac{n}{x_i^2}\right).$$

Проверим, что Гессиан $\mathbf{H}_{\mathrm{gm}}(\mathbf{x})$ отрицательно полуопределен:

$$\mathbf{y}^\mathsf{T}\mathbf{H}_{\mathrm{gm}}(\mathbf{x})\mathbf{y} = \frac{\mathrm{gm}(\mathbf{x})}{n^2}\left[\left(\sum_{i=1}^{n}\frac{y_i}{x_i}\right)^2 - n\sum_{i=1}^{n}\left(\frac{y_i}{x_i}\right)^2\right] \leqslant 0,\, \forall \mathbf{y} \in \mathbb{R}^n_{++},$$

неравенство следует из теоремы Коши-Буняковского для векторов $\mathbf{y} \oslash \mathbf{x}, \mathbf{1} \in \mathbb{R}^n_{++}$. Таким образом, геометрическое среднее вогнуто.

(f) $f(\mathbf{X}) = \mathrm{tr}(\mathbf{X}^{-1})$, $\mathbf{X} \in \mathbf{S}^n_{++}$.

Воспользуемся тем же подходом, что и в пункте (e). Рассмотрим произвольные матрицы $\mathbf{A}, \mathbf{B} \in \mathbf{S}^n_{++}$, $t \in \{t \in \mathbb{R} \,|\, \mathbf{A} + t\mathbf{B} \in \mathbf{S}^n_{++}\}$. Распишем $g(t)$:

$$\begin{aligned}
\mathrm{tr}((\mathbf{A} + t\mathbf{B})^{-1}) &= \mathrm{tr}((\mathbf{A}^{1/2}(\mathbf{I} + t\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})\mathbf{A}^{1/2})^{-1}) \\
&= \mathrm{tr}(\mathbf{A}^{-1}(\mathbf{I} + t\mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})^{-1}) \\
&= \mathrm{tr}(\mathbf{A}^{-1}\mathbf{Q}(\mathbf{I} + t\mathbf{\Lambda})^{-1}\mathbf{Q}^\mathsf{T}) \\
&= \mathrm{tr}(\mathbf{Q}^\mathsf{T}\mathbf{A}^{-1}\mathbf{Q}(\mathbf{I} + t\mathbf{\Lambda})^{-1}),
\end{aligned}$$

здесь использовалось разложение матрицы $\mathbf{C} = \mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2}$ по собственным векторам: $\mathbf{C} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^\mathsf{T}$. Последнее выражение является вычислением следа произведения положительно определенной и диагональной матрицы. Диагональные d_{ii} элементы положительно определенной матрицы $\mathbf{Q}^\mathsf{T}\mathbf{A}^{-1}\mathbf{Q}$ положительны, поэтому формулу можно переписать формулу в виде:

$$f(\mathbf{X}) = \sum_{i=1}^{n}\frac{d_{ii}}{1 + t\lambda_i},\, d_i i, \lambda_i \in \mathbb{R}_{++},$$

то есть в виде положительной линейной комбинации выпуклых функций $1/(1 + t\lambda_i)$ (вторая производная $2\lambda^2/(1 + t\lambda)^3 > 0$). Поэтому $f(\mathbf{X})$ — выпуклая функция.

(g) $f(\mathbf{x}, t) = -\log(t^2 - \mathbf{x}^\mathsf{T}\mathbf{x})$, $\mathrm{dom}(f) = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}_+ \,|\, \|\mathbf{x}\|_2 < t\}$.

