Методы оптимизации, ВМК, осень 2018

Дополнительный материал 4: Самосогласованные функции и метод Ньютона

1 Пролог: область квадратичной сходимости метода Ньютона

Ранее было установлено крайне важное качественное свойство метода Ньютона: при запуске из «достаточно близкой» к решению задачи начальной точки он обладает *квадратичной сходимостью*. Благодаря этому, во всех возникающих на практике ситуациях для достижения произвольной заданной точности методу Ньютона достаточно выполнить лишь 2–5 итераций. Возникают следующие естественные вопросы: а что же это за множество «достаточно близких» к решению начальных точек, для которых метод обладает квадратичной сходимостью, и насколько это множество большое? Чтобы ответить на эти и другие подобные вопросы, необходимо иметь четкое теоретическое описание *области квадратичной сходимости* метода Ньютона.

Для получения описания этой области наиболее естественным кажется обратиться к классической теореме о квадратичной сходимости метода Ньютона, которая была рассмотрена и доказана ранее:

Теорема 1.1 (Квадратичная сходимость метода Ньютона)**.** *Пусть V — конечномерное евклидово пространство (с соответствующей евклидовой нормой $\|\|$), Q — непустое открытое множество в V, и пусть $f : Q \to \mathbb{R}$ — трижды непрерывно дифференцируемая функция с точкой минимума $x^* \in Q$. Пусть при этом*

$$|D^3 f(x)[v, v, v]| \leq M\|v\|^3 \tag{1.1}$$

для некоторых $M, \delta > 0$ и всех $x \in Q$, таких, что $\|x - x^*\| \leq \delta$, и всех $v \in V$; и пусть

$$D^2 f(x^*)[v, v] \geq \mu\|v\|^2.$$

для некоторого $\mu > 0$ и всех $v \in V$. Тогда для любой начальной точки x_0 из множества

$$U := \left\{x \in Q : \|x - x^*\| \leq \min\left\{\frac{\mu}{2M}, \delta\right\}\right\} \tag{1.2}$$

последовательность $(x_k)_{k=0}^{\infty}$, заданная рекуррентно методом Ньютона

$$x_{k+1} := x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k), \tag{1.3}$$

определена корректно; при этом $x_k \in U$ и

$$\|x_k - x^*\| \leq \frac{\mu}{M}\left(\frac{1}{2}\right)^{2^k}$$

для всех $k \geq 0$.

Согласно этой теореме, область квадратичной сходимости U метода Ньютона, описываемая формулой (1.2), представляет собой шар относительно метрики, индуцированной нормой $\|\|$; центр этого шара есть решение x^*, а радиус определяется параметрами M, μ и δ.

Несмотря всю свою «естественность», оказывается, что теорема 1.1 имеет скрытый недостаток, который был обнаружен и подробно исследован в начале 90-х годов прошлого века Нестеровым и Немировским: описание области квадратичной сходимости дается полностью в терминах некоторой заранее выбранной евклидовой нормы $\|\|$, в то время как сам метод Ньютона ничего общего с нормой $\|\|$ не имеет! Другими словами, если вместо нормы $\|\|$ рассмотреть другую евклидову норму $\|\|'$, то при этом изменятся как величины M, μ и δ, так и само описание области квадратичной сходимости (1.2); однако последовательность $(x_k)_{k=0}^{\infty}$, построенная методом Ньютона, при этом останется прежней, поскольку итерация (1.3) может быть переписана в эквивалентной форме минимизации квадратичной модели:

$$x_{k+1} = \operatorname*{argmin}_{x \in V} \left\{ f(x_k) + \langle \nabla f(x_k), x - x_k \rangle + \frac{1}{2} \langle \nabla^2 f(x_k)(x - x_k), x - x_k \rangle \right\};$$

отсюда видно, что ньютоновский шаг определяется только первыми и вторыми производными $v \mapsto \langle \nabla f(x_k), v\rangle$ и $v \mapsto \langle \nabla^2 f(x_k)v, v\rangle$ функции f по направлению, которые никак от нормы $\|\|$ не зависят. Таким образом, рассмотрение разных евклидовых норм дает в итоге разные описания области квадратичной сходимости, хотя *реальная* область квадратичной сходимости метода Ньютона является некоторым фиксированным множеством, которое определяется исключительно самой функцией f.

Пример 1.2. Рассмотрим в пространстве $\mathbb{R}^2$ открытый прямоугольник

$$Q := \{x \in \mathbb{R}^2 : |x_1| < A;\ |x_2| < 1\},$$

где $A \geq 1$. Пусть $f : Q \to \mathbb{R}$ — функция

$$f(x) := -\ln(A^2 - x_1^2) - \ln(1 - x_2^2).$$

Точкой минимума этой функции является 0.

(a) Если применять теорему 1.1 для стандартной евклидовой нормы

$$\|v\| := (v_1^2 + v_2^2)^{1/2},$$

то для любого выбора δ область квадратичной сходимости U является подмножеством, скажем, круга

$$\left\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq \frac{1}{12A^2}\right\},$$

радиус которого уменьшается с ростом A. Таким образом, с ростом прямоугольника Q, область квадратичной сходимости U из теоремы 1.1 *уменьшается*.

(b) Если применять теорему 1.1 для нестандартной евклидовой нормы

$$\|v\|' := \left(\frac{v_1^2}{A^2} + v_2^2\right)^{1/2},$$

то для, скажем, $\delta = 0.1$ область квадратичной сходимости U представляет собой эллипс

$$U = \left\{x \in \mathbb{R}^2 : \frac{x_1^2}{A^2} + x_2^2 \leq 0.01\right\},$$

который является концентрическим к эллипсу, который вписан в прямоугольник, но только в 10 раз меньшего радиуса. Таким образом, *реальная* область квадратичной сходимости метода Ньютона с ростом прямоугольника Q *растет*, а не уменьшается.

Доказательство. Дифференцируя, находим

$$\begin{aligned}
Df(x)[v] &= \frac{2x_1}{A^2 - x_1^2}v_1 + \frac{2x_2}{1 - x_2^2}v_2, \\
D^2 f(x)[v, v] &= \frac{2(A^2 + x_1^2)}{(A^2 - x_1^2)^2}v_1^2 + \frac{2(1 + x_2^2)}{(1 - x_2^2)^2}v_2^2, \\
D^3 f(x)[v, v, v] &= \frac{4x_1(3A^2 + x_1^2)}{(A^2 - x_1^2)^3}v_1^3 + \frac{4x_2(3 + x_2^2)}{(1 - x_2^2)^3}v_2^3
\end{aligned}$$

для всех $x \in Q$ и всех $v \in \mathbb{R}^2$. В частности, $Df(0) = 0$; в силу выпуклости f, это доказывает, что 0 является точкой минимума.

(a) Предположим, чтобы прийти к противоречию, что для некоторого $\delta > 0$ радиус области квадратичной сходимости U из теоремы 1.1 больше, чем $1/(\sqrt{12}A)$. Тогда $\delta \geq 1/(\sqrt{12}A)$. Полагая $x = (0, \delta)$ и $v = (0, 1)$ получаем

$$D^3 f(x)[v, v, v] = \frac{4\delta(3 + \delta^2)}{(1 - \delta^2)^3} \geq 12\delta \geq \frac{\sqrt{12}}{A},$$

что показывает $M \geq \sqrt{12}/A$. В то же время

$$D^2 f(0)[v, v] = \frac{2}{A^2} v_1^2 + 2v_2^2,$$

что показывает $\mu = 2/A^2$ (достаточно рассмотреть $v = (1, 0)$). Таким образом, $\mu/(2M) \leq 1/(\sqrt{12}A)$. Но тогда радиус области квадратичной сходимости U не может быть больше $1/(\sqrt{12}A)$.

(b) Пусть $v \in \mathbb{R}^2$, и пусть $x \in Q$, причем $\|x\|' \leq \delta$. Тогда $|x_1| \leq \delta A$ и $|x_2| \leq \delta$. Отсюда

$$|D^3 f(x)[v, v, v]| \leq \frac{4(\delta A)(3A^2 + (\delta A)^2)}{(A^2 - (\delta A)^2)^3} v_1^3 + \frac{4\delta(3 + \delta^2)}{(1 - \delta^2)^3} v_2^3 = \frac{4\delta(3 + \delta^2)}{(1 - \delta^2)^3} \left(\frac{v_1^3}{A^3} + v_2^3 \right).$$

Используя стандартное неравенство $a^3 + b^3 \leq (a^2 + b^2)^{3/2}$ для $a, b > 0$ (утверждение 1.3), получаем

$$|D^3 f(x)[v, v, v]| \leq \frac{4\delta(3 + \delta^2)}{(1 - \delta^2)^3} (\|v\|')^3.$$

Полагая $\delta = 0.1$, получаем

$$\frac{4\delta(3 + \delta^2)}{(1 - \delta^2)^3} = 1.24\ldots \leq 2,$$

что дает $M \leq 2$. В то же время,

$$D^2 f(0)[v, v] = \frac{2}{A^2} v_1^2 + 2v_2^2 = 2(\|v\|')^2,$$

что дает $\mu = 2$. Таким образом, $\mu/(2M) \geq 0.5$, и радиус области U есть $\delta = 0.1$. $\square$

Утверждение 1.3. *Пусть* $a_1, \ldots, a_n \geq 0$, *и пусть* $0 < r \leq p$. *Тогда*

$$\left(\sum_{i=1}^n a_i^p \right)^{1/p} \leq \left(\sum_{i=1}^n a_i^r \right)^{1/r}. \tag{1.4}$$

Доказательство. Достаточно доказать (1.4) только для случая $\sum_{i=1}^n a_i^r = 1$; общий случай будет следовать из доказанного при рассмотрении $a_1/t, \ldots, a_n/t$, где $t := (\sum_{i=1}^n a_i^r)^{1/r}$.

