Методы оптимизации

Лектор: Андреев Юрий Александрович

_scarleteagle **imkochelorov** **smvfe**

asparagus_densiflorus

Многомерный навигатор

Метод Ньютона

или же "метод касательных"

Итеративный метод поиска корней уравнения $f(x) = 0$.
Также этим методом можно искать критические точки, задав $f := \varphi'$

В одномерном случае, итерация поиска нуля методом Ньютона достигается следующим правилом:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

Метод требует производную f, а в случае поиска экстремумов и вовсе φ'', но зато хорошо сходится.

Но мы можем обойти это требование, модифицировав метод Ньютона, заменив производную чем-то ещё: получим так называемые *квазиньютоновские методы*

Методы спуска

Пусть у нас есть $f'(x_k)$ (производная/градиент). Делаем шаг:

$$x_{k+1} = x_k - hf'(x_k) \ - \ \text{делаем шаг в сторону минимума}$$

$$x_{k+1} = x_k + hf'(x_k) \ - \text{делаем шаг в сторону максимума}$$

Эта h называется *learning rate* (темп обучения).

Шаг во многих случаях будет зависеть от итерации:

$$x_{k+1} = x_k - h(k)f'(x_k)$$

Направление есть, но как далеко по нему идти?

Для нахождения длины шага существуют правила *неточного поиска*:

Правило Армихо

Рассмотрим систему координат yO, где начало координат $O = x_k$.

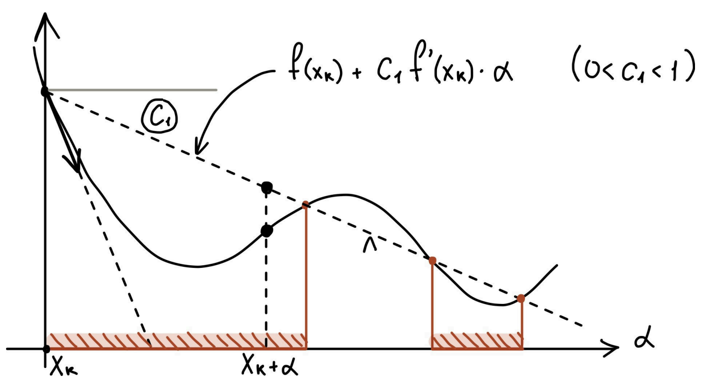
В начальном положении мы "стоим" в точке

$$(x_k, f(x_k))$$

Рассмотрим касательную через начальную точку

Уравнение касательной:

$$f(x_k) + \alpha f'(x_k)$$



Но мы берем не касательную, а прямую повыше:

$$l_a = f(x_k) + c_1 \alpha f'(x_k), \quad 0 < c_1 < 1$$

c_1 выбираем случайным образом

Нам нужно выбрать следующую точку в промежутках, где функция ниже прямой, желательно подальше

Заметим, что чем дальше мы пойдём, тем значение будет в промежутке с меньшей верхней границей

Как найти нужный α? Воспользуемся *backtracking*.

> **Метод backtracking**
>
> - Начинаем с заведомо большой $\alpha = \alpha_0$, проверяем условие в соответствующей точке $x_k + \alpha$. Можно, например, найти точку пересечения прямой и горизонтальной оси и взять соответствующую α_0
> - Если $f(x_k + \alpha_i)$ ниже, чем прямая l_a, тогда оставляем α_i
> - Если нет, то следующее значение α считаем так: $\alpha_{i+1} := q \cdot \alpha_i, \quad q \in (0; 1)$
>   Снова проверяем

Q: А что делать, если x_k попало в локальный максимум (или в седловую точку)?

