1

Двойственная функция

Двойственная функция в точке x^* показывает, насколько минимум надо опустить прямую $\langle x, x^* \rangle$ вниз, чтобы она целиком оказалась ниже функции $f(x)$.

$$f^*(x^*) = \inf_t \{ t \mid \langle x, x^* \rangle - t \leq f(x), \text{ для любого } x \}$$

Заметим, что определение легко переформулировать как поиск максимальной разницы от функции $f(x)$ вверх до прямой $\langle x, x^* \rangle$. Если в какой-то точке x прямая $\langle x, x^* \rangle$ лежит очень высоко относительно функции $f(x)$, то и минимальная величина, на которую придется опустить прямую очень велика.

Двойственная функция

Двойственная функция в точке x^* показывает, насколько минимум надо опустить прямую $\langle x, x^* \rangle$ вниз, чтобы она целиком оказалась ниже функции $f(x)$.

$$f^*(x^*) = \sup_x \{\langle x, x^* \rangle - f(x)\}$$

Конечно, двойственная функция может принимать особое значение ∞.

Примеры.

Упражнение

Как изменится двойственная функция, если сдвинуть исходную функцию на вектор b вдоль аргументов?

Решение упражнения

Определим сдвинутую функцию $g(x) = f(x - b)$.
Находим $g^*(x^*)$:

$$g^*(x^*) = \sup_x \{\langle x, x^* \rangle - g(x)\} = \sup_x \{\langle x, x^* \rangle - f(x - b)\}$$

Прибавим и вычтем $\langle x, b \rangle$:

$$g^*(x^*) = \sup_x \{\langle x - b, x^* \rangle - f(x - b)\} + \langle b, x^* \rangle = \sup_w \{\langle w, x^* \rangle - f(w)\} + \langle b, x^* \rangle$$

Следовательно,

$$g^*(x^*) = f^*(x^*) + \langle b, x^* \rangle$$

Теоремка

$$f^{**}(x) \leq f(x)$$

Доказательство неравенства

Для доказательства теоремы предположим, что $f(x) \leq a$, и докажем, что в этом случае $f^{**}(x) \leq a$.
На старт:

$$f(x) \leq a$$

Превратим левую и правую часть во что-то похожее на определение двойственной функции, рассмотрев произвольные угловые коэффициенты x^*.

$$\langle x, x^* \rangle - f(x) \geq \langle x, x^* \rangle - a$$

Левая часть не упадёт, если мы в ней вместо конкретного x возьмём супремум:

$$\sup_w \langle w, x^* \rangle - f(w) \geq \langle x, x^* \rangle - a$$

Это в чистом виде определение $f^*(x^*)$:

$$f^*(x^*) \geq \langle x, x^* \rangle - a$$

Вспомним, что вектор наклонов x^* был произвольным!
Получается, что для любого вектора x^*:

$$a \geq \langle x, x^* \rangle - f^*(x^*)$$

Если a выше каждого возможного значения правой части, то a выше максимально возможного.

$$a \geq \sup_{x^*} \{ \langle x, x^* \rangle - f^*(x^*) \}$$

А это и есть искомое утверждение:

$$a \geq f^{**}(x).$$

Для хороших выпуклых функций, оказывается, что $f(x) = f^{**}(x)$.

Теоремка

Если ...., то

$$f^{**}(x) = f(x)$$

Доказательство случая равенства

В доказательстве теоремки о неравенстве есть только один неравносильный переход. Обратим на него внимание.

2

> **ℹ Правильное определение двойственной задачи**
>
> Рассмотрим задачу поиска минимума функции ϕ по x с параметром y.
>
> $$\min_x \phi(x, y)$$
>
> Параметр y мы воспринимаем как возмущение, отклонение задачи от некоторой исходной.
> Исходная задача без возмущения имеет вид:
>
> $$\min_x \phi(x, 0)$$
>
> Двойственной задачей назовём задачу:
>
> $$\min_{y^*} \phi^*(0, y^*)$$

Заметим, что двойственная задача определена только после того, как мы договорились, как вносить возмущения в исходную задачу. Для разного типа возмущений будут возникать разные двойственные задачи.

2.1

2.2

2.3

2.4

3

Определим функцию $h(y)$ как оптимум исходной задачи про возмущении y,

$$h(y) = \min_x \phi(x, y)$$

В этом случае исходную задачу без возмущения можно представить как вычисление $h(0)$.

…

Таким образом, двойственная задача — это вычисление $h^{**}(0)$.

Помимо ещё одной абстрактной формулировки:

Правильное определение двойственной задачи

Рассмотрим задачу поиска минимума функции ϕ по x с параметром y.

$$h(y) = \min_x \phi(x, y)$$

Прямая задача — это поиск $h(0)$.
Двойственная задача — это поиск $h^{**}(0)$.

Вспомним, что $h^{**}(x) \leq h(x)$ для всех функций и $h^{**}(x) = h(x)$ для хороших … функций.

Следовательно, мы бесплатно получаем две теоремки:

Теоремка

Если $h(y) = \min_x \phi(x, y)$ — оптимум возмущённой задачи, то

$$h^{**}(0) = \min_{y^*} \phi^*(0, y^*) \leq \min_x \phi(x, 0) = h(0).$$

Теоремка

Если $h(y) = \min_x \phi(x, y)$ — оптимум возмущённой задачи, и …, то

$$h^{**}(0) = \min_{y^*} \phi^*(0, y^*) = \min_x \phi(x, 0) = h(0).$$

Источники: