🌲 Выпуклые функции - 1

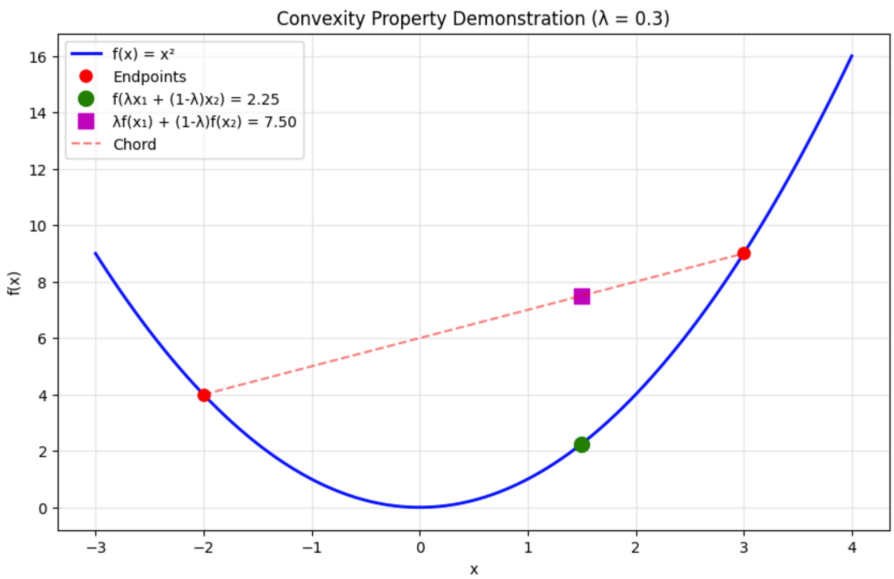
Определения, неравенство Йенсена, понятия эпиграфа и подуровня

Понятие выпуклой функции

Определение 1: Функция f(x), определенная на выпуклом множестве $S \subseteq \mathbb{R}^n$ **выпуклая**, если выполнено **неравенство Йенсена**:

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2)$$
$$\forall x_1, x_2 \in S, \lambda \in [0,1]$$

Если неравенство выполняется как строгое, $f(x)$ будет **строго выпукла** на S. Если неравенство выполняется в обратную сторону - функция называется вогнутой (выпуклой вниз).



Интуитивное нестрогое обоснование неравенства Йенсена: "геометрически" выпуклая - та функция, график которой лежит не выше любой своей хорды. То есть, левая часть неравенства - это значения функции (ie точка на самом графике) для некоторой точки $x_\lambda = \lambda x_1 + (1-\lambda)x_2$, лежащей между x_1 и x_2. Правая же часть неравенства принимает значения на хорде, которые, в случае выпуклой (ie выпуклой вниз) функции, будут больше чем значения на самой функции, как в левой части.

Если $f(x)$ выпукла на данном множестве, то любой её локальный минимум также является глобальным (возможно не единственным).

Определение 2: Функция $f(x)$ на множестве $S \subseteq \mathbb{R}^n$ называется **сильно выпуклой**, если при некотором $\theta > 0$:

$$f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2) - \frac{\theta}{2}\lambda(1-\lambda)\|x_1 - x_2\|^2$$
$$\forall x_1, x_2 \in S, \lambda \in [0,1]$$

Сильно выпуклая функция всегда строго выпуклая. Сильно выпуклая функция имеет единственный глобальный минимум.

Теорема 1 (Неравенство Йенсена в обобщённом виде): $f(x)$ - **выпуклая функция** на выпуклом множестве $S \subseteq \mathbb{R}^n$, если при $x_i \in S, 1 \le i \le m; \lambda = [\lambda_1, ..., \lambda_m]^T \in \Lambda^m$

$$f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \leq \sum_{i=1}^{m} \lambda_i f(x_i)$$

Где Λ^m — **вероятностный симплекс**, то есть множество вида:

$$\left\{\lambda \in \mathbb{R}_+^m : \sum_{i=1}^{m} \lambda_i = 1\right\}$$

Примеры выпуклых функций:
- $f(x) = x^p, p > 1, S = \mathbb{R}_{++}$
- $f(x) = -\ln x, S = \mathbb{R}_{++}$
- $f(x) = x \ln x, S = \mathbb{R}_{++}$
- $f(x) = e^{cx}, c \in \mathbb{R}, S = \mathbb{R}_{++}$

Прочие утверждения и свойства выпуклых функций:
1. (**Теорема**) Если $f_1(x), ...f_m(x)$ - выпуклые функции на S, $\alpha_i \geq 0, i \in [1, m]$, то будут выпуклыми функции:
   - $f(x) = \sum_{i=1}^{m} \alpha_i f_i(x)$
   - $f(x) = \max_i f_i(x)$
2. Пусть $g(y), h(x)$ - выпуклые функции на $\mathbb{R}, \mathbb{R}^n$ соответственно, тогда если $g(y)$ - неубывающая функция на $\mathbb{R} \Rightarrow f(x) = g(h(x))$ - тоже выпуклая на $\mathbb{R}^n$
3. Пусть $g(y)$ - выпуклая функция и невозрастающая функция на $S \subseteq \mathbb{R}$, а $h(x)$ — вогнутая функция на $\mathbb{R}^n$, тогда $f(x) = g(h(x))$ будет выпуклой на $\mathbb{R}^n$. Примером такой функции является $f(x) = -\ln(-h(x)), h(x) < 0$

Эпиграф(epigraph, надграфик)

