Методи оптимізації. Лекція 13.05.2022

Метод проекції градієнта

Постановка задачі:

$$f(x) \to min,\ x \in X \subseteq E^n \tag{1}$$

Проекцією точки a на множину X називається точка $y = P_X(a) \in X$, яка є найближча до точки a.

1. Проекція точки a на множину ***існує***, якщо множина X є ***замкнута*** множина.
2. Проекція точки a на множину X буде ***єдиною***, якщо множина X – ***опукла***.

Отже, знаходження проекції точки на множину еквівалентне такій задачі:

$$\|a - y\|^2 \to min(inf),\ y \in X. \tag{2}$$

Нехай $f(x) \in C^1(X)$, де X – опукла замкнута множина.

У методі **проекції градієнта** за наступну точку наближення вибирається проекція на множину X тієї точки, яка виходить за градієнтним методом.

Ітераційна формула методу проекції градієнта для розв'язання задачі (1) має вигляд:

$$x^{(k+1)} = P_X\left(x^{(k)} - \alpha_k f'\left(x^{(k)}\right)\right),\ k = 0, 1, \ldots \tag{3}$$

Отже, метод проекції градієнта є узагальненням градієнтних методів на випадок задачі умовної оптимізації.

Способи вибору крокового множника α_k є аналогічні відповідним способам для градієнтних методів. А саме:

1) довжина кроку α_k вибирається з умови:

$$g_k(\alpha_k) = \min_{\alpha \ge 0} g_k(\alpha),\ g_k(\alpha) = f\left(P_X\left(x^{(k)} - \alpha_k f'\left(x^{(k)}\right)\right)\right),$$

2) дробленням кроку: перевіряється умова монотонності:

$$f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right),$$

3) автоматичний вибір кроку за умовою:

$$f\left(x^{(k+1)}\right) - f\left(x^{(k)}\right) \le -\sigma\alpha_k \left\|x^{(k+1)} - x^{(k)}\right\|^2,$$

4) якщо $f(x) \in C^1(X)$ і константа Ліпшиця L для градієнту $f'(x)$ відома, то кроковий множник α_k може вибиратися за таким правилом:

$$0 < \varepsilon_0 \le \alpha_k \le \frac{2}{L + 2\varepsilon_1}, \tag{4}$$

де $\varepsilon_1 > 0$, ε_0, ε_1 – параметри методу.

5) апріорний вибір крокового множника

$$\alpha_k > 0, \sum_{k=0}^{\infty} \alpha_k = \infty, \ \sum_{k=0}^{\infty} \alpha_k^2 < \infty, \alpha_k \to 0, k \to \infty .$$

Теорема про збіжність методу проекції градієнта (3), (4). Нехай:

1) $f(x) \in C^1(X)$, X – опукла замкнута множина,
2) $f(x)$ обмежена знизу на X ,
3) $f(x)$ така, що її градієнт задовольняє умові Ліпшиця:

$$\| f'(x') - f'(x'') \| \le L \| x' - x'' \|, \forall x', x'' \in X .$$

Тоді послідовність $\{x_k\}$, що генерується за формулою (3) така, що

$$\left\| x^{(k+1)} - x^{(k)} \right\| \to 0, k \to \infty .$$

Якщо, крім того, ввести умови:

4) множина $M\left(x^{(0)}\right) = \left\{ x \in X : f(x) \le f\left(x^{(0)}\right) \right\}$ – є обмежена.

Тоді послідовність $\{x_k\}$ збігається до множини стаціонарних точок S_* функції $f(x)$:

$$\rho\left(x^{(k)}, S_*\right) \to 0, k \to \infty .$$

5) якщо $f(x)$ – опукла функція на множині X .