Пусть $\sum_{i=1}^n a_i^r = 1$. Тогда $a_i \leq 1$ для всех $1 \leq i \leq n$ (от противного). Поскольку $r \leq p$, то отсюда следует $a_i^p \leq a_i^r$ для всех $1 \leq i \leq n$. Складывая, получаем $\sum_{i=1}^n a_i^p \leq \sum_{i=1}^n a_i^r = 1$. Таким образом, $(\sum_{i=1}^n a_i^p)^{1/p} \leq 1$. $\square$

Пример 1.2 демонстрирует, что неудачный выбор нормы $\|\|$ в теореме 1.1 приводит к совершенно неправильному описанию области квадратичной сходимости метода Ньютона. Чтобы описать реальную область квадратичной сходимости более точно, необходимо рассматривать норму, так или иначе связанную с самой минимизируемой функцией.

Заметив упомянутый недостаток, Нестеров и Немировский предложили анализировать сходимость метода Ньютона с помощью специального семейства евклидовых норм, индуцированных самой минимизируемой функцией. Это привело к определению класса *самосогласованных функций*, для которого удается получить «правильное» описание области квадратичной сходимости метода Ньютона. Кроме того, для самосогласованных функций также удается получить не только локальные, но и *глобальные* оценки скорости сходимости метода Ньютона (чего не удается сделать для многих других стандартных классов функций).

2 Самосогласованные функции: определение и примеры

В основе теории самосогласованных функций лежит следующее важное наблюдение. Пусть V — конечномерное нормированное вещественное векторное пространство, Q — непустое открытое множество в V, и пусть $f : Q \to \mathbb{R}$ — дважды дифференцируемая функция. Тогда для любой фиксированной точки $x \in Q$ вторая производная $D^2 f(x)$, в силу стандартных свойств, задает в пространстве V симметричную билинейную форму $\langle , \rangle_{f,x} : V \times V \to \mathbb{R}$, определенную формулой

$$\langle v, w \rangle_{f,x} := D^2 f(x)[v, w].$$

Если при этом $D^2 f(x)$ является неотрицательной, то симметричная билинейная форма $\langle , \rangle_{f,x}$ становится *положительно полуопределенным скалярным произведением* в V. Таким образом, любая функция f, имеющая всюду неотрицательную вторую производную, индуцирует в пространстве V целое семейство евклидовых структур.

Каждая евклидова структура, в свою очередь, для произвольного $x \in Q$ индуцирует неотрицательную функцию $\| \|_{f,x} : V \to [0, +\infty)$, заданную формулой

$$\|v\|_{f,x} := \langle v, v \rangle_{f,x}^{1/2}.$$

Нетрудно видеть, что $\| \|_{f,x}$ является абсолютно однородной; кроме того, она удовлетворяет неравенству треугольника (упражнение 2.1). Таким образом, $\| \|_{f,x}$ задает в пространстве V евклидову *полунорму.*

Упражнение 2.1. (a) Пусть $(V, \langle , \rangle)$ — вещественное векторное пространство с положительно полуопределенным скалярным произведением, и пусть $v, w \in V$. Докажите *неравенство Коши–Буняковского*

$$|\langle v, w \rangle| \le \langle v, v \rangle^{1/2} \langle w, w \rangle^{1/2}.$$

(*Подсказка*: воспользуйтесь неотрицательностью $\langle v - \alpha w, v - \alpha w \rangle$, подобрав нужным образом α; при этом отдельно рассмотрите случай $\langle w, w \rangle = 0$.)

(b) Используя полученный результат, установите, что функция $\| \| : V \to [0, +\infty)$, определенная формулой $\|v\| := \langle v, v \rangle^{1/2}$, удовлетворяет *неравенству треугольника*

$$\|v + w\| \le \|v\| + \|w\|$$

для всех $v, w \in V$.

Чтобы подчеркнуть зависимость скалярного произведения $\langle , \rangle_{f,x}$ и полунормы $\| \|_{f,x}$ от выбранной точки $x \in Q$, будем называть их *локальными* скалярным произведением и полунормой соответственно. Заметим, что если вторая производная $D^2 f(x)$ является невырожденной, т. е. $D^2 f(x)[v, v] > 0$ для всех ненулевых $v \in V$, то $\langle , \rangle_{f,x}$ становится *положительно определенным* скалярным произведением, а полунорма $\| \|_{f,x}$ становится *нормой.*

Воспользуемся введенными локальными полунормами для определения самосогласованных функций:

Определение 2.2 (Самосогласованные функции)**.** Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое множество в V, и пусть $f : Q \to \mathbb{R}$ — трижды непрерывно дифференцируемая функция с неотрицательной второй производной. Функция f называется *самосогласованной*, если существует $M \ge 0$, такое, что

$$|D^3 f(x)[v, v, v]| \le M \|v\|_{f,x}^3. \tag{2.1}$$

для всех $x \in Q$ и всех $v \in V$. Число M при этом называется *параметром самосогласованности.* Если f является самосогласованной с параметром 2, то f называют *стандартной самосогласованной* функцией.

Замечание 2.3. Неравенство (2.1), грубо говоря, является аналогом неравенства (1.1) и означает, что для любой точки x третья производная $D^3 f(x)$ должна быть ограничена (как симметричное трилинейное преобразование) относительно локальной полунормы $\|\|_{f,x}$, причем одной и той же константой M для всех x. Тем не менее, между этими неравенствами есть одно важное различие: при варьировании точки x локальная полунорма $\|\|_{f,x}$ в правой части неравенства (2.1) *меняется*, в то время как норма $\|\|$ в правой части неравенства (1.1) остается *одинаковой* для всех x.

Замечание 2.4. Несмотря на то, что определение 2.2 можно использовать и для невыпуклых множеств Q, наибольший интерес все-таки представляет случай, когда множество Q выпуклое. Это связано с тем, что все основные свойства самосогласованных функций выводятся путем интегрирования различных функций по отрезкам из множества Q (для чего необходима выпуклость). В случае выпуклого множества Q неотрицательность второй производной эквивалентна выпуклости функции f.

Замечание 2.5. При определении самосогласованных функций некоторые авторы дополнительно требуют, чтобы функция f обладала *барьерным свойством* или, эквивалентно, являлась *замкнутой*. Как будет видно, для доказательства большинства свойств самосогласованных функций это требование оказывается не нужным. Более детальное обсуждение этого требования будет выполнено в разделе 4.

Замечание 2.6. Конкретная константа 2 в определении стандартной самосогласованной функции используется для, чтобы функция $t \to -\ln t$, играющая ключевую роль в рассматриваемой теории, являлась стандартной самосогласованной без дополнительной нормировки (см. пример 2.7).

Перейдем к рассмотрению наиболее важных примеров самосогласованных функций.

Пример 2.7 (Логарифмический барьер для луча)**.** Функция $f : (0, +\infty) \to \mathbb{R}$, заданная формулой

$$f(t) := -\ln t,$$

является стандартной самосогласованной.

Доказательство. Вычисляя производные, получаем

$$D^2 f(t)[v, v] = \frac{v^2}{t^2}, \qquad D^3 f(t)[v, v, v] = -\frac{2v^3}{t^3},$$

откуда

$$|D^3 f(t)[v, v, v]| = 2(D^2 f(t)[v, v])^{3/2}$$

для всех $t > 0$ и всех $v \in \mathbb{R}$. □

Пример 2.8. Функция $f : (0, +\infty) \to \mathbb{R}$, заданная формулой $f(t) := 1/t^p$, где $p > 0$, самосогласованной не является.

Доказательство. Действительно,

$$D^2 f(t)[v, v] = \frac{p(p+1)}{t^{p+2}} v^2, \qquad D^3 f(t)[v, v, v] = -\frac{p(p+1)(p+2)}{t^{p+3}} v^3$$

для всех $t > 0$ и всех $v \in \mathbb{R}$. Предположим, что $|D^3 f(t)[v, v, v]| \leq M(D^2 f(t)[v, v])^{3/2}$ для некоторого $M \geq 0$, всех $t > 0$ и всех $v \in \mathbb{R}$. Тогда $t^{p/2} \leq M'$ для некоторого $M' \geq 0$ (зависящего от M и p) и всех $t > 0$. Переходя к пределу при $t \to +\infty; t \in (0, +\infty)$ приходим к противоречию. □

Упражнение 2.9. Покажите, что функции $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ и $f_3 : (0, +\infty) \to \mathbb{R}$, заданные как

$$f_1(t) := e^t, \qquad f_2(t) := |t|^p, \qquad f_3(t) := \frac{1}{e^t - 1},$$

где $p \geq 3$, самосогласованными не являются.