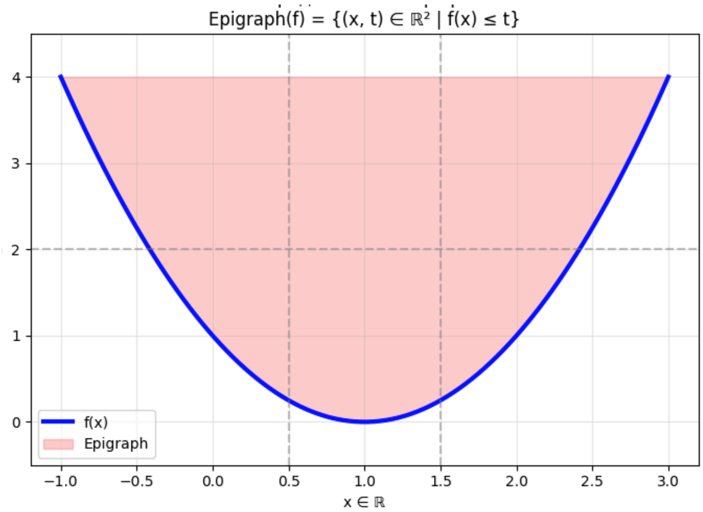
Определение 3: Эпиграфом функции $f(x)$ на $S \subseteq \mathbb{R}$ называется множество вида:

$$\operatorname{epi} f = \{[x, \mu] \in S \times \mathbb{R} : \mu \geq f(x)\}$$

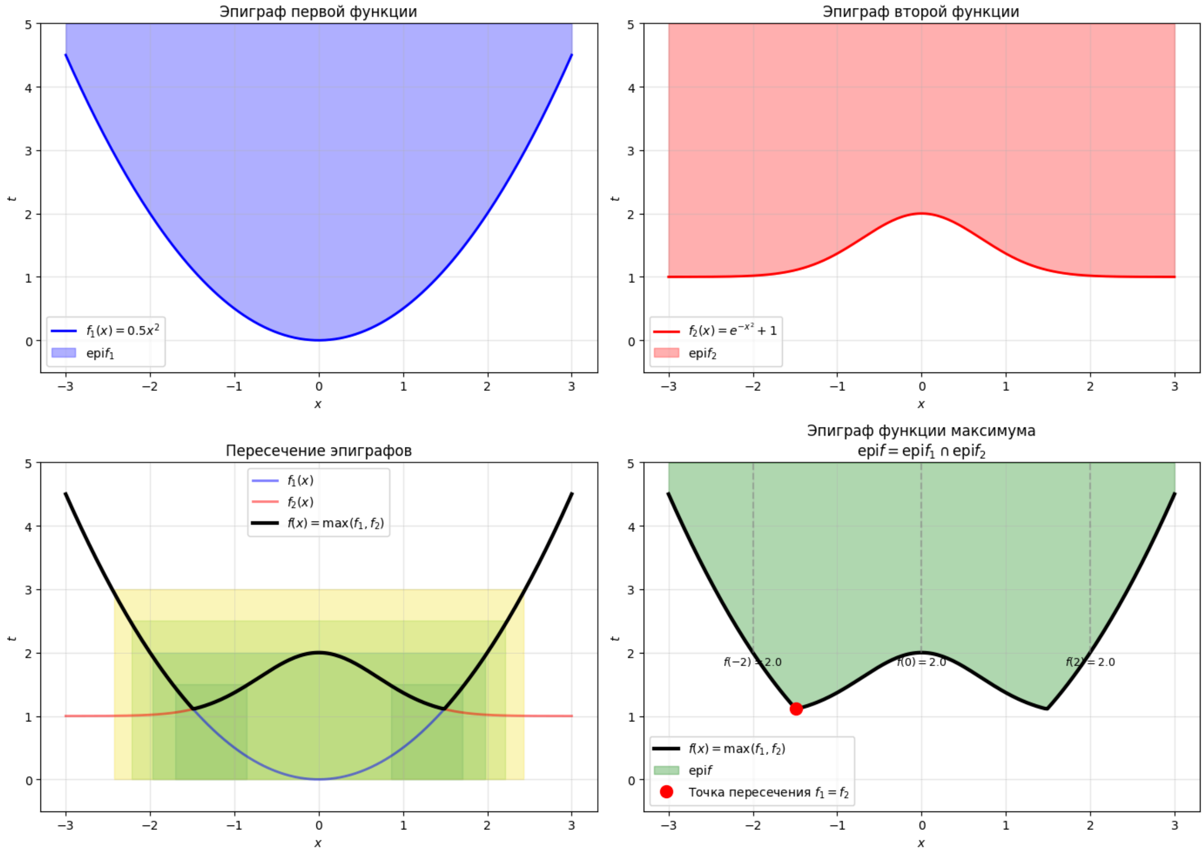
Определение важно, как один из критериев выпуклости функции. Об этом теорема.

Теорема 2: Чтобы $f(x)$ на выпуклом множестве S была выпукла необходимо и достаточно чтобы epi f был выпуклым множеством.

Над эпиграфом можно производить операции, сохраняющие выпуклость, как и над обычными выпуклыми множествами.



Если пересечь эпиграфы двух функций epi f_1, epi f_2, то в результате получаем множество epi f, где $f(x) = \max\{f_1(x), f_2(x)\}$ и определена она на множестве $S_1 \cap S_2$.



Множество подуровня функции

Определение 4: Пусть $f(x)$ — функция на множестве S, определим **множество Лебега/множество подуровня** как

$$\mathcal{L}_\beta = \{x \in S : f(x) \leq \beta\}$$

С этим определением связано ещё одно свойство выпуклости функции: если $f(x)$ - выпуклая функция на выпуклом множестве S, то для любого $\beta > 0$ множество подуровня выпукло. Обратное неверно, в отличие от критерия для эпиграфа.