Continuous Optimization
Homework 4

Lev Khoroshansky

Задача 1. На протяжении всего задания будем пользоваться утверждением 1 из семинара 1 без явных на то ссылок. Для всех пунктов очевидно, что $r^* = 0$.

1. $r_k = \dfrac{99^k}{100^k}$. Исследуем предел отношения:

$$\limsup_{k\to\infty} \frac{r_{k+1}}{r_k} = \limsup_{k\to\infty} \frac{99^{k+1}/100^{k+1}}{99^k/100^k} = \frac{99}{100} < 1$$

   Следовательно, последовательность обладает линейной скоростью сходимости.

2. $r_k = \dfrac{99^{k^2}}{100^{k^2}}$. Исследуем предел отношения:

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \frac{99^{(k+1)^2}/100^{(k+1)^2}}{99^{k^2}/100^{k^2}} = \lim_{k\to\infty} \frac{99^{2k+1}}{100^{2k+1}} = 0$$

   Таким образом, данная последовательность обладает сверхлинейной скоростью сходимости. Исследуем на квадратичную: пусть $M > 0$, тогда рассмотрим неравенство

$$r_{k+1} \leqslant M\, r_k^2 \quad \longrightarrow \quad \frac{r_{k+1}}{r_k^2} \leqslant M \quad \longrightarrow \quad \frac{99^{(k+1)^2}/100^{(k+1)^2}}{99^{2k^2}/100^{2k^2}} \leqslant M \quad \longrightarrow \quad \lim_{k\to\infty} \frac{100^{k^2-2k-1}}{99^{k^2-2k-1}} = \infty$$

   Видно, что квадратичной сходимости нет.

3. $r_k = \dfrac{1}{k}$. Исследуем предел отношения:

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \frac{k}{k+1} = 1$$

   Следовательно, последовательность обладает сублинейной скоростью сходимости.

4. $r_k = \dfrac{1}{\sqrt{k}}$. Исследуем предел отношения:

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \frac{\sqrt{k}}{\sqrt{k+1}} = 1$$

   Таким образом, последовательность обладает сублинейной скоростью сходимости.

5. $r_k = \dfrac{1}{k!}$. Исследуем предел отношения:

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \frac{k!}{(k+1)!} = \lim_{k\to\infty} \frac{1}{k+1} = 0$$

Следовательно, последовательность обладает сверхлинейной сходимостью.

Исследуем на квадратичную: пусть $M > 0$, тогда рассмотрим неравенство

$$r_{k+1} \leqslant M\, r_k^2 \quad\longrightarrow\quad \frac{r_{k+1}}{r_k^2} \leqslant M \quad\longrightarrow\quad \frac{k!\,k!}{(k+1)!} \leqslant M \quad\longrightarrow\quad \lim_{k\to\infty} \frac{k!}{k+1} = \infty$$

Видно, что квадратичной сходимости нет.

6. $r_k = \dfrac{1}{k^k}$. Исследуем предел отношения:

$$\lim_{k\to\infty} \frac{r_{k+1}}{r_k} = \lim_{k\to\infty} \frac{k^k}{(k+1)^{(k+1)}} = e^{-1} \cdot \lim_{k\to\infty} \frac{1}{k+1} = 0$$

Таким образом, последовательность обладает сверхлинейной сходимостью.

Исследуем на квадратичную: пусть $M > 0$, тогда рассмотрим неравенство

$$r_{k+1} \leqslant M\, r_k^2 \quad\longrightarrow\quad \frac{r_{k+1}}{r_k^2} \leqslant M \quad\longrightarrow\quad \frac{(k+1)^{2(k+1)}}{k^k} \leqslant M \quad\longrightarrow\quad \lim_{k\to\infty} \frac{(k+1)^{(k+1)}}{k^k} \cdot (k+1)^{(k+1)} = \infty$$

Видно, что квадратичной сходимости нет.

Задача 2. Введём следующие обозначения:

$$\begin{aligned}
&f_0(x) = c_1x_1 + c_2x_2 + c_3x_3 = \langle c, x\rangle \to \min\\
&g_1(x) = (x_1 - a_1)^2 + 2(x_2 - a_2)^2 + 3(x_3 - a_3)^2 - b \leqslant 0\\
&g_2(x) = -b < 0
\end{aligned}$$

Если существует $c_i = 0$, то соответствующая координата не влияет на значение целевой функции. Следовательно, (как станет видно дальше) x_i будет равен a_i и не повлияет на ограничение $g_1(x) \leqslant 0$. Таким образом, мы можем спроецировать задачу на множество одной размерностью ниже, где верны те же самые рассуждения, поэтому всюду дальше $c \neq 0$. Граничные точки, в которых $g_1(x) = 0$, нас так же не интересуют, так как это точки барьера, ограничивающего нашу область.

