😕 Дифференциальные критерии выпуклости функций

Если $f(x)$ - дифференцируемая функция на выпуклом множестве (в частности, один раз или два раза непрерывно дифференцируемая), то можно проверить на выпуклость данную функцию с помощью производных. Об этом - дифференциальные критерии.

Дифференциальный критерий выпуклости функции первого порядка

Теорема: Дифференцируемая функция $f(x)$ на выпуклом множестве $S \subseteq \mathbb{R}^n$ выпукла тогда и только тогда, когда $\forall x, y \in S$:

$$f(y) - f(x) \geq \nabla f^T(x) \cdot (y - x)$$

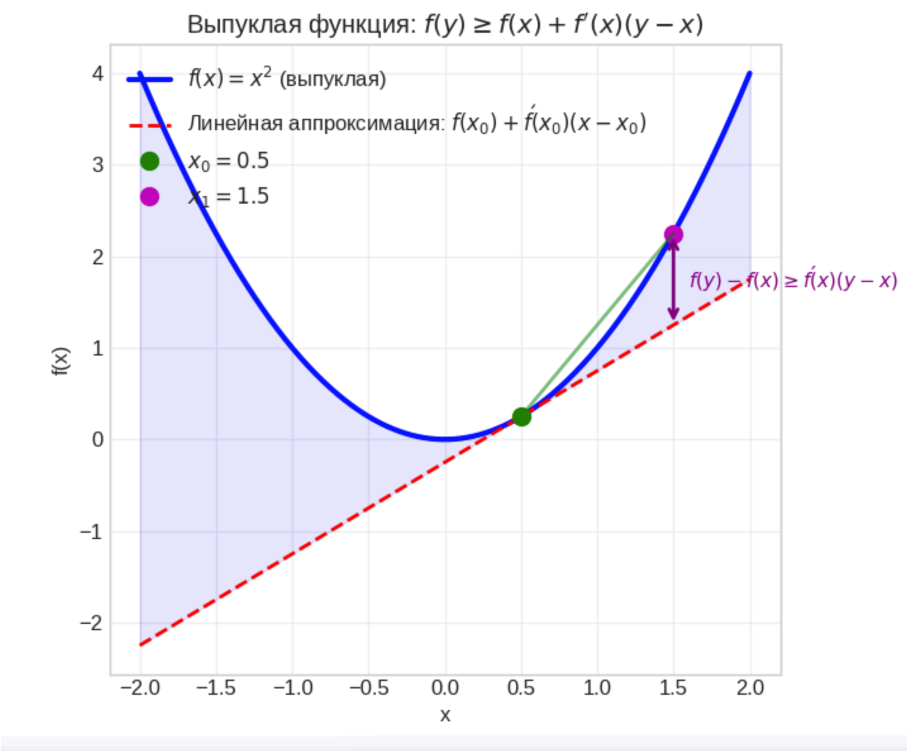
или

$$f(y) - f(x) \geq \langle f_x(x), y - x \rangle$$

Положив, для удобства, $y = x + \Delta x$:

$$f(x + \Delta x) - f(x) \geq \nabla f^T(x) \cdot (\Delta x)$$

То есть, выпуклая функция $f(x)$ больше или равна её линейной аппроксимации рядом Тейлора в любой точке данного множества S.



Дифференциальный критерий выпуклости функции второго порядка

Теорема: Дважды непрерывно дифференцируемая функция $f(x)$ на выпуклом множестве $S \subseteq \mathbb{R}^n$ выпукла тогда и только тогда, когда $\forall x \in \operatorname{int} S$:

$$\nabla^2 f(x) \succeq 0$$

или

$$\langle y, \nabla^2 f(x) y \rangle \geq 0$$

Где $\operatorname{int} S$ - внутренность множества S, $\nabla^2 f(x)$ - гессиан $f(x)$.

Сильная выпуклость

Напомним определение сильной выпуклости функции. Из сильной выпуклости следует обычная выпуклость функции.

Определение: Функция $f(x)$ называется сильно выпуклой (θ —выпуклой) на множестве S, если выполнено:

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

Существуют аналогичные критерии сильной выпуклости, из которых следуют описанные выше, при предположении $\theta = 0$.

Дифференциальные критерии сильной выпуклости функции первого и второго порядков

Теорема: Дифференцируемая функция $f(x)$ на выпуклом множестве $S \subseteq \mathbb{R}^n$ сильно выпукла с const $\theta > 0$ тогда и только тогда, когда:

$$f(y) - f(x) \geq \nabla f^T(x)(y-x) + \frac{\theta}{2}\|y-x\|^2$$

или

$$f(y) - f(x) \geq \langle f_x(x), y - x\rangle + \frac{\theta}{2}\|y-x\|^2$$

Положив $y = x + \Delta x$, имеем:

$$f(x+\Delta x) - f(x) \geq \nabla f^T(x)\Delta x + \frac{\theta}{2}\|\Delta x\|^2$$

Теорема: Дважды непрерывно дифференцируемая функция $f(x)$ на выпуклом множестве $S \subseteq \mathbb{R}^n$: $\text{int}\, S \neq \emptyset$ сильно выпукла с const $\theta > 0$ тогда и только тогда, когда:

$$\nabla^2 f(x) \succeq \theta I$$

или

$$\langle y, \nabla^2 f(x) y\rangle \geq \theta\ \|y\|^2$$
$$\forall x, y \in S$$

То есть, спектр (ie множество собственных значений) гессиана строго отделено от нуля. Аналогично, условие теоремы выполнено при $\theta = \lambda_{\min}(\nabla^2 f)$

Резюме методов исследования функции на выпуклость

Были рассмотрены следующие методы для определения выпуклых (и выпуклых вверх/вогнутых) функций:

- По определению/неравенство Йенсена
- Дифференциальные критерии обычной и сильной выпуклости
- Через другие выпуклые функции и операции, сохраняющие выпуклость, над ними
- Через связь выпуклости функции с её эпиграфом
- Через связь выпуклости функции с её множеством Лебега/множеством подуровня функции