Перепишем функцию в следующем виде:

$$f(\mathbf{x},t) = -\log(t^2 - \mathbf{x}^\mathsf{T}\mathbf{x}) = -\log(t) - \log\left(t - \frac{\mathbf{x}^\mathsf{T}\mathbf{x}}{t}\right).$$

Логарифм — вогнутая функция, поэтому первое слагаемое $(-\log(t))$ является выпуклой функцией. Рассмотрим функцию $g(\mathbf{x}, t) = \mathbf{x}^\mathsf{T}\mathbf{x}/t = \sum_{i=0}^{n} x_i^2/t = \sum_{i=0}^{n} g_i(x_i, t)$. Каждое слагаемое $g_i(x_i, t)$ является выпуклой функцией, что можно доказать через положительную полуопределенность Гессиана:

$$\mathbf{H}_{g_i}(x_i,t) = \begin{pmatrix} \frac{2}{t} & -\frac{2x_i}{t^2} \\ -\frac{2x_i}{t^2} & \frac{2x_i^2}{t^3} \end{pmatrix};\ \Delta_1 = \frac{2}{t} > 0,\, \Delta_2 = 0,\, t > 0.$$

Таким образом, $g(\mathbf{x}, t)$ является выпуклой как положительная линейная комбинация выпуклых функций. Тогда $(-\mathbf{x}^\mathsf{T}\mathbf{x}/t)$ — вогнутая, как и $t - \mathbf{x}^\mathsf{T}\mathbf{x}/t$ (сумма линейной и вогнутой функции вогнута). Далее, второе слагаемое в $f(\mathbf{x}, t)$, $(-\log\left(t - \mathbf{x}^\mathsf{T}\mathbf{x}/t\right))$ — выпукло как композиция выпуклой невозрастающей функции $(-\log(y))$, $y > 0$ и вогнутой функции. Получаем, что $f(\mathbf{x}, t)$ — выпукла как положительная линейная комбинация выпуклых функций (или, эквивалентно, отрицательной линейной комбинации вогнутых функций).

Дополнение. Докажем следующее утверждение: пусть $g(x)$ — вогнутая функция, $h(x)$ — выпуклая и невозрастающая. Тогда $f(x) = h(g(x))$ — выпуклая функция.

Выведем это по определению: $x, y \in \mathrm{dom}(g)$; $g(x), g(y) \in \mathrm{dom}(h)$; $\forall\theta \in [0; 1]$:

$$h(g(\theta x + (1-\theta)y)) \leqslant h(\theta g(x) + (1-\theta)g(y)) \leqslant \theta h(g(x)) + (1-\theta)h(g(y)).$$

Это доказывает выпуклость $f(x)$. Аналогично доказываются следующие утверждения (в список включено доказанное):

* $g(x)$ — выпуклая, $h(x)$ — выпуклая и неубывающая. Тогда $h(g(x))$ — выпуклая.
* $g(x)$ — вогнутая, $h(x)$ — выпуклая и невозрастающая. Тогда $h(g(x))$ — выпуклая.
* $g(x)$ — вогнутая, $h(x)$ — вогнутая и неубывающая. Тогда $h(g(x))$ — вогнутая.
* $g(x)$ — выпуклая, $h(x)$ — вогнутая и невозрастающая. Тогда $h(g(x))$ — вогнутая.

2. Пусть задан ориентированный взвешенный граф $G = (V,E)$. Рассмотрим $p_{ij}(\mathbf{c})$ — функцию кратчайшего расстояния между некоторой парой вершин $v_i,v_j \in V$, зависящую от вектора весов $\mathbf{c} \in \mathbb{R}^{|E|}$ ребер графа. Докажем, что данная функция вогнута.

   Функция кратчайшего расстояния между парой вершин $v_i,v_j \in V$ является минимумом по всевозможным путям $\Pi_{ij}(\mathbf{c}) \subseteq E$, соединяющим эти вершины:

$$p_{ij}(\mathbf{c}) = \min_{\Pi_{ij}(\mathbf{c})} \sum_{e \in \Pi_{ij}(\mathbf{c})} c_e.$$