Применим логарифмический барьер:

$$F(r, x) = f_0(x) - r\left[\ln(-g_1(x)) + \ln(-g_2(x))\right]$$

Найдём точки обнуления градиента:

$$\begin{aligned}
\frac{\partial}{\partial x_1} F(r, x) = c_1 - r\,\frac{2(x_1 - a_1)}{g_1(x)} = 0\\
\frac{\partial}{\partial x_2} F(r, x) = c_2 - r\,\frac{4(x_2 - a_2)}{g_1(x)} = 0\\
\frac{\partial}{\partial x_3} F(r, x) = c_3 - r\,\frac{6(x_3 - a_3)}{g_1(x)} = 0
\end{aligned}$$

Выразим $(x_1 - a_1)$ и $(x_2 - a_2)$ через $(x_3 - a_3)$:

$$\begin{aligned}
c_1 &= \frac{2r(x_1-a_1)}{g_1(x)} \longrightarrow \frac{(x_1-a_1)}{c_1} = \frac{g_1(x)}{2r} \longrightarrow (x_1-a_1) = \frac{3c_1}{c_3}(x_3-a_3) \\
c_2 &= \frac{4r(x_2-a_2)}{g_1(x)} \longrightarrow \frac{(x_2-a_2)}{c_2} = \frac{g_1(x)}{4r} \longrightarrow (x_2-a_2) = \frac{3c_2}{2c_3}(x_3-a_3) \\
c_3 &= \frac{6r(x_3-a_3)}{g_1(x)} \longrightarrow \frac{(x_3-a_3)}{c_3} = \frac{g_1(x)}{6r}
\end{aligned}$$

Решим для $v_3 = (x_3 - a_3)$:

$$6rv_3 = c_3\left[\frac{9c_1^2}{c_3^2}v_3^2 + \frac{9c_2^2}{2c_3^2}v_3^2 + 3v_3^2 - b\right] \longrightarrow \underbrace{\left(18c_1^2 + 9c_2^2 + 6c_3^2\right)}_{z} v_3^2 - 12c_3rv_3 - 2c_3^2b = 0$$

Откуда имеем

$$v_3 = \frac{6c_3r \pm |c_3|\sqrt{36r^2 + 2zb}}{z}$$

Знак выберем тот, что противоположен $\operatorname{sign}(c_3)$ (так как конечное выражение, которое будет зависеть от этого выбора, неотрицательно). Устремим r к нулю и получим

$$w_3 = \lim_{r\to 0} v_3 = \frac{-\operatorname{sign}(c_3)|c_3|\sqrt{2b}}{\sqrt{z}} \longrightarrow x_1^* = \frac{3c_1}{c_3}w_3 + a_1, \quad x_2^* = \frac{3c_2}{2c_3}w_3 + a_2, \quad x_3^* = w_3 + a_3$$

Выразим минимум целевой функции:

$$\begin{aligned}
f_0(x^*) &= \frac{3c_1^2}{c_3}w_3 + c_1a_1 + \frac{3c_2^2}{2c_3}w_3 + c_2a_2 + c_3w_3 + c_3a_3 \\
&= -\left[\underbrace{\frac{3c_1^2\sqrt{2b}}{\sqrt{z}} + \frac{3c_2^2\sqrt{b}}{\sqrt{2z}} + \frac{3c_3^2\sqrt{2b}}{\sqrt{z}}}_{\geqslant 0}\right] + \langle c, a\rangle
\end{aligned}$$

Задача 3. Введём следующие обозначения:

$$f_0(x) = x_1 + 2x_2, \quad g_1(x) = 1 - x_1x_2 \leqslant 0, \quad g_2(x) = x_1 + x_2 - 4 \leqslant 0, \quad g_3(x) = -x_1 \leqslant 0$$

Применим логарифмический барьер:

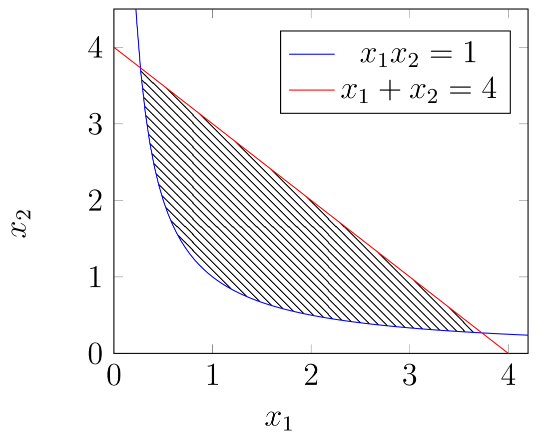
$$F(r,x) = f_0(x) - r\sum_{i=1}^{3}\ln(-g_i(x))$$