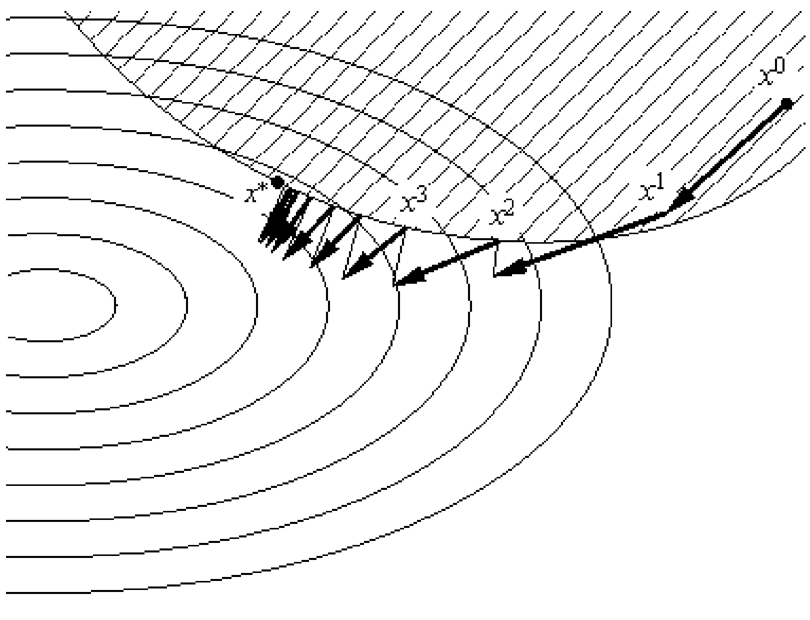
тоді послідовність $\{x_k\}$ є мінімізуючою послідовністю:

$$\lim_{k \to \infty} f\left(x^{(k)}\right) = f_*, \quad \rho\left(x^{(k)}, X_*\right) \to 0, k \to \infty .$$

Також справедлива така оцінка за функцією:

$$f\left(x^{(k)}\right) - f_* \le \frac{c}{k}, \quad k = 1, 2, ..., \quad c = const .$$

Геометрична ілюстрація методу проекції градієнта



Множини простої структури

У загальному випадку задача (2) за складністю еквівалентна задачі (1). Тому має сенс застосовувати метод (3) тільки у випадку, коли проекція точки на множину легко знаходиться.

Існують множини, які називаються ***множинами простої структури***, для яких проекцію можна виписати явно.

Нехай $a=(a_1,\ldots,a_n)\in E^n$, тоді

- ***невід'ємний октант*** $X=\left\{\, x\in E^n : x_i\ge 0,\quad i=\overline{1,n} \,\right\}$.

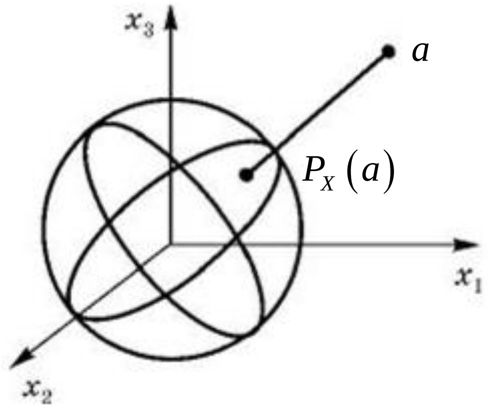
Проекцією точки a на множину X буде точка

$$P_X(a)=\left(max(0,a_1),\ldots,max(0,a_n)\right).$$

- ***куля у n–вимірному просторі*** $X=\left\{\, x\in E^n : \|x-x_c\|\le R \,\right\}$, де x_c – центр, R – радіус кулі.

Проекцією точки $a\notin X$ (точка не належить кулі) на множину X буде точка

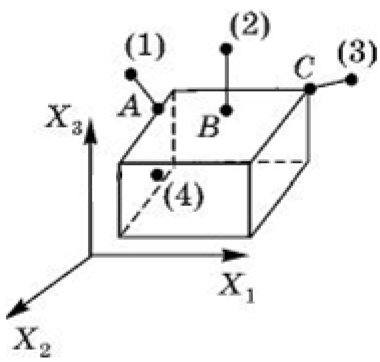
$$P_X(a)=x_c+\frac{a-x_c}{\|a-x_c\|}R.$$



- ***n–вимірний паралелепіпед*** $X=\left\{\, x\in E^n : b_j\le x_j\le c_j,\ j=\overline{1,n} \,\right\}$.

