1

Постановка задачи безусловной оптимизации

Пусть $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \Omega$ – вектор параметров оптимизации

$\Omega \in \mathbb{R}^n$ – множество допустимых решений, $n \geq 1$

$f(\boldsymbol{x})$ – целевая функция, $f : \mathbb{R}^n \to \mathbb{R}$

Тогда задачу безусловной оптимизации запишем следующим образом:

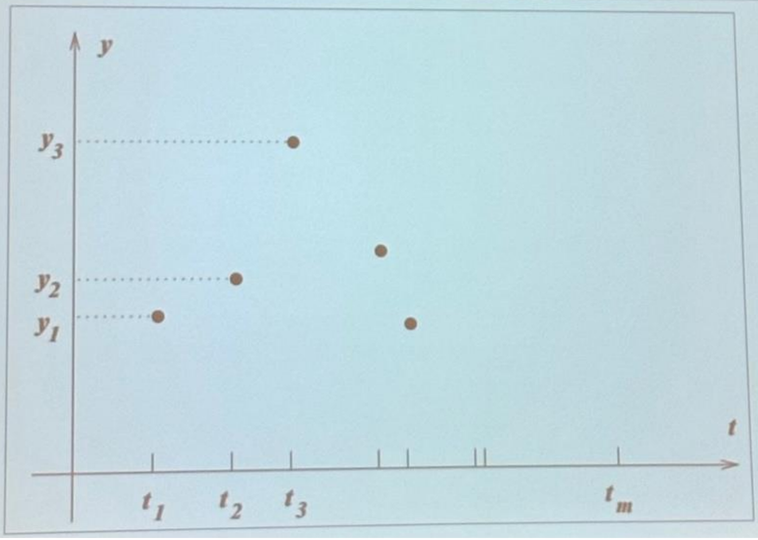
$$\min_{x \in \Omega} f(\boldsymbol{x})$$

2

Пример

Практическая задача безусловной оптимизации

Задача: имеется набор экспериментальных данных $(t_i, y_i),\ i = 1 \ldots m$. *Необходимо аппроксимировать геометрическое место точек исходных данных некоторой кривой* $f(\boldsymbol{x})$



3

Пример

Практическая задача безусловной оптимизации

Пусть из характера эксперимента известно, что данные могут соответствовать модели, заданной следующей функцией:

$$\phi(t;x) = x_1 + x_2 e^{-(x_3-t)^2/x_4} + x_5 \cos(x_6 t),$$

где $x_1, x_2, \ldots, x_6$ *- параметры модели.*

Тогда расхождение экспериментально полученных данных $(t_j, y_j),\ i = 1 \ldots m$ *со значением модели, полученным от функции* $\phi(t;x)$ *при заданных параметрах* $x_1, x_2, \ldots, x_6$, *может быть выражено как:*

$$r_j(x) = y_j - \phi(t_j;x),\ j = 1 \ldots m$$

Следовательно, для решения задачи оптимизации необходимо минимизировать значения расхождений $r_j(x)$, *а именно – найти минимум целевой функции* $f(x) = r_1^2(x) + \ldots + r_m^2(x)$

Таким образом, задача безусловной оптимизации имеет вид:

$$\min_{x \in R^6} \left(r_1^2(x) + \ldots + r_m^2(x) \right)$$

В данном примере рассмотрена нелинейная задача наименьших квадратов (nonlinear least-squares problem)

5

Подходы итеративному поиску оптимального решения

Поиск оптимального решения начинается с выбранного начального значения x_0. *Выбор начального значения, как правило, основан на заведомо известной информации о поведении целевой функции или о положении её минимума.*

Начиная со значения x_0, *алгоритм оптимизации генерирует последовательность итераций* $\{x_k\}_{k=0}^{\infty}$ *завершающуюся по достижении условия останова (будут разобраны позже). На каждой итерации алгоритм может использовать информацию о значении целевой функции на текущем и предыдущем(-их) шаге. На основе полученной информации алгоритм выполняет следующий шаг и выбирает следующее значение* x_{k+1}, *в котором целевая функция будет меньше, чем на предыдущем (кроме немонотонных методов).*

Существуют две стратегии выбора итеративного шага алгоритма оптимизации: стратегия линейного поиска и стратегия доверительной области.

7

Подходы итеративному поиску оптимального решения

Стратегия доверительной области

В рамках стратегии доверительной области информация, известная о целевой функции $f(x)$, используется для её аппроксимации с помощью некоторой функции $m_k(x)$, поведение которой рядом с точкой x_k схоже (совпадает) с целевой функцией. После аппроксимации, поиск минимума ведут уже для функции $m_k(x)$:

$$\min_p m_k(x_k + p),$$

Где $m_k(x)$, как правило, находят в виде:

$$m_k(x_k + p) = f_k + p^T \nabla f_k + \frac{1}{2} p^T B_k p$$

где $f_k, \nabla f_k$ – значения целевой функции и градиента целевой функции в точке x_k соответственно, B_k – гессиан (или некоторое приближение гессиана) целевой функции в точке x_k.