Найдём точки обнуления градиента:

$$\begin{aligned}
\frac{\partial}{\partial x_1}F(r,x) &= 1 + r\left[\frac{-x_2}{g_1(x)} + \frac{1}{g_2(x)} + \frac{-1}{g_3(x)}\right] = 0 \\
\frac{\partial}{\partial x_2}F(r,x) &= 2 + r\left[\frac{-x_1}{g_1(x)} + \frac{1}{g_2(x)}\right] = 0
\end{aligned}$$

И дальше сходим с ума, пытаясь найти невыражаемые переменные (многочасовые эксперименты показали, проблема метода барьеров в этой задаче заключается в том, что целевая функция безусловной оптимизации стремится к бесконечности).

Попробуем по-другому – изобразим область, задаваемую ограничениями:



Градиент целевой функции равен $(1, 2)$, откуда следует, что минимум лежит на грани $x_2 = {}^1\!/_{x_1}$ (градиент указывает в сторону максимума, но нам нужен минимум, а антиградиент указывает на синюю грань). Выразим x_2 через x_1 и приравняем производную к нулю, после чего найдём точку минимума:

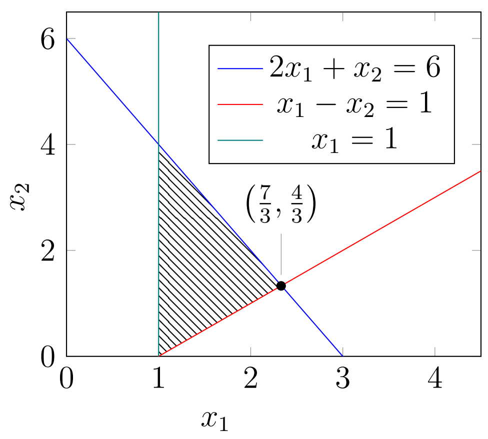
$$x_2 = \frac{1}{x_1} \longrightarrow \widehat{f_0}(x) = x_1 + \frac{2}{x_1} \longrightarrow \nabla\widehat{f_0}(x) = 1 - \frac{2}{x_1^2} = 0 \longrightarrow x_1 = \sqrt{2}, x_2 = \frac{1}{\sqrt{2}}$$

Тогда сам минимум целевой функции равен $f_0\left(\sqrt{2}, \frac{1}{\sqrt{2}}\right) = 2\sqrt{2}$.

Задача 4. Для начала найдём градиент:

$$f_0(x) = -\frac{x_1^2}{8} - \frac{x_2^2}{3} \longrightarrow \nabla f_0(x) = \left(-\frac{x_1}{4}, -\frac{2x_2}{3}\right)$$

Теперь изобразим область, в которой ищется минимум:



Заметим, что начальная точка $x_0 = (2, 3)$ не принадлежит данной области (хотя, согласно условиям метода проекции градиента, должна, спишем на некорректное условие), поэтому спроецируем её на ближайшую (синюю) грань – $x_0 = (1.6, 2.8)$.

Положим шаг $\alpha = 1$ постоянным и сделаем первый шаг:

$$x_1 = \pi_X\left(x_0 - \alpha\nabla f(x_0)\right) = \pi_X\left(\left(2,\ 4\,\frac{2}{3}\right)\right) = (1, 4)$$

Второй шаг:

$$x_2 = \pi_X\left(x_1 - \alpha\nabla f(x_1)\right) = \pi_X\left(\left(\frac{5}{4}, \frac{20}{3}\right)\right) = (1, 4)$$

Мы сошлись к точке $x^* = (1, 4)$, минимум целевой функции равен $\dfrac{-131}{24}$.

(Все проекции были найдены при помощи проведения перпендикуляров и нахождения ближайших вершин в случае отсутствия перпендикуляров. Данные вычисления тривиальны и очевидны, поэтому они были опущены.)

Задача 5. Введём следующие обозначения:

$$f_0(x) = 4x_1^2 + 3x_2^2 - 4x_1x_2 + x_1$$

Найдём градиент:

$$\nabla f_0(x) = (8x_1 - 4x_2 + 1, 6x_2 - 4x_1)$$

Шаг на каждой итерации будем находить одномерной оптимизацией:

$$\alpha_k = \arg\min_{\alpha > 0} f(x_k - \alpha\nabla f(x_k))$$

Имплементация самого поиска минимума приведена в приложенном файле task5.py.