Пример 2.10 (Постоянная, аффинная и квадратичная функции)**.** Пусть V — конечномерное евклидово пространство, $A : V \to V$ — самосопряженный положительно полуопределенный оператор, $b \in V$, $c \in \mathbb{R}$. Тогда функция $f : V \to \mathbb{R}$, заданная формулой

$$f(x) := \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c,$$

является самосогласованной с параметром 0. В частности, аффинная функция $x \mapsto \langle b, x\rangle + c$ и постоянная функция $x \mapsto c$ являются самосогласованными с параметром 0.

Пример 2.11 (Логарифмический барьер для множества решений квадратичного неравенства)**.** Пусть V — конечномерное евклидово пространство, $A : V \to V$ — самосопряженный положительно полуопределенный оператор, $b \in V$, $c \in \mathbb{R}$. Пусть $\phi : V \to \mathbb{R}$ — функция

$$\phi(x) := \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c.$$

Пусть $Q := \{x \in V : \phi(x) < 0\}$. Тогда функция $f : Q \to \mathbb{R}$, заданная формулой

$$f(x) := -\ln(-\phi(x)),$$

является стандартной самосогласованной.

Доказательство. Действительно, множество Q является открытым в V как прообраз открытого множества $(0, +\infty)$ в $\mathbb{R}$ (см. упражнение A.1).

Будем считать, что $Q \neq \emptyset$, поскольку иначе утверждение тривиально. Пусть $x \in Q$, $v \in V$. Дифференцируя, находим

$$D^2 f(x)[v, v] = \omega_1 + \omega_2^2, \qquad D^3 f(x)[v, v, v] = 3\omega_1\omega_2 + 2\omega_2^3,$$

где

$$\omega_1 := \frac{\langle Av, v\rangle}{-\phi(x)}, \qquad \omega_2 := \frac{\langle Ax + b, v\rangle}{-\phi(x)}.$$

Поскольку $\omega_1 \geq 0$ в силу положительной полуопределенности A (и того, что $\phi(x) < 0$), то отсюда следует, что $D^2 f(x)[v, v] \geq 0$; так как x и v произвольные, то это доказывает выпуклость f. Чтобы показать, что f — стандартная самосогласованная, осталось показать

$$|3\omega_1\omega_2 + 2\omega_2^3| \leq 2(\omega_1 + \omega_2^2)^{3/2}.$$

Если $\omega_2 = 0$, то неравенство тривиально, поэтому будем считать, что $\omega_2 \neq 0$. Обозначим $t := \omega_1/\omega_2^2 \geq 0$. Тогда

$$\frac{|3\omega_1\omega_2 + 2\omega_2^3|}{(\omega_1 + \omega_2^2)^{3/2}} = 2\frac{1 + \frac{3}{2}t}{(1 + t)^{3/2}}.$$

Остается воспользоваться неравенством Бернулли $(1 + t)^{3/2} \geq 1 + \frac{3}{2}t$. □

Упражнение 2.12 (Логарифмический барьер для конуса Лоренца)**.** Пусть V — конечномерное евклидово пространство, и пусть

$$K := \{(x, t) \in V \times \mathbb{R} : \|x\| < t\},$$

где $\|\| : V \to \mathbb{R}$ — евклидова норма $\|x\| := \langle x, x\rangle^{1/2}$. Покажите, что функция $f : K \to \mathbb{R}$, заданная формулой

$$f(x, t) := -\ln(t^2 - \|x\|^2),$$

является стандартной самосогласованной (в пространстве $V \oplus \mathbb{R}$). (*Подсказка*: адаптируйте доказательство из примера 2.11.)

Пример 2.13 (Логарифмический барьер для конуса $\mathbb{S}^n_+$)**.** Функция $f : \mathbb{S}^n_{++} \to \mathbb{R}$, заданная формулой

$$f(X) := -\ln \operatorname{Det}(X),$$

является стандартной самосогласованной (относительно пространства $\mathbb{S}^n$).

Доказательство. Пусть $X \in \mathbb{S}^n_{++}$, $V \in \mathbb{S}^n$ — произвольные. Дифференцируя, получаем

$$D^2 f(X)[V,V] = \sum_{i=1}^{n} \lambda_i^2, \qquad D^3 f(X)[V,V,V] = -2 \sum_{i=1}^{n} \lambda_i^3,$$

где $\lambda_1 \geq \cdots \geq \lambda_n$ — собственные значения матрицы $X^{-1/2} V X^{-1/2}$. Чтобы показать, что f — стандартная самосогласованная, осталось показать

$$\left| \sum_{i=1}^{n} \lambda_i^3 \right| \leq \left(\sum_{i=1}^{n} \lambda_i^2 \right)^{3/2}.$$

Но это следует из неравенства треугольника и утверждения 1.3. □

Как показывают предыдущие примеры, процесс проверки самосогласованности функции и вычисление ее параметра самосогласованности может быть довольно утомительным. К счастью, в нашем распоряжении имеются следующие правила, которые в большинстве случаев позволяют быстро установить самосогласованность функций, возникающих путем комбинирования функций, для которых уже известно, что они самосогласованы, и определить параметр их самосогласованности:

Упражнение 2.14 (Правила комбинирования самосогласованных функций)**.** Пусть V и W — конечномерные нормированные вещественные векторные пространства, Q и S — открытые множества в V и W соответственно, и пусть $f : Q \to \mathbb{R}$ и $g : S \to \mathbb{R}$ — функции.

(a) (Умножение на положительную константу) Если f является самосогласованной с параметром M, и $\alpha > 0$, то функция $\alpha f : Q \to \mathbb{R}$ является самосогласованной с параметром $M/\sqrt{\alpha}$.

(b) (Сумма) Пусть $V = W$. Если f и g являются самосогласованными с параметрами M и N соответственно, то сумма $f + g : Q \cap S \to \mathbb{R}$ является самосогласованной с параметром $\max\{M, N\}$.

(c) (Композиция с аффинным преобразованием) Если функция $g : S \to \mathbb{R}$ является самосогласованной с параметром N, и $A : V \to W$ — аффинное преобразование, то композиция $g \circ A : A^{-1}(S) \to \mathbb{R}$ также является самосогласованной с параметром N.

(d) (Сепарабельная сумма) Если f и g являются самосогласованными с параметрами M и N соответственно, то функция $h : Q \times S \to \mathbb{R}$, заданная формулой $h(x, y) := f(x) + g(y)$, является самосогласованной (относительно пространства $V \oplus W$) с параметром $\max\{M, N\}$.

Замечание 2.15. Согласно упражнению 2.14, любую самосогласованную функцию можно сделать стандартной самосогласованной, умножив ее на определенную константу. При этом, как показывают последние три правила, стандартные самосогласованные функции замкнуты относительно операций суммирования, композиции с аффинным преобразованием и сепарабельной суммы. По этой причине для удобства все самосогласованные функции предварительно нормируют, и далее работают лишь со стандартными самосогласованными функциями.

Пример 2.16 (Логарифмический барьер для полупространства)**.** Пусть V — конечномерное евклидово пространство, $a \in V$, $b \in \mathbb{R}$. Тогда функция $f : \{x \in V : \langle a, x \rangle < b\} \to \mathbb{R}$, заданная формулой

$$f(x) := -\ln(b - \langle a, x \rangle),$$

является стандартной самосогласованной как композиция логарифмического барьера для луча (пример 2.7) с аффинным преобразованием $A : V \to \mathbb{R}$, заданным формулой $A(x) := b - \langle a, x \rangle$.

Пример 2.17 (Логарифмический барьер для выпуклого многогранника)**.** Пусть V — конечномерное евклидово пространство, $a_1, \dots, a_m \in V$, $b_1, \dots, b_m \in \mathbb{R}$. Тогда функция $f : \{x \in V : \langle a_1, x\rangle < b_1;\ \dots;\ \langle a_m, x\rangle < b_m\} \to \mathbb{R}$, заданная формулой

$$f(x) := -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x\rangle),$$

является стандартной самосогласованной как сумма логарифмических барьеров для полупространств (пример 2.16).

Пример 2.18 (Логарифмический барьер для множества решений линейного матричного неравенства)**.** Пусть $A_1, \dots, A_m, B \in \mathbb{S}^n$ — симметричные матрицы, и пусть

$$Q := \{x \in \mathbb{R}^m : x_1 A_1 + \dots + x_m A_m \prec B\},$$

где запись $M_1 \prec M_2$ для $M_1, M_2 \in \mathbb{S}^n$ означает, что $M_2 - M_1 \in \mathbb{S}^n_{++}$. Тогда функция $f : Q \to \mathbb{R}$, заданная формулой

$$f(x) := -\ln \operatorname{Det}(B - x_1 A_1 - \dots - x_m A_m),$$

является стандартной самосогласованной (относительно пространства $\mathbb{R}^m$) как композиция логарифмического барьера для конуса $\mathbb{S}^n_+$ (пример 2.13) с аффинным преобразованием $A : \mathbb{R}^m \to \mathbb{S}^n_{++}$, заданным формулой

$$A(x) := B - x_1 A_1 - \dots - x_m A_m.$$

Пример 2.19. Функция $f : \mathbb{S}^n_{++} \times (0, +\infty) \to \mathbb{R}$, заданная формулой

$$f(X, t) := -\ln \operatorname{Det}(X) - \ln t,$$

является стандартной самосогласованной (относительно пространства $\mathbb{S}^n \oplus \mathbb{R}$) как сепарабельная сумма логарифмического барьера для конуса $\mathbb{S}^n_+$ (пример 2.13) и логарифмического барьера для луча (пример 2.7).