A: Метод модифицируется добавлением импульса к шагу

Правило Вулфа

в других источниках также встречается "Вулфа" или "Вольфе"

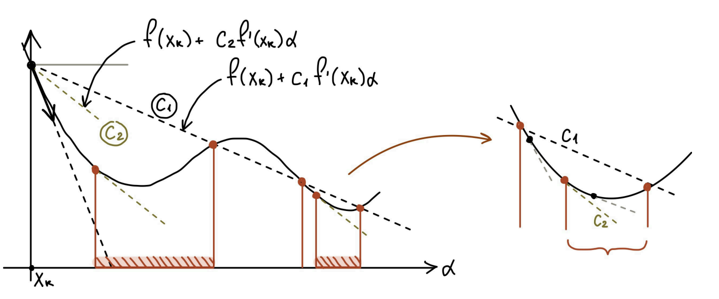
Зададим две константы:

$$0 < c_1 < c_2 < 1$$

Рассмотрим систему координат $yO\alpha$, где начало координат $O = x_k$.

В начальном положении мы "стоим" в точке $(x_k, f(x_k))$.

Рассмотрим касательную через начальную точку.



Уравнение касательной:

$$f(x_k) + f'(x_k)\alpha$$

Зададим две прямые:

$$l_{c_1} : f(x_k) + c_1 f'(x_k)\alpha$$

$$l_{c_2} : f(x_k) + c_2 f'(x_k)\alpha$$

Ищем точку, которая будет ниже прямой l_{c_1}, причём касательная в которой положе прямой l_{c_2}

Прямая l_{c_2} задаёт начальный наклон, относительно которого мы будем рассматривать наклоны касательной в промежутке полученном от прямой l_{c_1}.

Как это записать формулами?

Ищем такое α', что $f'(x_k + \alpha') \geq c_2 f'(x_k)$

Как мы ищем α'? Можно бинпоиском, можно ввести коэффициент q и увеличивать/уменьшать свои шаги

Правило Голдстейна

Зададим две константы:

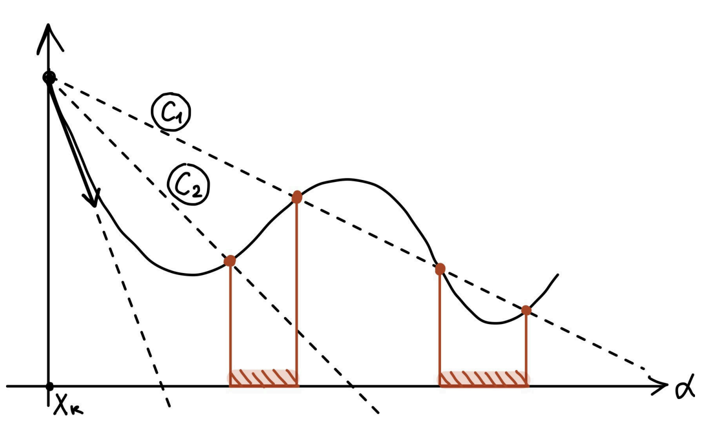
$$0 < c_1 < c_2 < 1$$

Аналогично предыдущим способам введём l_{c_1} и l_{c_2}

Давайте искать, где f лежит между прямыми:

Идем справа налево, чтобы выполнилась верхняя граница

Идем слева направо, чтобы выполнилась нижняя граница



Линейный поиск в $\mathbb{R}^n$

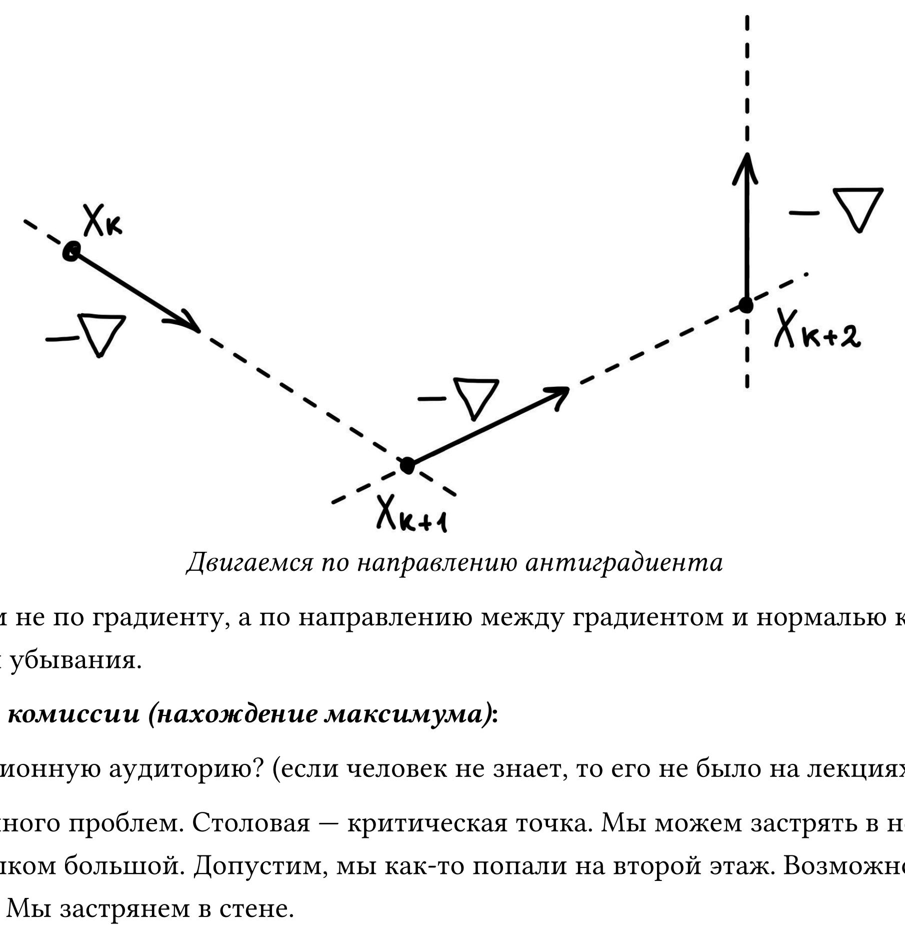
Применим то, чему мы научились, в многомерном случае:

Мы находимся в x_i, переходим в точку x_{i+1}

Куда идти?

- Покоординатный поиск: берём любую координату и меняем её так, чтобы значение стало меньше.
- По направлению убывания. Лучшее убывание — *антиградиент*.

Можно рассмотреть функцию в данном направлении: $g(t) = f(x_k) + tu$. Мы свели к одномерному случаю.



Двигаемся по направлению антиградиента

Вообще, можно идти не по градиенту, а по направлению между градиентом и нормалью к нему, это тоже будет направлением убывания.

Пример задания на комиссии (нахождение максимума):

- Как пройти в лекционную аудиторию? (если человек не знает, то его не было на лекциях)

Может быть очень много проблем. Столовая — критическая точка. Мы можем застрять в ней, если шаг с позиции входа слишком большой. Допустим, мы как-то попали на второй этаж. Возможно, мы попали в коридор (нечаянно). Мы застрянем в стене.

Можно учитывать в данном случае так называемый *"условный градиент"* — застряли в стене, идём в сторону, куда указывает тень (проекция) от градиента на стене.

Метод Ньютона помогает нам сразу подниматься на этаж вверх, а не по каждой лестнице отдельно. Квазиньютоновские методы бывают не точны и на 3 этаже: мы можем попасть в коворкинг или пойти в коридор.

Ещё у нас коридоры обычно ровные, без набора высоты. Поэтому там градиента нет. В этом случае можно использовать **метод импульса** — учитывать то, как мы двигались в прошлый раз (формально — следующее изменение будем принимать как линейную комбинацию градиента и предыдущего изменения). Так мы, кстати, можем бороться с максимумами.

Пример задания на комиссии по пожарной безопасности (нахождение минимума):

- Пройти из аудитории на улицу. Направление антиградиента указано на планах эвакуации.