В конечном итоге имеем

```
==== step 32 ====
Found minimum:
x* = [-0.1875 -0.125], f0(x*) = -0.09375
```

Задача 6. Вычислим градиент:

$$\nabla f(x) = \frac{1}{2(1 + x^T x)^2} 2x = 4f(x)^2 x$$

И гессиан:

$$\nabla^2 f(x) = 4\,\nabla\left[f(x)^2 x\right] = \frac{1}{(1 + x^T x)^2}\left[I_n - \frac{4xx^T}{1 + x^T x}\right]$$

Найдём обратную к гессиану матрицу:

$$[\nabla^2 f(x)]^{-1} = (1 + x^T x)^2\left[I_n - \frac{4xx^T}{1 + x^T x}\right]^{-1}$$

Положим $y = -\dfrac{4x}{1 + x^T x}$ и воспользуемся формулой Шэрмана-Моррисона:

$$\begin{aligned}[\nabla^2 f(x)]^{-1} &= (1 + x^T x)^2\,\left[I_n + xy^T\right]^{-1} = (1 + x^T x)^2\left[I_n - \frac{xy^T}{1 + x^T y}\right] \\ &= (1 + x^T x)^2\left[I_n + \frac{4xx^T}{1 - 3x^T x}\right] = \frac{1}{4f(x)^2}\left[I_n + \frac{4xx^T}{1 - 3x^T x}\right]\end{aligned}$$

Теперь рассмотрим метод наискорейшего спуска:

$$\alpha_k = \operatorname*{argmin}_{\alpha \geqslant 0} f(x_k - \alpha \nabla f(x_k))$$
$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) = x_k - 4\,\alpha_k f(x_k)^2 x_k = x_k(1 - 4\,\alpha_k f(x_k)^2)$$

В данной задаче множество $B = \{x \mid f(x) \leqslant f(x_0)\}$ ограничено (так как является шаром), откуда следует, что метод всегда сходится к точке минимума (по теореме 4 из семинаров 5-6). Иными словами, данный метод не нуждается в ограничениях на x_0.

Очередь метода Ньютона (помним, что минимум достигается в нуле):

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) = x_k - \left[I_n + \frac{4x_k x_k^T}{1 - 3x_k^T x_k} \right] x_k = \underbrace{\frac{4x_k^T x_k}{3x_k^T x_k - 1}}_{\beta} x_k$$

Откуда выходит, что $\|x_{k+p}\| = |\beta|^p \|x_k\|$. Если x_k - точка минимума, то $|\beta|$ может принимать значение 1, иначе должно выполняться неравенство $|\beta| < 1$ (чтобы норма убывала):

$$\left| \frac{4x_k^T x_k}{3x_k^T x_k - 1} \right| < 1 \longrightarrow x_k^T x_k < \frac{1}{7} \longrightarrow \|x_0\| < \frac{1}{\sqrt{7}}$$

Таким образом, в данной задаче метод Ньютона предъявляет требование на максимально возможное расстояние до начальной точки.

Задача 7. Выразим матрицу и сдвиг квадратичной формы:

$$A = 2 \begin{pmatrix} 1 & -1/2 \\ -1/2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 2 \end{pmatrix}$$

Найдём градиент целевой функции:

$$\nabla f(x) = Ax - b = (2x_1 - x_2 + 2, -x_1 + 2x_2 - 2)$$

Вычислим базисный вектор p_1:

$$p_1 = r_1 = -\nabla f(x_0) = b - Ax_0 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

Теперь найдём p_2:

$$\alpha_1 = \frac{\langle \nabla f(x_0), \nabla f(x_0) \rangle}{\langle A \nabla f(x_0), \nabla f(x_0) \rangle} = \frac{1}{2} \longrightarrow x_1 = x_0 + \alpha_1 p_1 = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}$$
$$r_2 = r_1 - \alpha_1 A p_1 = \begin{pmatrix} 0 \\ -1/2 \end{pmatrix} \longrightarrow \beta_1 = \frac{\langle r_2, r_2 \rangle}{\langle r_1, r_1 \rangle} = \frac{1}{4} \longrightarrow p_2 = r_2 + \beta_1 p_1 = \begin{pmatrix} -1/4 \\ -1/2 \end{pmatrix}$$

И вычислим x_2:

$$\alpha_2 = \frac{\langle r_2, r_2 \rangle}{\langle Ap_2, p_2 \rangle} = \frac{2}{3} \longrightarrow x_2 = x_1 + \alpha_2 p_2 = \begin{pmatrix} -2/3 \\ 2/3 \end{pmatrix}$$

В данной задачи размерность пространства X равна 2, следовательно, x_2 – искомая точка минимума. Значение целевой функции: $f(x_2) = -\dfrac{4}{3}$.