Проекцією точки a на множину X буде точка

$$\left(P_X(a)\right)_j=\begin{cases} b_j, & a_j<b_j, \\ a_j, & b_j\le a_j\le c_j, \\ c_j, & a_j>c_j, \end{cases}$$



- ***n*–вимірний півпростір** $X = \{ x \in E^n : (p, x) \geq \beta \}$, p – вектор, β – число.

Проекцією точки a на множину X буде точка

$$P_X(a) = a + max\left(0, \beta - (p, x)\right) \frac{p}{\|p\|^2}.$$

Модифікація методу проекції градієнта

Існує модифікація методу проекції градієнта, коли наступне наближення знаходиться за формулою:

$$x^{(k+1)} = \beta_k P_X\left(x^{(k)} - \alpha_k f'\left(x^{(k)}\right)\right) + \left(1 - \beta_k\right) x^{(k)},\ 0 < \beta_k < 1. \qquad (5)$$

Зауважимо, що, використовуючи ідею проектування, можна модифікувати стосовно до задач умовної оптимізації й інші методи безумовної оптимізації, в тому числі метод Ньютона, метод спряжених градієнтів.

Метод проекції Ньютона

$$x^{(k+1)} = P_X\left(x^{(k)} - \left[f''\left(x^{(k)}\right)\right]^{-1} f'\left(x^{(k)}\right)\right),\ k = 0,1,\ldots \qquad (6)$$

або

$$x^{(k+1)} = P_X\left(x^{(k)} - \alpha_k \left[f''\left(x^{(k)}\right)\right]^{-1} f'\left(x^{(k)}\right)\right),\ k = 0,1,\ldots, \qquad (7)$$

Метод проекції спряжених градієнтів

Ітераційна формула методу проекції спряжених градієнтів:

$$x^{(k+1)} = P_X\left(x^{(k)} - \alpha_k p^{(k)}\right),\ k = 0,1,\ldots \qquad (8)$$

де

$$p^{(k)} = \begin{cases} f'\left(x^{(k)}\right), k = 0, \\ f'\left(x^{(k)}\right) - \beta_k p^{(k-1)},\ k = 1,2,.. \end{cases}$$

$$\beta_k = \begin{cases} \dfrac{\left(f'\left(x^{(k)}\right), f'\left(x^{(k-1)}\right) - f'\left(x^{(k)}\right)\right)}{\left\|f'\left(x^{(k-1)}\right)\right\|^2}, k \in I_1 \\ 0, k \in I_2 \end{cases}$$

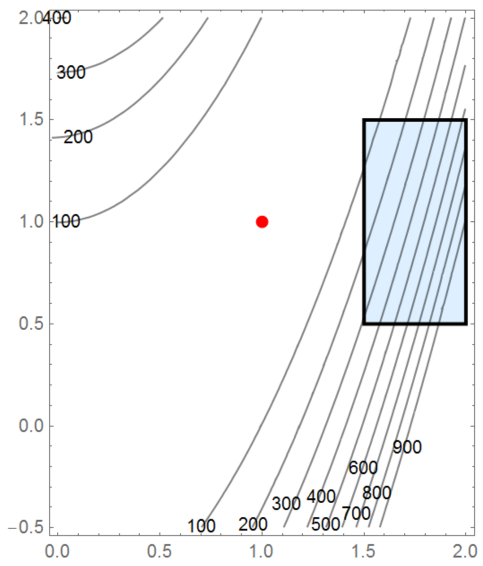
$$\alpha_k: \qquad f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right).$$

I_1, I_2 – множина індексів, $I_1 \cup I_2 = \{0,1,2,\ldots\}$, $I_1 \cap I_2 = \varnothing$.

Приклад. Розв'язати методом проекції градієнта для функції Розенброка:

$$f(x) = 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2 \to min,\ x \in X,$$

$$X = \left\{\left(x_1, x_2\right):\ 1.5 \le x_1 \le 2,\ 0.5 \le x_2 \le 1.5\right\}.$$



На рис. наведено лінії рівня функції Розенброка, точку мінімуму $(1,1)$ (червоним кольором), область X. Видно, що точка $(1,1)$ не належить області X.