   Заметим, что кратчайший путь в графе (упорядоченный набор ребер, соединяющих данные вершины) очевидным образом зависит от вектора весов ребер. Докажем вогнутость p_{ij} по определению. Рассмотрим произвольные вектора весов $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{R}^{|E|}$ и произвольную $\theta \in [0; 1]$, их выпуклая комбинация равна $\mathbf{b}_\theta = \theta\mathbf{b}_1 + (1-\theta)\mathbf{b}_2$. Требуется доказать:

$$p_{ij}(\mathbf{b}_\theta) = \min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e \in \Pi_{ij}(\mathbf{b}_\theta)} b_{\theta e} \geqslant \theta p_{ij}(\mathbf{b}_1) + (1-\theta)p_{ij}(\mathbf{b}_2).$$

   По линейности выпишем левую часть неравенства, $p_{ij}(\mathbf{b}_\theta)$, как сумму двух слагаемых:

$$p_{ij}(\mathbf{b}_\theta) = \min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e \in \Pi_{ij}(\mathbf{b}_\theta)} \theta b_{1e} + (1-\theta)b_{2e} = \theta \left[\min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e \in \Pi_{ij}(\mathbf{b}_\theta)} b_{1e}\right] + (1-\theta)\left[\min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e \in \Pi_{ij}(\mathbf{b}_\theta)} b_{2e}\right].$$

   Теперь по определению распишем правую часть, выпуклую комбинацию кратчайших расстояний в графе для векторов весов $\mathbf{b}_1$ и $\mathbf{b}_2$:

$$\theta p_{ij}(\mathbf{b}_1) + (1-\theta)p_{ij}(\mathbf{b}_2) = \theta \left[\min_{\Pi_{ij}(\mathbf{b}_1)} \sum_{e \in \Pi_{ij}(\mathbf{b}_1)} b_{1e}\right] + (1-\theta)\left[\min_{\Pi_{ij}(\mathbf{b}_2)} \sum_{e \in \Pi_{ij}(\mathbf{b}_2)} b_{2e}\right].$$

Заметим теперь, что $\theta p_{ij}(\mathbf{b}_1)$ и $\theta p_{ij}(\mathbf{b}_2)$ соответственно не больше, чем первое и второе слагаемые в записи $\theta p_{ij}(\mathbf{b}_\theta)$, так как в первом случае речь идет о кратчайшем расстоянии в графе между данными вершинами для векторов $\mathbf{b}_1$ и $\mathbf{b}_2$, но для пути $\Pi_{ij}(\mathbf{b}_\theta)$, тогда как во втором случае кратчайший путь выбирается отдельно для каждого вектора весов (соответственно $\Pi_{ij}(\mathbf{b}_1)$, $\Pi_{ij}(\mathbf{b}_2)$). Иными словами:

$$\min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e\in\Pi_{ij}(\mathbf{b}_\theta)} b_{1e} \geqslant \min_{\Pi_{ij}(\mathbf{b}_1)} \sum_{e\in\Pi_{ij}(\mathbf{b}_1)} b_{1e}; \quad \min_{\Pi_{ij}(\mathbf{b}_\theta)} \sum_{e\in\Pi_{ij}(\mathbf{b}_\theta)} b_{2e} \geqslant \min_{\Pi_{ij}(\mathbf{b}_2)} \sum_{e\in\Pi_{ij}(\mathbf{b}_2)} b_{2e}.$$

Это доказывает неравенство $p_{ij}(\mathbf{b}_\theta) \geqslant \theta p_{ij}(\mathbf{b}_1) + (1-\theta)p_{ij}(\mathbf{b}_2)$, то есть $p_{ij}(\mathbf{c})$ вогнута.

3. Покажем, что следующая функция выпуклая и убывает (покоординатно) для любого конечного n на области определения, где каждый знаменатель положителен:

$$f_n(\mathbf{x}) = \cfrac{1}{x_1 - \cfrac{1}{x_2 - \cfrac{1}{x_3 - ...}}}$$

Функция имеет рекурсивный вид и может быть записана как композиция следующих функций:

$$f_n(x_1,x_2,...,x_n) = \frac{1}{x_1 - f_{n-1}(x_2,...,x_n)},$$
$$f_1(x_1) = \frac{1}{x_1}, \quad f_2(x_1,x_2) = \frac{1}{x_1 - \frac{1}{x_2}}, \quad f_3(x_1,x_2,x_3) = \frac{1}{x_1 - \frac{1}{x_2 - \frac{1}{x_3}}}, \quad ...$$

Для доказательства убывания функции воспользуемся индукцией по n (числу аргументов) и по m (порядковый номер аргумента в функции). Для $n=1$, $m=1$ (знаменатели положительны, $x_1>0$):

$$f_1(x_1) = \frac{1}{x_1}, \quad f_1'(x_1) = -\frac{1}{x_1^2} < 0, \quad f_2''(x_1) = \frac{2}{x_1^3} > 0,$$

то есть $f_1(x_1)$ выпуклая и убывающая.

Пусть свойства доказаны для $n \leqslant k$, докажем для $n = k+1$. Сначала рассмотрим частную производную по x_1, $\frac{\partial f_{k+1}}{\partial x_1}$:

$$\frac{\partial f_{k+1}(x_1,...,x_{k+1})}{\partial x_1} = -\frac{1}{(x_1 - f_k(x_2,...,x_{k+1}))^2} < 0,$$

то есть f_{k+1} убывает по x_1. Теперь рассмотрим частную производную по x_2:

$$\frac{\partial f_{k+1}(x_1,...,x_{k+1})}{\partial x_2} = \frac{1}{(x_1 - f_k(x_2,...,x_{k+1}))^2} \cdot \frac{\partial f_k(x_2,...,x_{k+1})}{\partial x_2} < 0,$$

она меньше нуля, так как на предыдущих шагах было доказано $\frac{\partial f_k}{\partial x_2} < 0$, а первый множитель больше нуля. Аналогично доказываются знаки частных производных для $3 \leqslant m \leqslant k+1$, ведь $\frac{\partial f_k}{\partial x_m} < 0$ было доказано ранее:

$$\frac{\partial f_{k+1}(x_1,...,x_{k+1})}{\partial x_m} = \frac{1}{(x_1 - f_k(x_2,...,x_{k+1}))^2} \cdot \frac{\partial f_k(x_2,...,x_{k+1})}{\partial x_m} < 0.$$

Таким образом, функция $f_{k+1}(x_1,...,x_{k+1})$ — убывающая по каждому аргументу, шаг индукции доказан. Значит, функция $f_n(x_1,...,x_n)$ убывает по каждому аргументу для любого n.

Теперь докажем выпуклость функции f_n. Воспользуемся индукцией по n, числу аргументов. База для $n = 1$ уже была доказана выше (для $f_1(x_1) = 1/x_1$). Пусть функция выпукла для $n \leqslant k$. Докажем выпуклость для $n = k + 1$:

$$f_{k+1}(x_1,x_2,...,x_n) = \frac{1}{x_1 - f_k(x_2,...,x_{k+1})}.$$

Ранее было доказано, что f_k убывает по каждому аргументу, по предположению индукции она также выпукла. Поэтому $g(x_1,...,x_{k+1}) = (x_1 - f_k(x_2,...,x_{k+1}))$ — вогнутая как сумма линейной и вогнутой (выпуклой с отрицательным коэффициентом). Ранее было показано, что функция $h(x) = 1/x$ — выпуклая и убывающая для $x > 0$. Тогда из свойств композиции (см. 3.1(g)) следует, что $f_{k+1}(x_1,x_2,...,x_n) = h(g(x_1,...,x_{k+1}))$ — выпуклая. Шаг индукции доказан. Таким образом, $f_n(x_1,...,x_n)$ — выпуклая функция для любого n.

4. Покажем выпуклость функции:

$$h(\mathbf{x}) = \frac{f^2(\mathbf{x})}{g(\mathbf{x})},$$

где $f(\mathbf{x})$ выпуклая и неотрицательная, а $g(\mathbf{x})$ вогнутая и положительная.

В задаче не делаются предположения о дифференцируемости функций, поэтому докажем выпуклость из свойств композиции.

Рассмотрим функцию $p(u,v) = u^2/v$, $\mathrm{dom}(p) = \{(u,v) \in \mathbb{R} \times \mathbb{R}_{++}\}$. Эта функция выпуклая, доказывается с использованием критерия второго порядка (и это было доказано в 3.1(g)). Кроме того, данная функция является возрастающей по u для $u \geqslant 0$ и убывающей по v для $y > 0$.

Теперь рассмотрим исходную функцию как композицию: $h(\mathbf{x}) = p(f(\mathbf{x}), g(\mathbf{x}))$. Функция $p(u,v)$ — выпуклая, возрастающая по u для $u \in \mathrm{dom}(f)$ и убывающая по v для $v \in \mathrm{dom}(g)$. Воспользовавшись правилами композиции (доказательство аналогично 3.1(g), по определению, но по двум аргументам внешней функции), получаем, что $h(\mathbf{x})$ — выпуклая функция, что и требовалось.

4 Приложение (2.1(c))

∡ конус $K_{d_0} = \{(a,b,c) \in (-\mathbb{R}_{++}) \times \mathbb{R} \times \mathbb{R}_+ \mid \exp\left(\frac{b}{a}\right) \le -\frac{e \cdot c}{a}\}$

∡ $K_d = K_{d_0} \cup \mathcal{A}$, где $\mathcal{A} = \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+$

- Тот факт, что $cl(K_{d_0}) = K_d$ доказывается абсолютно аналогично п. (a).

- Докажем, что $(cl(K))^* = K_d$. Для этого докажем $K^* \supseteq K_d$ и $K^* \subseteq K_d$.

① Начнём с $(cl(K))^* \supseteq K_d$.] $\overset{(a,b,c)}{v} \in K_d = K_{d_0} \cup \{0\} \times \mathbb{R}_+ \times \mathbb{R}_+$. Если $v \in \mathcal{A}$, то:

$$v^T u = \overset{=0}{a}x + \overset{\ge 0}{b}\overset{\ge 0}{y} + \overset{\ge 0}{c}\overset{\ge 0}{z} \ge 0, \quad \forall u = (x,y,z) \in cl(K).$$

Теперь рассмотрим $v \in K_{d_0}$, $u \in cl(K)$. Оценим снизу $v^T u$:

$$v^T u = ax + by + c\underline{z} \ge ax + by + \underline{c\, y \cdot \exp\left(\frac{x}{y}\right)} \ge ax - a y\, \underline{\exp\left(\frac{a-b}{a}\right)} \exp\left(\frac{x}{y}\right) + by$$

Упростим выражение:

$$v^T u \ge ax + by - ay \exp\left(-\frac{a-b}{a} + \frac{x}{y}\right) = ax + by - ay \exp\left(-\frac{ay - by - ax}{ay}\right) \ge$$

$$\ge \left[\begin{matrix} e^t \ge 1+t \\ (\forall t) \end{matrix}\right] \ge ax + by - ay \cdot \left(-\frac{ay - by - ax}{ay} + 1\right) = 0.$$

Таким образом, $v \in (cl(K))^* \Rightarrow (cl(K))^* \supseteq K_d$.

② Теперь докажем, что $(cl(K))^* \subseteq K_d$.

Докажем от противного, пусть $v = (a,b,c) \in (cl(K))^*$, но $v \notin K_d$.

K_d – объединение 2 выпуклых конусов K_{d_0} и $\mathcal{A}$. Значит, $[v \notin K_d] \Leftrightarrow [v \notin K_{d_0}] \& [v \notin \mathcal{A}]$.