В заключение этого раздела приведем следующую техническую лемму, которая будет активно использоваться в дальнейшем:

Лемма 2.20 (Эквивалентное определение самосогласованных функций)**.** *Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое множество в V, и пусть $f : Q \to \mathbb{R}$ — трижды непрерывно дифференцируемая функция. Тогда f является самосогласованной с параметром $M \geq 0$, если и только если*

$$|D^3 f(x)[v_1, v_2, v_3]| \leq M \|v_1\|_{f,x} \|v_2\|_{f,x} \|v_3\|_{f,x}$$

для всех $x \in Q$ и всех $v_1, v_2, v_3 \in V$.

Доказательство. В одну сторону утверждение очевидно. Доказательство в другую сторону является довольно объемным, и поэтому не будет здесь приведено. □

3 Свойства самосогласованных функций

Выясним, как связаны локальные полунормы $\|\|_{f,x}$ и $\|\|_{f,y}$ для двух различных точек $x, y \in Q$. В силу непрерывности второй производной $D^2 f$, естественно ожидать, что, чем ближе друг к другу точки x и y, тем ближе друг к другу должны быть соответствующие полунормы. Следующая лемма показывает, что это, действительно, так и предоставляет конкретные неравенства:

Лемма 3.1 (Связь между локальными полунормами)**.** *Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция. Пусть $x, y \in Q$, причем $\|y - x\|_{f,x} < 1$. Тогда*

$$(1 - \|y - x\|_{f,x})\|v\|_{f,x} \leq \|v\|_{f,y} \leq (1 - \|y - x\|_{f,x})^{-1}\|v\|_{f,x} \tag{3.1}$$

для любого $v \in V$.

Для доказательства леммы 3.1 сперва докажем вспомогательное утверждение, которое полезно само по себе и говорит о том, что в специальном случае $v = y - x$ первое неравенство в (3.1) может быть усилено:

Лемма 3.2. *Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция. Тогда*

$$\frac{\|y - x\|_{f,x}}{1 + \|y - x\|_{f,x}} \leq \|y - x\|_{f,y} \leq \frac{\|y - x\|_{f,x}}{1 - \|y - x\|_{f,x}},$$

где первое неравенство верно для всех $x, y \in Q$, а второе неравенство верно для всех $x, y \in Q$, таких, что $\|y - x\|_{f,x} < 1$.

Доказательство. Докажем только второе неравенство; первое доказывается аналогично.

Пусть $\xi : [0, 1] \to \mathbb{R}$ — функция

$$\xi(t) := \|y - x\|^2_{f,x+t(y-x)} = D^2 f(x + t(y - x))[y - x, y - x].$$

Заметим, что функция ξ определена корректно благодаря выпуклости Q.

Поскольку f трижды непрерывно дифференцируема, то ξ непрерывно дифференцируема. Дифференцируя, находим

$$\xi'(t) = D^3 f(x + t(y - x))[y - x, y - x, y - x]$$

для всех $0 \leq t \leq 1$. Поскольку f является стандартной самосогласованной, то

$$|\xi'(t)| \leq 2\xi(t)^{3/2} \tag{3.2}$$

для всех $0 \leq t \leq 1$.

Пусть $\varepsilon > 0$, такое, что $\|y - x\|^2_{f,x} + \varepsilon < 1$ (оно обязательно существует, поскольку $\|y - x\|_{f,x} < 1$). Тогда из (3.2) получаем

$$\frac{|\xi'(t)|}{2(\xi(t) + \varepsilon)^{3/2}} \leq 1$$

для всех $0 \leq t \leq 1$. Отсюда

$$\int_0^1 \frac{\xi'(t)}{2(\xi(t) + \varepsilon)^{3/2}} dt \leq 1.$$

С другой стороны, по формуле Ньютона–Лейбница,

$$\int_0^1 \frac{\xi'(t)}{2(\xi(t) + \varepsilon)^{3/2}} dt = \frac{1}{(\xi(0) + \varepsilon)^{1/2}} - \frac{1}{(\xi(1) + \varepsilon)^{1/2}}.$$

Таким образом,

$$\frac{1}{(\xi(0) + \varepsilon)^{1/2}} - \frac{1}{(\xi(1) + \varepsilon)^{1/2}} \leq 1,$$

или, используя определение ξ,

$$\frac{1}{(\|y - x\|^2_{f,x} + \varepsilon)^{1/2}} - \frac{1}{(\|y - x\|^2_{f,y} + \varepsilon)^{1/2}} \leq 1.$$

Выполняя алгебраические преобразования и используя $\|y-x\|^2_{f,x}+\varepsilon<1$, получаем

$$(\|y-x\|^2_{f,y}+\varepsilon)^{1/2} \le \frac{(\|y-x\|^2_{f,x}+\varepsilon)^{1/2}}{1-(\|y-x\|^2_{f,x}+\varepsilon)^{1/2}}.$$

В силу произвольности ε, полученное неравенство верно для всех $\varepsilon \in (0, 1-\|y-x\|^2_{f,x})$. Остается перейти к пределу при $\varepsilon \to 0; \varepsilon \in (0, 1-\|y-x\|^2_{f,x})$. □

Доказательство леммы 3.1. Пусть $\xi : [0,1] \to \mathbb{R}$ — функция

$$\xi(t) := \|v\|^2_{f,x+t(y-x)} = D^2 f(x+t(y-x))[v,v].$$

Заметим, что ξ определена корректно в силу выпуклости Q.

Поскольку f трижды непрерывно дифференцируема, то ξ непрерывно дифференцируема. Дифференцируя, находим

$$\xi'(t) = D^3 f(x+t(y-x))[v,v,y-x]$$

для всех $0 \le t \le 1$. Используя лемму 2.20 и лемму 3.2 (для $y' = x+t(y-x)$), получаем

$$|\xi'(t)| \le 2\xi(t)\|y-x\|_{f,x+t(y-x)} \le \frac{2r}{1-tr}\xi(t)$$

для всех $0 \le t \le 1$. Пусть $\varepsilon > 0$ — произвольное число. Тогда

$$\frac{|\xi'(t)|}{\xi(t)+\varepsilon} \le \frac{2r}{1-tr}$$

для всех $0 \le t \le 1$. Отсюда

$$\left|\int_0^1 \frac{\xi'(t)}{\xi(t)+\varepsilon} dt\right| \le \int_0^1 \frac{2r}{1-tr} dt = -2\ln(1-r).$$

По формуле Ньютона–Лейбница

$$\int_0^1 \frac{\xi'(t)}{\xi(t)+\varepsilon} dt = \ln(\xi(1)+\varepsilon) - \ln(\xi(0)+\varepsilon) = \ln\frac{\xi(1)+\varepsilon}{\xi(0)+\varepsilon}.$$

Таким образом,

$$2\ln(1-r) \le \ln\frac{\xi(1)+\varepsilon}{\xi(0)+\varepsilon} \le -2\ln(1-r),$$

или, эквивалентно,

$$(1-r)^2(\xi(0)+\varepsilon) \le \xi(1)+\varepsilon \le \frac{\xi(0)+\varepsilon}{(1-r)^2}.$$

Используя определение ξ и r, это принимает вид

$$(1-\|y-x\|_{f,x})^2(\|v\|^2_{f,x}+\varepsilon) \le \|v\|^2_{f,y}+\varepsilon \le \frac{\|v\|^2_{f,x}+\varepsilon}{(1-\|y-x\|_{f,x})^2}.$$

Переходя к пределу при $\varepsilon \to 0; \varepsilon > 0$, получаем (3.1). □

Напомним, что для стандартного класса μ-сильно выпуклых функций с L-липшицевым градиентом для всех x, y справедливы неравенства

$$\mu\|y-x\|^2 \le \langle \nabla f(y) - \nabla f(x), y-x\rangle \le L\|y-x\|^2,$$
$$\frac{\mu}{2}\|y-x\|^2 \le f(y) - f(x) - \langle \nabla f(x), y-x\rangle \le \frac{L}{2}\|y-x\|^2,$$

которые означают, что вариация $\langle \nabla f(y) - \nabla f(x), y - x\rangle$, а также погрешность линеаризации $f(y) - f(x) - \langle \nabla f(x), y - x\rangle$ ведут себя примерно так же, как и квадрат нормы $\|y - x\|^2$. Оказывается, что при близких x и y такой же результат имеет место и для самосогласованных функций, причем в пределе $\mu = L = 1$:

Утверждение 3.3 (Основные функциональные неравенства)**.** *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция. Тогда*

$$\frac{\|y - x\|_{f,x}^2}{1 + \|y - x\|_{f,x}} \le \langle \nabla f(y) - \nabla f(x), y - x\rangle \le \frac{\|y - x\|_{f,x}^2}{1 - \|y - x\|_{f,x}}, \tag{3.3}$$

$$\omega(\|y - x\|_{f,x}) \le f(y) - f(x) - \langle \nabla f(x), y - x\rangle \le \omega^*(\|y - x\|_{f,x}), \tag{3.4}$$

где левые неравенства верны для всех $x, y \in Q$, а правые неравенства верны для всех $x, y \in Q$, таких, что $\|y - x\|_{f,x} < 1$. Здесь $\omega : [0, +\infty) \to \mathbb{R}$ и $\omega^* : [0, 1) \to \mathbb{R}$ — функции

$$\begin{aligned} \omega(t) &:= t - \ln(1 + t) = \frac{t^2}{2} - \frac{t^3}{3} + \frac{t^4}{4} - \dots, \\ \omega^*(s) &:= -s - \ln(1 - s) = \frac{s^2}{2} + \frac{s^3}{3} + \frac{s^4}{4} + \dots. \end{aligned} \tag{3.5}$$

Доказательство. Пусть $\xi : [0, 1] \to \mathbb{R}$ — функция

$$\xi(t) := \langle \nabla f(x + t(y - x)), y - x\rangle.$$

Заметим, что ξ определена корректно в силу выпуклости Q.

