§ 1.1 О выпуклых функциях

Выпуклой (вверх) называется функция, надграфик которой является выпуклым множеством. Из теории выпуклых множеств следует, что это достигается при выполнении неравенства Йенсена

$$f(\alpha x + (1-\alpha)\ y) \le \alpha f(x) + (1-\alpha) f(y)$$

где $0 \le \alpha \le 1$. Если неравенство выполняется для $-f(x)$, то функция называется вогнутой (выпуклой вниз).

Доказываются следующие важные свойства выпуклых функций:

1. Выпуклая функция имеет единственную подозрительную на экстремум точку, причем она будет являться глобальным экстремумом (минимумом для выпуклой вверх и максимумом для выпуклой вниз). Следовательно существует единственная точка x_0, для которой верно $\left.\frac{\partial f}{\partial x}\right|_{x=x_0} = 0$.

2. Вторая производная сохраняет свой знак (> 0 для выпуклой и < 0 для вогнутой функции) и может обратиться в нуль лишь в точке x_0.

Аналогичные заключения справедливы и для функции $f(\vec{x})$ многих переменных. Условие экстремума тогда можно записать в виде системы

$$\frac{\partial f(\vec{x})}{\partial x_i} = 0 \quad \forall i = 1..n$$

Возводя каждое уравнение в квадрат и складывая почленно, получим функцию невязки системы

$$\mathcal{E}[f(\vec{x})] = \sum_{i=1}^{n} \left(\frac{\partial f(\vec{x})}{\partial x_i} \right)^2$$

Покажем, что функция $\mathcal{E}$ тоже имеет единственный экстремум. Запишем условие экстремума:

$$\frac{\partial \mathcal{E}}{\partial x_j} = \frac{\partial \sum\limits_{i=1}^{n} \left(\frac{\partial f(\vec{x})}{x_i} \right)^2}{\partial x_j} = \sum_{i=1}^{n} \frac{\partial f(\vec{x})}{\partial x_i} \frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} = 0 \quad \forall i = 1..n$$

Но производные $\frac{\partial f(\vec{x})}{\partial x_i}$ обращаются в нуль только при $x = x_0$, а значит в любой другой точке $x = x1 \ne x0$ выполнено $\frac{\partial f(\vec{x})}{\partial x_i} \ne 0$. Вторые производные $\frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j}$ сохраняют свой знак и могут обратиться в нуль лишь при $x = x_0$. Следовательно при $x \ne x_0$ нельзя удовлетворить $\frac{\partial \mathcal{E}}{\partial x_j} = 0$, что и требовалось доказать.

§ 1.2 О выпуклости стабилизирующих функционалов

Линейные функции вида $f(x) = \mp x$ являются выпуклыми (вогнутыми). Аналогично определяются и линейные функционалы.

То же верно и для квадратичных функций и функционалов. Покажем это для стабилизирующего функционала

$$\Omega[f(x)] = \int_a^b \sum_{n=0}^{N} q_n(x) \left(\frac{d^{(n)} f(x)}{dx^n} \right)^2 dx$$

Вторая вариация функционала

$$\delta^2\Omega[f(x)] = \frac{\partial^2\Omega[f(x)+\alpha\delta f(x)]}{\partial\alpha^2} = \int\limits_a^b \sum_{n=0}^{N} q_n(x)\big(\delta f^{(n)}(x)\big)^2 dx > 0$$

при $q_n(x) > 0 \quad \forall x \in [a,b]$. Следовательно в общем случае выбора правильных $q_n(x)$ стабилизирующий функционал является выпуклым.

Покажем выпуклость функционала Ω непосредственно по определению выпуклости:

$$\Omega[\alpha f + \beta g] \le \alpha\Omega[f] + \beta\Omega[g]$$

Действительно, для n-го члена суммы в левой части имеем

$$q_n(\alpha f^{(n)} + \beta g^{(n)})^2 = q_n(\alpha^2(f^{(n)})^2 + \beta^2(g^{(n)})^2 + 2\alpha\beta f^{(n)}g^{(n)})$$

а для правой

$$q_n(\alpha(f^{(n)})^2 + \beta(g^{(n)})^2)$$

Вычитая одно из другого, получим

$$q_n(\alpha(\alpha-1)(f^{(n)})^2 + \beta(\beta-1)(g^{(n)})^2 + 2\alpha\beta f^{(n)}g^{(n)})$$

Или, пользуясь тем, что $\beta = 1-\alpha$,

$$-q_n(\alpha\beta(f^{(n)})^2 + \beta\alpha(g^{(n)})^2 - 2\alpha\beta f^{(n)}g^{(n)}) = -q_n\alpha\beta(f^{(n)} - g^{(n)})^2 < 0 \quad \text{iff } q_n(x) > 0 \ \forall x \in [a,b]$$

что и требовалось показать.

Выпуклость стабилизирующего функционала N-го порядка позволяет использовать его в качестве стабилизатора и в случае метода максимума энтропии, не нарушая выпуклости результирующего функционала.