Квадратичная функция на единичной сфере
Дополнения
Лукашевич Александр
МФТИ 476

1. Движение по дуге большого круга

$$
\begin{aligned}
x^{k+1} &= P_S(x^k - \gamma_k \nabla f(x^k)) = \frac{x^k - \gamma_k \nabla f(x^k)}{\|x^k - \gamma_k \nabla f(x^k)\|} \\
\gamma_k &= \underset{\gamma \geq 0}{argmin} f(P_S(x^k - \gamma \nabla f(x^k))) = \underset{\gamma \geq 0}{argmin} f\left(\frac{x^k - \gamma \nabla f(x^k))}{\|x^k - \gamma \nabla f(x^k))\|}\right).
\end{aligned} \tag{1.1}
$$

Учтем вид функции: $f(x) = \frac{1}{2} x^T A x$ при расчете шага:

$$\gamma_k = \underset{\gamma \geq 0}{argmin}\ \frac{1}{2} \left(\frac{x^k - \gamma A x^k}{\|x^k - \gamma A x^k\|}\right)^T A \frac{x^k - \gamma A x^k}{\|x^k - \gamma A x^k\|} = \tag{1.2}$$

$$= \underset{\gamma \geq 0}{argmin}\ \frac{1}{2}\ \frac{(x^k)^T A x^k - 2\gamma (x^k)^T A^2 x^k + \gamma^2 (x^k)^T A^3 x^k}{< x^k - \gamma A x^k, x^k - \gamma A x^k >} = \tag{1.3}$$

$$= \underset{\gamma \geq 0}{argmin}\ \frac{1}{2}\ \frac{(x^k)^T A x^k - 2\gamma (x^k)^T A^2 x^k + \gamma^2 (x^k)^T A^3 x^k}{(x^k - \gamma A x^k)^T E (x^k - \gamma A x^k)} = \tag{1.4}$$

$$= \underset{\gamma \geq 0}{argmin}\ \frac{1}{2}\ \frac{(x^k)^T A x^k - 2\gamma (x^k)^T A^2 x^k + \gamma^2 (x^k)^T A^3 x^k}{(x^k)^T x^k - 2\gamma (x^k)^T A x^k + \gamma^2 (x^k)^T A^2 x^k} = \tag{1.5}$$

$$= \underset{\gamma \geq 0}{argmin}\ \frac{1}{2}\ F(x^k, \gamma). \tag{1.6}$$

Поскольку в знаменателе стоит квадрат нормы вектора $x^k - \gamma A x^k$, то, вводя обозначения $a_i(x^k) = (x^k)^T A^i x^k,\ i = 0, \dots 3$, знаменатель примет следующий вид:

$$a_2(x^k)\gamma^2 - 2\gamma a_1(x^k) + a_0(x^k) > 0. \tag{1.7}$$

Этот трехчлен должен быть строго больше нуля, чтобы сохранить смысл выражения.

Выясним, что нужно для того, чтобы (1.7) выполнялось. Для этого действительных корней не должно быть, а коэффициент при γ^2 должен быть больше нуля:

$$
\begin{aligned}
&d = (2a_1(x^k))^2 - a_2(x^k) a_0(x^k) < 0, \\
&a_2(x^k) > 0.
\end{aligned} \tag{1.8}
$$

Заметим, что последнее неравенство означает, что A^2 положительно определена.

Согласно необходимому условию экстремума, γ_k должно быть решением уравнения $\nabla_\gamma F(x^k, \gamma) = 0$:

$$\left(\frac{\gamma^2 a_3 - 2\gamma a_2 + a_1}{a_2 \gamma^2 - 2\gamma a_1 + a_0}\right)' = 2\frac{\gamma^2(a_2^2 - a_3 a_1) + \gamma(a_3 a_0 - a_2 a_1) + \overset{<0,(1.8)}{(a_1^2 - a_0 a_2)}}{(a_2\gamma^2 - 2\gamma a_1 + a_0)^2} = 0 \tag{1.9}$$

$$\gamma_k = \frac{(a_2 a_1 - a_3 a_0) \pm \sqrt{(a_2 a_1 - a_3 a_0)^2 - (a_2^2 - a_3 a_1)(a_1^2 - a_0 a_2)}}{2(a_2^2 - a_3 a_1)}. \tag{1.10}$$