Так как f трижды непрерывно дифференцируема, то ξ непрерывно дифференцируема. Дифференцируя, находим

$$\xi'(t) = \langle \nabla^2 f(x + t(y - x))(y - x), y - x\rangle = \|y - x\|_{f,x+t(y-x)}^2$$

для всех $0 \le t \le 1$. Используя однородность полунормы, а также лемму 3.2 (для $y' = x + t(y - x)$), получаем

$$\|y - x\|_{f,x+t(y-x)}^2 \le \frac{\|y - x\|_{f,x}^2}{(1 - t\|y - x\|_{f,x})^2}$$

для всех $0 \le t \le 1$. (Заметим, что $t\|y - x\|_{f,x} < 1$ для всех $0 \le t \le 1$, поскольку $\|y - x\|_{f,x} < 1$.)

Обозначим $r := \|y - x\|_{f,x}$. Таким образом,

$$\xi'(t) \le \frac{r^2}{(1 - tr)^2}$$

для всех $0 \le t \le 1$. Применяя формулу Ньютона–Лейбница, получаем

$$\xi(1) - \xi(0) = \int_0^1 \xi'(t)dt \le \int_0^1 \frac{r^2}{(1 - tr)^2} dt = \frac{r^2}{1 - r}.$$

Подставляя определение ξ и r, получаем второе неравенство в (3.3); первое неравенство доказывается аналогично.

Для доказательства (3.4) нужно воспользоваться формулой Ньютона–Лейбница и неравенствами (3.3); подробный вывод предоставляется в качестве упражнения. □

Упражнение 3.4 (Функции ω и ω^* являются сопряженными)**.** Покажите, что функции ω и ω^*, определенные в (3.5), для всех $t \ge 0$ и всех $0 \le s < 1$ удовлетворяют соотношениям

$$\omega(t) = \max_{0 \le s' < 1} \{ts' - \omega^*(s')\}, \qquad \omega^*(s) = \max_{t' \ge 0} \{t's - \omega(t')\}.$$

Интересной особенностью самосогласованных функций является то, что некоторые локальные характеристики также являются и глобальными. Например, справедливо следующее утверждение:

Утверждение 3.5 (Невырожденность в точке эквивалентна невырожденности всюду)**.** *Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция. Если $D^2 f(x)[v, v] = 0$ для некоторого $x \in Q$ и некоторого $v \in V$, то $D^2 f(y)[v, v] = 0$ для всех $y \in Q$. Другими словами, если вторая производная $D^2 f$ является невырожденной в некоторой точке $x \in Q$, то она также является невырожденной всюду на Q.*

Доказательство. Пусть $\xi : [0, 1] \to \mathbb{R}$ — функция

$$\xi(t) := D^2 f(x + t(y - x))[v, v].$$

В силу выпуклости Q, функция ξ определена корректно. Кроме того, поскольку f трижды непрерывно дифференцируема, то ξ является непрерывно дифференцируемой.

Дифференцируя, находим

$$\xi'(t) = D^3 f(x + t(y - x))[v, v, y - x].$$

Применяя лемму 2.20, получаем

$$|\xi'(t)| \leq 2\xi(t)\|y - x\|_{f,x+t(y-x)}.$$

для всех $0 \leq t \leq 1$.

В силу непрерывности, функция $t \mapsto \|y - x\|_{f,x+t(y-x)}$ ограничена на интервале $[0, 1]$ (по теореме Вейерштрасса). Поэтому

$$|\xi'(t)| \leq A\xi(t)$$

для некоторого $A > 0$ и всех $0 \leq t \leq 1$.

Пусть $\varepsilon > 0$ — произвольное число. Тогда

$$\left|\frac{\xi'(t)}{\xi(t) + \varepsilon}\right| \leq A$$

для всех $0 \leq t \leq 1$. Отсюда

$$\left|\int_0^1 \frac{\xi'(t)}{\xi(t) + \varepsilon}\right| \leq A.$$

С другой стороны, по формуле Ньютона–Лейбница,

$$\int_0^1 \frac{\xi'(t)}{\xi(t) + \varepsilon} dt = \ln(\xi(1) + \varepsilon) - \ln(\xi(0) + \varepsilon) = \ln\frac{\xi(1) + \varepsilon}{\xi(0) + \varepsilon}.$$

Таким образом,

$$e^{-A}(\xi(0) + \varepsilon) \leq \xi(1) + \varepsilon \leq e^{A}(\xi(0) + \varepsilon).$$

Остается воспользоваться определением ξ и перейти к пределу при $\varepsilon \to 0; \varepsilon > 0$. □

В дальнейшем будем называть самосогласованную функцию *невырожденной*, если ее вторая производная всюду невырождена.

Напомним, что каждый вектор $s \in V$ в конечномерном евклидовом пространстве V порождает линейный функционал $v \mapsto \langle s, v\rangle$. Если в пространстве V помимо скалярного произведения также задана некоторая норма $\|\|$ (не обязательно порожденная скалярным произведением), то *сопряженная норма* вектора s определяется как операторная норма этого функционала:

$$\|s\|^* := \max_{v \in V : \|v\|=1} |\langle s, v\rangle|.$$

(Это определение корректное, поскольку максимум достигается по теореме Вейерштрасса в силу компактности единичной сферы.)

Упражнение 3.6 (Основные свойства сопряженной нормы)**.** Пусть V — конечномерное евклидово пространство, и пусть $\|\| : V \to \mathbb{R}$ — норма в V (не обязательно индуцированная скалярным произведением). Покажите, что:

(a) Сопряженная норма $\|\|^*$, действительно, является нормой в V.

(b) (Неравенство Гельдера) Для всех $s, v \in V$ выполнено

$$|\langle s, v \rangle| \leq \|s\|^* \|v\|.$$

(c) (Эквивалентные формы записи сопряженной нормы) Для всех $s \in V$ выполнено

$$\|s\|^* = \max_{\|v\|=1} \langle s, v \rangle = \max_{\|v\| \leq 1} |\langle s, v \rangle| = \max_{\|v\| \leq 1} \langle s, v \rangle.$$

(d) (Сопряжение является инволюцией) $(\|\|^*)^* = \|\|$.

Упражнение 3.7. Пусть V — конечномерное евклидово пространство, $B : V \to V$ — самосопряженный положительно определенный оператор, и пусть $\|\| : V \to \mathbb{R}$ — норма

$$\|v\| := \langle Bv, v \rangle^{1/2}.$$

Пусть $s \in V$. Покажите, что

$$\|s\|^* = \langle B^{-1} s, s \rangle^{1/2}.$$

(*Подсказка*: Используя $B^{1/2}$, оцените $|\langle s, v \rangle|$ сверху с помощью неравенства Коши–Буняковского; далее предъявите v, для которого неравенство переходит в равенство.)

Напомним, что если самосогласованная функция f невырожденная, то для любой точки $x \in Q$ локальная полунорма $\|\|_{f,x}$ является нормой в V. Таким образом, можно ввести *локальную сопряженную норму* $\|\|_{f,x} : V \to \mathbb{R}$, заданную как

$$\|s\|_{f,x}^* := \max_{v \in V : \|v\|_{f,x}=1} |\langle s, v \rangle| = \langle \nabla^2 f(x)^{-1} s, s \rangle^{1/2}.$$

В дальнейшем наиболее важную роль будет играть локальная норма градиента:

$$\|\nabla f(x)\|_{f,x}^* = \langle \nabla^2 f(x)^{-1} \nabla f(x), \nabla f(x) \rangle^{1/2}.$$

Интуитивно норма градиента $\nabla f(x)$ показывает, насколько точка x далеко от решения. Следующее утверждение подтверждает эту мысль:

Утверждение 3.8. *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция. Пусть f достигает своего минимума на Q, и пусть f^* — соответствующее минимальное значение. Тогда*

$$\omega(\|\nabla f(x)\|_{f,x}^*) \leq f(x) - f^* \leq \omega^*(\|\nabla f(x)\|_{f,x}^*),$$

где первое неравенство выполнено для всех $x \in Q$, а второе неравенство выполнено для всех $x \in Q$, таких, что $\|\nabla f(x)\|_{f,x}^* < 1$.

Доказательство сразу же следует из следующей леммы:

Лемма 3.9. *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция. Тогда*

$$\omega(\|\nabla f(y) - \nabla f(x)\|_{f,y}^*) \leq f(y) - f(x) - \langle \nabla f(x), y - x \rangle \leq \omega^*(\|\nabla f(y) - \nabla f(x)\|_{f,y}^*),$$

где первое неравенство верно для $x, y \in Q$, а второе неравенство верно для всех $x, y \in Q$, таких, что $\|\nabla f(y) - \nabla f(x)\|_{f,y}^* < 1$.

Доказательство. Докажем только второе неравенство; доказательство первого неравенства предоставляется в качестве упражнения.