То есть условие $v \notin K_d$ эквивалентно следующим условиям:

$$[v \notin K_d] \Leftrightarrow \left[\left((a \ge 0) \vee \left(\exp\left(\frac{b}{a}\right) > -\frac{ec}{a}\right) \vee (c < 0)\right) \& \left((a \ne 0) \vee (b < 0) \vee (c < 0)\right)\right]$$

Всего есть 9 вариантов (слагаемых) (можно раскрыть скобки по дистрибутивности), проверим каждый из них:

1–5) Те варианты, в которых $(c < 0)$ — 5 шт.

∡ $u = (0,0,1) \in cl(K)$. Тогда $v^T u = c < 0 \Rightarrow v \notin (cl(K))^*$

6) $(a \ge 0) \& (b < 0)$ — 1 шт.

> Если $c < 0$: ∡ $u = (0,0,1) \in cl(K)$; $v^T u = c < 0 \Rightarrow v \notin (cl(K))^*$

> Если $c = 0$: ∡ $u = (0,1,1) \in cl(K)$; $v^T u = b + c = b < 0 \Rightarrow v \notin (cl(K))^*$

> Если $c > 0$; $a > 0$: ∡ $u = (-1;0,0) \in cl(K)$; $v^T u = -a < 0 \Rightarrow v \notin (cl(K))^*$

> Если $c>0;\ a=0$: $\nexists\, u = \left(\ln\left(\frac{-b}{2c}\right),\ 1;\ \frac{-b}{2c}\right) \in cl(K)$ (легко проверить)

$$v^T u = b + c\cdot\frac{-b}{2c} = \frac{1}{2}b < 0 \Rightarrow v \notin (cl(K))^*$$

7) $(a\geq 0)\,\&\,(a\neq 0) = (a>0)$ — 1шт.

$\nexists\, u = (-1,0,0) \in cl(K)$; $v^T u = -a < 0 \Rightarrow v \notin (cl(K))^*$

$\left(\exp\left(\frac{b}{a}\right) > -\frac{ec}{a}\right)\,\&\,(a\neq 0) \vee \cancel{(b<0)}$

$=$

8-9) $\underbrace{\left(\exp\left(\frac{b}{a}\right) > -\frac{ec}{a}\right)\,\&\,(a\neq 0)} \vee \left(\exp\left(\frac{b}{a}\right) > -\frac{ec}{a}\right)\,\&\,(b<0)$ — 2шт.

> Если $a>0$: $\nexists\, u = (-1,0,0) \in cl(K)$; $v^T u = -a < 0 \Rightarrow v \notin (cl(K))^*$

> Если $a<0;\ c<0$: $\nexists\, u = (0,0,1) \in cl(K)$; $v^T u = c < 0 \Rightarrow v \notin (cl(K))^*$

> Если $a<0;\ c=0$: $\nexists\, u = \left(1,\ 1/\gamma;\ e^{\gamma}/\gamma\right)$, $\gamma>0$. Очевидно, что $u \in cl(K)$

$$v^T u = \overset{<0}{a} + b/\gamma < 0 \text{ для } \gamma > -b/a. \Rightarrow v \notin (cl(K))^*$$

> Если $a<0,\ c>0$: $\nexists\, u = \left(\frac{a-b}{ac}\cdot\exp\left(-\frac{a-b}{a}\right);\ \exp\left(-\frac{a-b}{a}\right)\cdot\frac{1}{c};\ \frac{1}{c}\right) \in cl(K)$

$$v^T u = ax + by + cz = \frac{a-b}{c}\exp\left(-\frac{a-b}{a}\right) + \frac{b}{c}\exp\left(-\frac{a-b}{a}\right) + 1$$

$$\Leftrightarrow \frac{a}{c}\exp\left(-\frac{a-b}{a}\right) + 1 < 0 \Leftrightarrow \underbrace{\exp\left(-\frac{a-b}{a}\right) > -\frac{c}{a}}_{\Updownarrow}$$

$$\exp\left(\frac{b}{a}\right) > -\frac{ec}{a}$$

Значит, $v \notin (cl(K))^*$

Таким образом, $(cl(K))^* \subseteq K_d$. $(cl(K))^* = K_d$ доказано.