В случае, когда $(a_2^2 - a_3 a_1) = 0$, получим одно решение

$$\gamma = \frac{a_1^2 - a_0 a_2}{a_3 a_0 - a_2 a_1}$$

Заметим, что первый множитель в (1.10) во втором слагаемом под корнем есть произведение детерминантов квадратных уравнений в (1.6). Итак, получены две стационарные точки. Теперь нужно выяснить какая из них даёт минимум. Для этого нужно определить каков знак второй производной в каждой из стационарных точек. Точка с положительной второй производной будет точкой минимума:

$$F'' = \frac{\gamma^3(-2a_2\widehat{a_2}) + \gamma^2(-3a_2\widehat{a_1}) + \gamma(2\widehat{a_2}a_0 - 4a_2\widehat{a_0} + 2a_1\widehat{a_1}) + \widehat{a_1}a_0 + 4a_1\widehat{a_0}}{(\gamma^2 a_2 - 2\gamma a_1 + a_0)^3}, \tag{1.11}$$

где $\widehat{a_2} = a_2^2 - a_3 a_1$, $\widehat{a_1} = a_3 a_0 - a_2 a_1$, $\widehat{a_0} = a_1^2 - a_0 a_2$ Можно указать промежутки положительности F'' используя, например, формулу Кардано (напомним, что знаменатель положителен, см. (1.7), поэтому рассматриваем только числитель):

- Сперва приведем числитель к каноническому виду $y^3 + py + q$:

  Сделаем замену $\gamma = y - \frac{b}{3a} = y - \frac{3a_2\widehat{a_1}}{6a_2\widehat{a_2}}$.

Тогда

$$p = \frac{6a_2\widehat{a_2}(2\widehat{a_2}a_0 - 4a_2\widehat{a_0} + 2a_1\widehat{a_1}) - (3a_2\widehat{a_1})}{(6a_2\widehat{a_2})}$$
$$q = \frac{2(-3a_2\widehat{a_1})^3 - 9(-2a_2\widehat{a_2})(-3a_2\widehat{a_1})(2\widehat{a_2}a_0 - 4a_2\widehat{a_0} + 2a_1\widehat{a_1}) + 27(-2a_2\widehat{a_2})^2(\widehat{a_1}a_0 + 4a_1\widehat{a_0})}{27(-2a_2\widehat{a_2})^3}$$

- Вычислим $Q = \left(\frac{p}{3}\right)^3 + \left(\frac{q}{2}\right)^2$ и определим его знак:

  $Q > 0$: один вещественный корень и два сопряженных комплексных корня,

  $Q = 0$: один однократный вещественный корень и один двукратный, или, если p = q = 0, то один трёхкратный вещественный корень,

  $Q < 0$ три вещественных корня.

- Введем обозначения:

$$\alpha = \sqrt[3]{-\frac{q}{2} + \sqrt{Q}}$$
$$\beta = \sqrt[3]{-\frac{q}{2} - \sqrt{Q}}$$

  Тогда корни уравнения выражаются следующим образом:

$$\begin{aligned} y_1 &= \alpha + \beta, \\ y_{2,3} &= -\frac{\alpha+\beta}{2} \pm i\frac{\alpha-\beta}{2}\sqrt{3}. \end{aligned} \tag{1.12}$$

  Переходя к исходной переменной:

$$\begin{aligned} \gamma_1 &= \alpha + \beta - \frac{3a_2\widehat{a_1}}{6a_2\widehat{a_2}}, \\ \gamma_{2,3} &= -\frac{\alpha+\beta}{2} \pm i\frac{\alpha-\beta}{2}\sqrt{3} - \frac{3a_2\widehat{a_1}}{6a_2\widehat{a_2}}. \end{aligned} \tag{1.13}$$

Считая, что имеет место случай $Q < 0$, а так же считая, что корни упорядочены по возрастанию, промежутки выпуклости будут следующими: $(\gamma_1,\ \gamma_2) \cup (\gamma_3, +\inf)$.

Итак, найдя стационарные точки, нужно будет проверить их на принадлежноть вышеуказанным интервалам.

Возможно, дабы не погрязть в комплексных корнях, проще будет просто подставить найденные стационарные точки в (1.11) и посмотреть знак выражения.