Пусть $g : Q \to \mathbb{R}$ — функция
$$g(z) := f(z) - \langle \nabla f(x), z \rangle.$$
Согласно упражнению 2.14 и примеру 2.10, функция g является стандартной самосогласованной (как сумма стандартной самосогласованной функции f и линейного функционала $z \mapsto -\langle \nabla f(x), z \rangle$). Заметим, что функции f и g имеют одинаковые вторые производные, поэтому $\|\|_{g,y} = \|\|_{f,y}$.

В силу утверждения 3.3, имеем
$$g(x) \geq g(y) + \langle \nabla g(y), x - y \rangle + \omega(\|x - y\|_{f,y})$$
В силу определения двойственной нормы,
$$\langle \nabla g(y), x - y \rangle \geq -\|\nabla g(y)\|^*_{f,y} \|x - y\|_{f,y}$$
Таким образом,
$$g(x) \geq g(y) - \|\nabla g(y)\|^*_{f,y} \|x - y\|_{f,y} + \omega(\|x - y\|_{f,y}).$$
Из упражнения 3.4, получаем
$$-\|\nabla g(y)\|^*_{g,y} \|x - y\|_{f,y} + \omega(\|x - y\|_{f,y}) \geq -\omega^*(\|\nabla g(y)\|^*_{f,y}).$$
В итоге,
$$g(x) \geq g(y) - \omega^*(\|\nabla g(y)\|^*_{f,y}).$$
Остается подставить определение g. □

Наконец, обсудим вопросы о существовании и единственности минимума самосогласованной функции.

Упражнение 3.10 (Единственность минимума)**.** Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция. Покажите, что если точка минимума функции f существует, то она единственная. (*Подсказка*: используйте утверждение 3.3.)

Утверждение 3.11 (Существование минимума)**.** *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция. Тогда точка минимума функции f существует, если и только если найдется $x_0 \in Q$, такое, что $\|\nabla f(x_0)\|^*_{f,x_0} < 1$.*

Доказательство. В одну сторону утверждение тривиально: если $x^* \in Q$ — точка минимума f, то, согласно условию оптимальности первого порядка, $\nabla f(x^*) = 0$, откуда $\|\nabla f(x^*)\|^*_{f,x^*} = 0$.

Теперь докажем обратное. Пусть $\|\nabla f(x_0)\|^*_{f,x_0} < 1$ для некоторого $x_0 \in Q$. Пусть
$$L_0 := \{x \in Q : f(x) \leq f(x_0)\}.$$
Покажем, что функция f имеет точку минимума на множестве L_0; отсюда будет следовать утверждение (почему?). Поскольку f непрерывна, то, согласно теореме Вейерштрасса, достаточно показать, что множество L_0 компактно. Так как пространство V является конечномерным, то достаточно показать, что L_0 замкнуто и ограничено. Замкнутость L_0 сразу же следует из того, что L_0 является прообразом замкнутого множества $(0, f(x_0)]$ в $\mathbb{R}$ при непрерывном отображении f.

Докажем, что L_0 ограничено. Пусть $x \in L_0$ — произвольное. Согласно упражнению 3.3,
$$f(x) \geq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \omega(\|x - x_0\|_{f,x_0}).$$
Далее
$$\langle \nabla f(x_0), x - x_0 \rangle \geq -\|\nabla f(x_0)\|^*_{f,x_0} \|x - x_0\|_{f,x_0}.$$

Таким образом,
$$f(x) \geq f(x_0) - \|\nabla f(x_0)\|^*_{f,x_0}\|x - x_0\|_{f,x_0} + \omega(\|x - x_0\|_{f,x_0}).$$
С другой стороны, так как $x_0 \in L_0$, то $f(x) \leq f(x_0)$. Объединяя, получаем, что
$$\phi(\|x - x_0\|_{f,x_0}) \leq \|\nabla f(x_0)\|^*_{f,x_0},$$
где $\phi : [0, +\infty) \to \mathbb{R}$ — функция
$$\phi(t) := \frac{\omega(t)}{t} = 1 - \frac{\ln(1+t)}{t}.$$
Заметим, что функция ϕ является непрерывной и строго монотонно возрастающей; таким образом, существует обратная функция $\phi^{-1} : [0, +\infty) \to \mathbb{R}$. Значит,
$$\|x - x_0\|_{f,x_0} \leq \phi^{-1}(\|\nabla f(x_0)\|^*_{f,x_0}).$$
Отсюда следует ограниченность L_0 (в конечномерном пространстве все нормы эквивалентны). □

4 Барьерное свойство и эллипсоид Дикина

В том случае, когда область определения Q функции $f : Q \to \mathbb{R}$ является собственным подмножеством пространства V, при построении итеративных методов минимизации функции f необходимо следить за тем, чтобы точки, генерируемые методом, принадлежали множеству Q. Для самосогласованных функций наиболее простой способ контролировать эту принадлежность появляется при добавлении следующего дополнительного свойства, представляющего самостоятельный интерес:

Определение 4.1 (Барьерное свойство)**.** Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое множество в V, и пусть $f : Q \to \mathbb{R}$ — непрерывная функция. Говорят, что f обладает *барьерным свойством*, если для любой граничной точки $x_0 \in \partial Q$ выполнено
$$\lim_{x \to x_0; x \in Q} f(x) = +\infty.$$
В этом случае функцию f также называют *барьерной функцией* или просто *барьером* для множества Q.

Пример 4.2. В случае $Q = V$, барьерное свойство выполняется бессодержательно, поскольку $\partial V = \emptyset$. Таким образом, любая непрерывная функция, определенная на всем пространстве, автоматически обладает барьерным свойством.

Пример 4.3 (Барьеры для луча)**.** Каждая из функций $f_1, f_2, f_3 : (0, +\infty) \to \mathbb{R}$, заданных как
$$f_1(t) := -\ln t, \qquad f_2(t) := \frac{1}{t^p}, \qquad f_3(t) := \frac{1}{e^t - 1},$$
где $p > 0$, является барьером для луча $(0, +\infty)$. (В этом случае $\partial(0, +\infty) = \{0\}$.) Однако самосогласованной функцией является только f_1 (см. примеры 2.7, 2.8 и 2.9).

Пример 4.4. Функции $f_1, f_2 : (0, +\infty) \to \mathbb{R}$, заданные как
$$f_1(t) := t^p, \qquad f_2(t) := e^{-t},$$
где $p > 0$, барьерным свойством не обладают.

Пример 4.5 (Логарифмический барьер для множества решений квадратичного неравенства)**.** Пусть V — конечномерное евклидово пространство, $A : V \to V$ — самосопряженный положительно полуопределенный оператор, $b \in V$, $c \in \mathbb{R}$. Пусть $\phi : V \to \mathbb{R}$ — функция

$$\phi(x) := \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c.$$

Пусть

$$Q := \{x \in V : \phi(x) < 0\}.$$

Тогда функция $f : Q \to \mathbb{R}$, заданная формулой

$$f(x) := -\ln(-\phi(x)),$$

является барьером для множества Q.

Доказательство. Если $\partial Q = \emptyset$, то барьерное свойство выполняется бессодержательно, поэтому пусть $\partial Q \neq \emptyset$, и пусть $x_0 \in \partial Q$. Тогда $\phi(x_0) = 0$ (см. упражнение A.2). В силу непрерывности ϕ, отсюда следует, что $\phi(x) \to 0$ при $x \to x_0; x \in Q$. Остается воспользоваться непрерывностью логарифма. □

Упражнение 4.6. Покажите, что логарифмический барьер для конуса Лоренца и логарифмический барьер для конуса положительно полуопределенных матриц $\mathbb{S}^n_+$, определенные в примерах 2.12 и 2.13 соответственно, действительно, являются барьерами.

Упражнение 4.7 (Правила комбинирования барьерных функций)**.** Покажите, что для барьерных функций справедливы те же самые правила комбинирования, что и для самосогласованных функций (упражнение 2.14).

Самосогласованные барьерные функции обладают очень интересным геометрическим свойством.

Пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция, определенная на открытом выпуклом множестве в конечномерном евклидовом пространстве V, и пусть $x \in Q$. Тогда открытый шар относительно полунормы $\|\|_{f,x}$ радиуса $r > 0$ с центром в точке x, т. е. множество

$$B_f(x, r) := \{y \in V : \|y - x\|_{f,x} < r\},$$

называется открытым *эллипсоидом Дикина* радиуса r с центром в точке x (относительно функции f).

В том случае, когда полунорма $\|\|_{f,x}$ является нормой (т. е. в случае невырожденной f), из открытости множества Q, конечно же, следует, что для достаточно малого $r_x > 0$ эллипсоид Дикина $B_f(x, r_x)$ целиком лежит внутри множества Q (в силу эквивалентности любых норм в конечномерном пространстве). Особенностью *барьерных* самосогласованных функций является то, что эллипсоид Дикина $B_f(x, r_x)$ лежит внутри Q даже в случае вырожденной функции f, и, что самое интересное, радиус r_x может быть выбран одним и тем же для всех точек x:

Утверждение 4.8 (Область определения содержит единичный открытый эллипсоид Дикина)**.** *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x \in Q$. Тогда $B_f(x, 1) \subseteq Q$. Другими словами, если $y \in V$ и $\|y - x\|_{f,x} < 1$, то $y \in Q$.*

Доказательство. Пусть $y \in V$ и $\|y - x\|_{f,x} < 1$. Покажем, что в этом случае $y \in Q$.