Градієнт функції Розенброка обчислюється за формулами

$$f'(x) = \begin{pmatrix} -400x_1\left(x_2 - x_1^2\right) - 2\left(1 - x_1\right) \\ 200\left(x_2 - x_1^2\right) \end{pmatrix}.$$

Виберемо початкове наближення $x^{(0)} = (1.8,\ 1.3)$, $f\left(x^{(0)}\right) = 377$.

Зауважимо, що початкове наближення належить області X. Нехай константа зупинки ітераційного процесу $\varepsilon = 0.01$.

Ітерація 1. Ітераційна формула методу проекції градієнта для розв'язання задачі має вигляд:

$$x^{(k+1)} = P_X\left(x^{(k)} - \alpha_k f'\left(x^{(k)}\right)\right),\quad k = 0,1,\ldots$$

Кроковий множник α_k будемо визначати за умовою монотонності

$$f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right).$$

Знайдемо $x^{(1)} = P_X\left(x^{(0)} - \alpha_0 f'\left(x^{(0)}\right)\right)$. Нехай $\alpha_0 = 0.002$, тоді

$$x^{(0)} - \alpha_0 f'\left(x^{(0)}\right) = \begin{pmatrix} 1.8 \\ 1.3 \end{pmatrix} - 0.002 \cdot \begin{pmatrix} 1398.4 \\ -388 \end{pmatrix} \approx \begin{pmatrix} -0.9968 \\ 2.076 \end{pmatrix}.$$

Точка $\begin{pmatrix} -0.9968 \\ 2.076 \end{pmatrix}$ виходить за межі області X. Так як множина X – координатний паралелепіпед, то можна вказати явний вигляд проекції точки на множину.

Проекцією точки a на множину X буде точка

$$\left(P_X(a)\right)_j = \begin{cases} b_j, & a_j < b_j, \\ a_j, & b_j \le a_j \le c_j, \\ c_j, & a_j > c_j. \end{cases}$$

Для заданої області маємо:

$$\left(P_X(a)\right)_1 = \begin{cases} 1.5, & a_1 < 1.5, \\ a_j, & 1.5 \le a_1 \le 2, \\ 2, & a_1 > 2, \end{cases} \qquad \left(P_X(a)\right)_2 = \begin{cases} 0.5, & a_2 < 0.5, \\ a_j, & 0.5 \le a_2 \le 1.5, \\ 1{,}5, & a_2 > 1.5, \end{cases}$$

$$x^{(1)} = P_X\begin{pmatrix} -0.9968 \\ 2.076 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix},\ f\left(x^{(1)}\right) = 56.5.$$

Так як $f\left(x^{(1)}\right) < f\left(x^{(0)}\right)$, то перевіряємо критерій зупинки:

$$\left\| x^{(1)} - x^{(0)} \right\| = 0.360555 > 0.01.$$

Критерій зупинки не виконується, значить, переходимо до наступної ітерації.

Ітерація 2. Визначаємо $x^{(2)} = P_X\left(x^{(1)} - \alpha_1 f'\left(x^{(1)}\right)\right)$. Нехай $\alpha_1 = 0.002$, тоді

$$x^{(1)} - \alpha_1 f'\left(x^{(1)}\right) = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix} - 0.002 \cdot \begin{pmatrix} 451 \\ -150 \end{pmatrix} = \begin{pmatrix} 0.898 \\ 1.6 \end{pmatrix}.$$

Точка $x^{(1)} - \alpha_1 f'\left(x^{(1)}\right)$ виходить за межі області X.

$$x^{(2)} = P_X\begin{pmatrix} 0.898 \\ 1.6 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 1.5 \end{pmatrix},\ x^{(2)} = x^{(1)}.$$

$$\left\| x^{(2)} - x^{(1)} \right\| = 0.$$

Критерій зупинки виконується, значить, процес обчислень закінчено. На рис. зображено траєкторію руху до розв’язку для заданої задачі умовної оптимізації. Через $\tilde{x}^{(1)}$ та $\tilde{x}^{(2)}$ позначено точки, які не належать допустимій області, а $x^{(1)}$ та $x^{(2)}$ їх проекції відповідно.