Однако функция $m_k(x)$ аппроксимирует целевую функцию достаточно точно лишь в некоторой окрестности точки x_k. По этой причине поиск минимума аппроксимированной функции ведут в некоторой доверительной области точки x_k. Как правило, доверительную область определяю как $\|p\|_2 \le \Delta$, где $\Delta > 0$ – радиус доверительной области.

9

Методы линейного поиска

Длина шага

При выборе длины шага оптимизации возникает конфликт интересов:

- Большая величина шага позволяет уменьшить количество обращений к целевой функции (снижение вычислительной сложности), но уменьшает точность алгоритма (можем «перескочить» минимум целевой функции).
- Малая величина шага обеспечивает бóльшую точность, но увеличивает вычислительную сложность и количество итераций. В частности, наиболее точным выбором станет решение локальной задачи минимизации функции от длины шага при фиксированных значениях x_k и p_k:

$$\phi(\alpha) = f_k\left(x_k + \alpha p_k\right), \ \alpha > 0$$

Однако такой подход иррационален с точки зрения вычислительных и временных ресурсов.

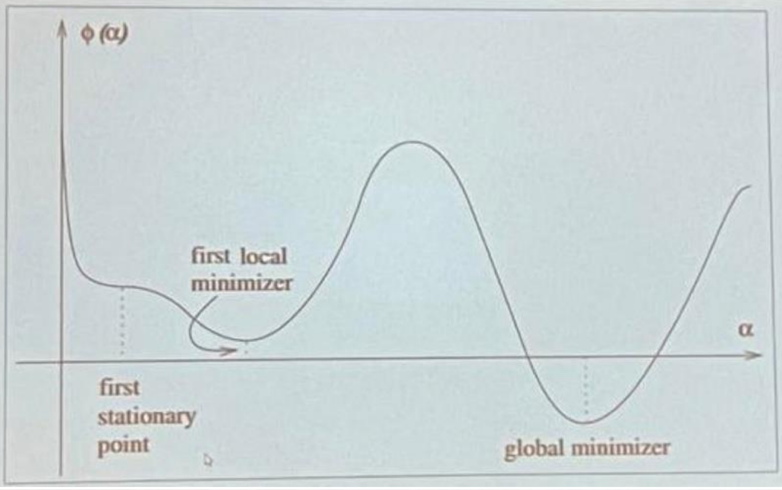
10

Методы линейного поиска

Длина шага. Пример

Оптимальный выбор длины шага оптимизации – глобальный минимум функции $\phi(\alpha)$ при фиксированных x_k и p_k:

$$\phi(\alpha)=f_k\left(x_k+\alpha p_k\right),\ \alpha>0$$



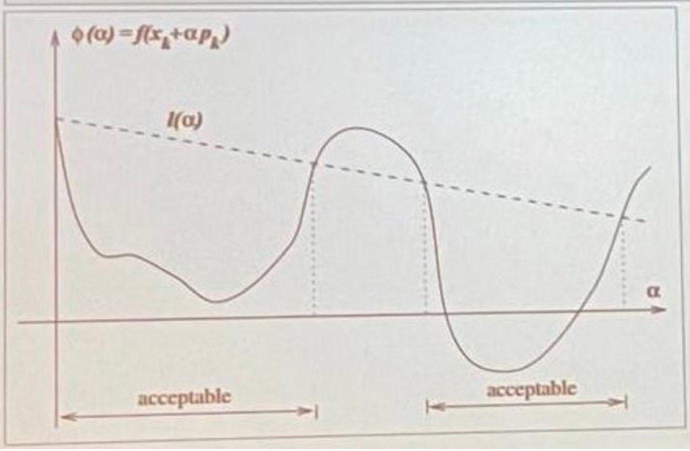
11

Методы линейного поиска

Длина шага. Условия Вольфе

Можно показать, что для того, чтобы длина шага α_k алгоритма оптимизации обеспечивала достаточное уменьшение функции, должно быть выполнено следующее условие:

$$f\left(x_k+\alpha p_k\right)\le \underbrace{f\left(x_k\right)+c_1\alpha\nabla f_k^T p_k}_{l(\alpha)}$$



Однако условия не обеспечивает гарантированный прогресс в поиске минимума, поскольку будет выполнено при любом достаточно малом значении шага α_k.

Необходимо добавить условие, исключающее дробление шага.