Чтобы прийти к противоречию, предположим, что $y \notin Q$. Положим

$$t' := \sup\{t \in [0, 1] : x + t(y - x) \in Q\}.$$

Тогда точка $y' := x + t'(y - x)$ является граничной точкой множества Q (почему?), причем $\{x + t(y - x) : 0 \leq t < t'\} \subseteq Q$. Согласно утверждению (3.3), имеем

$$f(x + t(y - x)) \leq f(x) + t\langle\nabla f(x), y - x\rangle + \omega^*(t\|y - x\|_{f,x}).$$

для всех $0 \le t < t'$. Переходя к пределу (используя непрерывность ω^*), получаем

$$\lim_{t \to t'; 0 \le t < t'} f(x + t(y - x)) \le f(x) + t' \langle \nabla f(x), y - x \rangle + \omega^*(t' \|y - x\|_{f,x}).$$

Но поскольку f обладает барьерным свойством и y' — граничная точка Q, то $\lim_{t \to t'; 0 \le t < t'} f(x + t(y - x)) = +\infty$, что противоречит выписанному неравенству. $\square$

Рассмотрим, как с помощью этого утверждения можно контролировать, скажем, допустимость точек, генерируемых методом Ньютона. Пусть $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, и пусть

$$x_1 := x_0 - \nabla^2 f(x_0)^{-1} \nabla f(x_0).$$

Тогда

$$\|x_1 - x_0\|_{f,x_0} = \langle \nabla^2 f(x_0)(x_1 - x_0), x_1 - x_0 \rangle^{1/2} = \langle \nabla^2 f(x_0)^{-1} \nabla f(x_0), \nabla f(x_0) \rangle^{1/2} = \|\nabla f(x_0)\|^*_{f,x_0}.$$

Таким образом, в случае $\|\nabla f(x_0)\|^*_{f,x_0} < 1$ можно гарантировать, что $x_1 \in Q$.

Упражнение 4.9. Пусть V — конечномерное нормированное вещественное векторное пространство, Q — открытое выпуклое множество в V, и пусть $f : Q \to \mathbb{R}$ — стандартная самосогласованная функция, обладающая барьерным свойством. Покажите, что если Q не содержит прямых (т. е. множеств вида $\{x + \alpha v : \alpha \in \mathbb{R}\}$, где $x \in Q$, $v \in V$, $v \ne 0$), то f является невырожденной. (*Подсказка*: рассуждайте от противного и используйте утверждение 4.8.)

5 Стандартный метод Ньютона

Теперь в нашем распоряжении имеется вся необходимая теория для того, чтобы доказать аналог теоремы 1.1 для самосогласованных функций. Базовым элементом в этом доказательстве является следующая лемма, показывающая, как изменяется локальная норма градиента за один шаг стандартного метода Ньютона:

Лемма 5.1 (Прогресс одного шага стандартного метода Ньютона)**.** *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, и пусть $\|\nabla f(x_0)\|^*_{f,x_0} < 1$. Положим*

$$x_1 := x_0 - \nabla^2 f(x_0)^{-1} \nabla f(x_0).$$

Тогда $x_1 \in Q$, и

$$\|\nabla f(x_1)\|^*_{f,x_1} \le \left(\frac{\|\nabla f(x_0)\|^*_{f,x_0}}{1 - \|\nabla f(x_0)\|^*_{f,x_0}} \right)^2. \tag{5.1}$$

Доказательство. Нетрудно видеть, что

$$\|x_1 - x_0\|_{f,x_0} = \|\nabla f(x_0)\|^*_{f,x_0}. \tag{5.2}$$

Поскольку $\|\nabla f(x_0)\|^*_{f,x_0} < 1$, то из утверждения 4.8 заключаем, что $x_1 \in Q$.

По определению сопряженной нормы $\|\|^*_{f,x_1}$, чтобы показать (5.1), нужно доказать, что

$$|\langle \nabla f(x_1), v \rangle| \le \left(\frac{\|\nabla f(x_0)\|^*_{f,x_0}}{1 - \|\nabla f(x_0)\|^*_{f,x_0}} \right)^2 \|v\|^2_{f,x_1} \tag{5.3}$$

для всех $v \in V$.

Пусть $v \in V$ — произвольный вектор, и пусть $\xi : [0,1] \to \mathbb{R}$ — функция

$$\xi(t) := \langle \nabla f(x_0 + t(x_1 - x_0)), v \rangle.$$

Дифференцируя, находим

$$\xi'(t) = \langle \nabla^2 f(x_0 + t(x_1 - x_0))(x_1 - x_0), v \rangle = D^2 f(x_0 + t(x_1 - x_0))[v, x_1 - x_0].$$

и

$$\xi''(t) = D^3 f(x_0 + t(x_1 - x_0))[v, x_1 - x_0, x_1 - x_0]$$

для всех $0 \le t \le 1$. Согласно лемме 2.20, получаем

$$|\xi''(t)| \le 2\|v\|_{f, x_0 + t(x_1 - x_0)} \|x_1 - x_0\|^2_{f, x_0 + t(x_1 - x_0)}$$

для всех $0 \le t \le 1$. Применяя лемму 3.1 и тождество (5.2), получаем

$$\|v\|_{f, x_0 + t(x_1 - x_0)} \le \frac{\|v\|_{f, x_0}}{1 - t\|\nabla f(x_0)\|^*_{f, x_0}}$$

и

$$\|x_1 - x_0\|_{f, x_0 + t(x_1 - x_0)} \le \frac{\|\nabla f(x_0)\|^*_{f, x_0}}{1 - t\|\nabla f(x_0)\|^*_{f, x_0}}$$

для всех $0 \le t \le 1$. Таким образом,

$$|\xi''(t)| \le \frac{2(\|\nabla f(x_0)\|^*_{f, x_0})^2 \|v\|_{f, x_0}}{(1 - t\|\nabla f(x_0)\|^*_{f, x_0})^3} \tag{5.4}$$

для всех $0 \le t \le 1$.

Используя формулу Ньютона–Лейбница дважды, получаем

$$\xi(1) = \xi(0) + \xi'(0) + \int_0^1 (\xi'(\tau) - \xi'(0)) d\tau = \xi(0) + \xi'(0) + \int_0^1 \int_0^\tau \xi''(t) dt d\tau. \tag{5.5}$$

Применяя (5.4), получаем

$$\left| \int_0^1 \int_0^\tau \xi''(t) dt d\tau \right| \le \int_0^1 \int_0^\tau |\xi''(t)| dt d\tau \le \int_0^1 \int_0^\tau \frac{2(\|\nabla f(x_0)\|^*_{f, x_0})^2 \|v\|_{f, x_0}}{(1 - t\|\nabla f(x_0)\|^*_{f, x_0})^3} dt d\tau.$$

Вычисляя интеграл, находим

$$\int_0^1 \int_0^\tau \frac{2(\|\nabla f(x_0)\|^*_{f, x_0})^2}{(1 - t\|\nabla f(x_0)\|^*_{f, x_0})^3} dt d\tau = \frac{(\|\nabla f(x_0)\|^*_{f, x_0})^2}{1 - \|\nabla f(x_0)\|^*_{f, x_0}}.$$

Подставляя определение ξ и полученный результат в (5.5), в итоге получаем

$$|\langle \nabla f(x_1), v \rangle| \le \frac{(\|\nabla f(x_0)\|^*_{f, x_0})^2}{1 - \|\nabla f(x_0)\|^*_{f, x_0}} \|v\|_{f, x_0}.$$

Применяя утверждение 3.3 снова, получаем

$$\|v\|_{f, x_0} \le \frac{\|v\|_{f, x_1}}{1 - \|\nabla f(x_0)\|^*_{f, x_0}}.$$

Объединяя последние два неравенства, получаем (5.4). $\square$

Рассмотрим теперь всю последовательность $(x_k)_{k=0}^{\infty}$, построенную методом Ньютона:

$$x_{k+1} := x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k).$$

Для того, чтобы применить лемму 5.1 индуктивно ко всем шагам метода, необходимо обеспечить базовое неравенство $\|\nabla f(x_k)\|_{f,x_k}^* < 1$ для всех $k \geq 0$. Поскольку в начальный момент времени это неравенство выполнено, то достаточно потребовать монотонного убывания локальных норм $\|\nabla f(x_k)\|_{f,x_k}^*$ градиентов. Согласно формуле (5.1), для этого достаточно потребовать, чтобы выполнялось

$$\frac{\|\nabla f(x_0)\|_{f,x_0}^*}{(1 - \|\nabla f(x_0)\|_{f,x_0}^*)^2} \leq 1,$$

или, разрешая,

$$\|\nabla f(x_0)\|_{f,x_0}^* \leq \frac{3 - \sqrt{5}}{2} = 0.3819\ldots$$

Пусть $\|\nabla f(x_0)\|_{f,x_0}^* \leq \beta$ для некоторого $0 < \beta < \bar{\lambda}$. Тогда, применяя индукцию и лемму 5.1, нетрудно показать, что

$$\|\nabla f(x_{k+1})\|_{f,x_{k+1}}^* \leq (1 - \beta)^2 \|\nabla f(x_k)\|_{f,x_k}^*$$

для всех $k \geq 0$, откуда

$$\|\nabla f(x_k)\|_{f,x_k}^* \leq (1 - \beta)^2 \left(\frac{\beta}{(1 - \beta)^2} \right)^{2^k}.$$

для всех $k \geq 0$. Выбирая β таким образом, чтобы, скажем,

$$\frac{\beta}{(1 - \beta)^2} = \frac{1}{2},$$

приходим к следующему результату:

Теорема 5.2 (Квадратичная сходимость метода Ньютона для минимизации самосогласованной функции)**.** *Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, и пусть*

$$\|\nabla f(x_0)\|_{f,x_0}^* \leq 2 - \sqrt{3} = 0.2679\ldots.$$

Тогда последовательность $(x_k)_{k=0}^{\infty}$, заданная рекуррентно стандартным методом Ньютона

$$x_{k+1} := x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k),$$

определена корректно; при этом $x_k \in Q$ и

$$\|\nabla f(x_k)\|_{f,x_k}^* \leq \left(\frac{1}{2} \right)^{2^k}$$

для всех $k \geq 0$.

Несмотря на то, что эта теорема в терминах локальной нормы градиента, из нее следуют аналогичные результаты и в других терминах (невязка по функции, по точкам и т.д.). Например:

Упражнение 5.3. Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, и пусть $\|\nabla f(x_0)\|_{f,x_0}^* \leq 2 - \sqrt{3} = 0.2679\ldots$. Тогда последовательность $(x_k)_{k=0}^{\infty}$, заданная рекуррентно стандартным методом Ньютона

$$x_{k+1} := x_k - \nabla^2 f(x_k)^{-1} \nabla f(x_k),$$

определена корректно; при этом $x_k \in Q$ и

$$f(x_k) - f^* \leq \left(\frac{1}{4}\right)^{2^k}.$$

для всех $k \geq 0$. (Подсказка: воспользуйтесь утверждением 3.8 и неравенством $\omega^*(s) \leq \frac{s^2}{2(1-s)}$ для $0 \leq s < 1$.)

Вернемся к примеру из пролога:

Пример 5.4. Пусть $A \geq 1$, и пусть Q — открытый прямоугольник

$$Q := \{x \in \mathbb{R}^2 : |x_1| < A;\ |x_2| < 1\},$$

и пусть $f : Q \to \mathbb{R}$ — функция

$$f(x) := -\ln(A^2 - x_1^2) - \ln(1 - x_2^2).$$

Согласно примеру 2.7 и упражнению 2.14, функция f является стандартной самосогласованной, а также обладает барьерным свойством. Кроме того, согласно утверждению 4.9, она невырожденная.

Вычисляя, находим

$$(\|\nabla f(x)\|_{f,x}^*)^2 = \frac{2x_1^2}{A^2 + x_1^2} + \frac{2x_2^2}{1 + x_2^2} \leq 2\frac{x_1^2}{A^2} + 2x_2^2.$$

Таким образом, область квадратичной сходимости $\|\nabla f(x)\|_{f,x}^* < \beta$, даваемая теоремой 5.2, не меньше, чем эллипс

$$\{x \in \mathbb{R}^2 : \frac{x_1^2}{A^2} + x_2^2 \leq 2\beta^2\}$$

Для $\beta = 2 - \sqrt{3}$ получаем $2\beta^2 = 0.143\ldots$. То есть мы получили тот же результат, что и в начале, даже немного лучше ($0.143\ldots$ вместо 0.01).

6 Демпфированный метод Ньютона

Как прийти в зону квадратичной сходимости метода Ньютона?

Пусть $x_0 \in Q$. Согласно упражнению 3.3, для самосогласованной функции справедлива следующая оценка сверху:

$$f(x) \leq f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \omega^*(\|x - x_0\|_{f,x_0})$$

для всех $x \in Q$, таких, что $\|x - x_0\|_{f,x_0} < 1$. Минимизируя эту оценку, можно получить итерацию демпфированного метода Ньютона, причем с конкретной длиной шага:

Упражнение 6.1. Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция. Пусть $x_0 \in Q$. Покажите, что минимум функции

$$x \mapsto f(x_0) + \langle \nabla f(x_0), x - x_0 \rangle + \omega^*(\|x - x_0\|_{f,x_0})$$

на $\{y \in V : \|y - x_0\|_{f,x_0} < 1\}$ достигается в точке

$$x_1 := x_0 - \frac{1}{1 + \|\nabla f(x_0)\|_{f,x_0}^*}[\nabla^2 f(x_0)]^{-1} \nabla f(x_0).$$

(*Подсказка*: оцените функционал снизу по неравенству Гельдера, а затем воспользуйтесь упражнением 3.4; далее убедитесь, что нижняя оценка достигается на x_1.)

Из доказанных неравенств уже легко оценить прогресс одного шага метода Ньютона:

Упражнение 6.2 (Прогресс демпфированного метода Ньютона)**.** Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, и пусть

$$x_1 := x_0 - \frac{1}{1 + \|\nabla f(x_0)\|^*_{f,x_0}} [\nabla^2 f(x_0)]^{-1} \nabla f(x_0).$$

Покажите, что $x_1 \in Q$, и

$$f(x_0) - f(x_1) \geq \omega(\|\nabla f(x_0)\|^*_{f,x_0}).$$

(*Подсказка:* Воспользуйтесь утверждениями 4.8 и 3.3.)

В силу монотонности функции ω отсюда вытекает, что если $\|\nabla f(x_0)\|^*_{f,x_0} \geq \beta$, то

$$f(x_0) - f(x_1) \geq \omega(\beta).$$

Тогда общая трудоемкость попадания в область квадратичной сходимости составляет не более

$$\frac{f(x_0) - f^*}{\omega(\beta)}$$

шагов.

В заключение отметим, что демпфированный метод Ньютона сам переключается на квадратичную сходимость:

Упражнение 6.3 (Локальная сходимость демпфированного метода Ньютона)**.** Пусть V — конечномерное евклидово пространство, Q — открытое выпуклое множество в V, $f : Q \to \mathbb{R}$ — невырожденная стандартная самосогласованная функция, обладающая барьерным свойством. Пусть $x_0 \in Q$, причем $\|\nabla f(x_0)\|^*_{f,x_0} < 1$, и пусть

$$x_1 := x_0 - \frac{1}{1 + \|\nabla f(x_0)\|^*_{f,x_0}} [\nabla^2 f(x_0)]^{-1} \nabla f(x_0).$$

Покажите, что $x_1 \in Q$, и

$$\|\nabla f(x_1)\|^*_{f,x_1} \leq 2(\|\nabla f(x_0)\|^*_{f,x_0})^2.$$

(*Подсказка:* адаптируйте доказательство утверждения 5.1.)

A Краткий обзор основных топологических понятий

Для простоты сосредоточимся лишь на метрических пространствах, поскольку только они нас интересуют. Это включает в себя нормированные пространства как частный случай.

Обозначения: $\mathrm{int}(E)$ обозначает внутренность, $\overline{E}$ обозначает замыкание множества E (множество всех точек соприкосновения). Напомним, что $\overline{E} = \mathrm{int}(E) \cup \partial E$, $\partial E = \overline{E} \setminus \mathrm{int}(E)$. Вспомогательная лемма:

В прямой сумме $V \oplus W$ пространств введена стандартная норма/метрика $d_{V \oplus W}((x, y)) := \max\{d_V(x), d_W(y)\}$.

Преобразование $T : V \to W$ называется непрерывным, если для любой точки $x_0 \in V$ и любого $\varepsilon > 0$ найдется $\delta > 0$, такое, что $T(x) \in B_W(T(x_0), \varepsilon)$ для всех $x \in B_V(x_0, \delta)$.

Упражнение A.1. Пусть V и W — метрические пространства.

(a) Если E и S — открытые множества в V, то пересечение $E \cap S$ также является открытым множеством в V.

(b) Если S — открытое множество в W и $T : V \to W$ — непрерывное преобразование, то прообраз $T^{-1}(S)$ является открытым множеством в V.

(c) Если S — замкнутое множество в W и $T : V \to W$ — непрерывное преобразование, то прообраз $T^{-1}(S)$ является замкнутым множеством в V.

(d) Если E и S — открытые множества в V и W соответственно, то декартово произведение $E \times S$ является открытым множеством в $V \oplus W$.

Упражнение A.2 (Граница результата при операциях над множествами)**.** Пусть V и W — метрические пространства.

(a) Если E и S — множества в V, то

$$\partial(E \cap S) \subseteq (\partial E \cup \partial S) \cap \overline{E} \cap \overline{S} \subseteq \partial E \cup \partial S. \tag{A.1}$$

(b) Если S — множество в W и $T : V \to W$ — непрерывное преобразование, то

$$\partial T^{-1}(S) \subseteq T^{-1}(\partial S). \tag{A.2}$$

(c) Если E и S — множества в V и W соответственно, то

$$\partial(E \times S) = (\partial E \times \overline{S}) \cup (\overline{E} \times \partial S).$$

Приведите примеры, когда вложения (A.1) и (A.2) являются строгими.

Упражнение A.3. Покажите, что множество $\mathbb{S}^n_{++}$ является открытым, и

$$\partial \mathbb{S}^n_{++} = \{X \in \mathbb{S}^n_{+} : \mathrm{Det}(X) = 0\}.$$

(*Подсказка*: рассмотрите собственные значения; удобно работать в терминах операторной нормы.)