Харьковский национальный университет имени В.Н.Каразина

“ІНФОРМАТИКА І ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ для хіміків”

Лекция №8
Методы оптимизации. Нелинейный МНК.

В.В.Иванов

Безусловная оптимизация - нахождение оптимума (минимизация значения функции)

$$f(x) = 3x^4 - 16x^3 + 18x^2 + 24$$

$$f'(x) = 12x^3 - 48x^2 + 36x = 0$$

$$x^3 - 4x^2 + 3x = 0$$

$$x_1 = 0; \quad x_2 = 1; \quad x_3 = 3$$

$$f''(x) = 36x^2 - 96x - 36$$

$f''(0) = 36$ $\quad f(0) = 24$

$f''(1) = -24$ $\quad f(1) = 29$

$f''(3) = 72$ $\quad f(3) = -3$

Однако не всегда все так просто. В общем случае, когда речь идет о функции многих переменных

$$f(x_1, x_2, ..., x_N) = f(\vec{x})$$

имеется ряд осложнений.

Например

- функция может не имеет простого аналитического выражения
- не удается, (или слишком сложно) вычислить первую производную
- не удается просто решить уравнение $f'(x) = 0$

Пример. Выход реакции, в зависимости от макроскопических параметров системы: температура, давление. В этом случае можно провести эксперимент, который устанавливает соответствие выхода реакции (η(%)) и температура (T^o) и давление ($P^{(атм.)}$).

$$\eta = f(T, P)$$

однако заранее не известно где именно находится минимум. В связи с этим возникает задача как используя минимальные экспериментальные данные найти макроскопические параметры, которые обеспечивают максимум выхода реакции.

Термины

- Глобальный оптимум (минимум) функции f(x)

$$f(x_*) < f(x), \quad \forall x \in V, x \neq x_*$$

- сильный локальный оптимум (минимум)

$$f(x_*) < f(x), \quad \forall x \in N(\eta), \quad x \neq x_*$$

- слабый локальный оптимум (минимум)

$$f(x_*) \leq f(x), \quad \forall x \in N(\eta), \quad x \neq x_*$$

- слабый локальный оптимум (минимум)

$$f'(\vec{x}_*) = \mathrm{grad}(f(\vec{x}_*)) = 0$$

$$f'(\vec{x}_*) = \mathrm{grad}(f(\vec{x}_*)) = 0$$

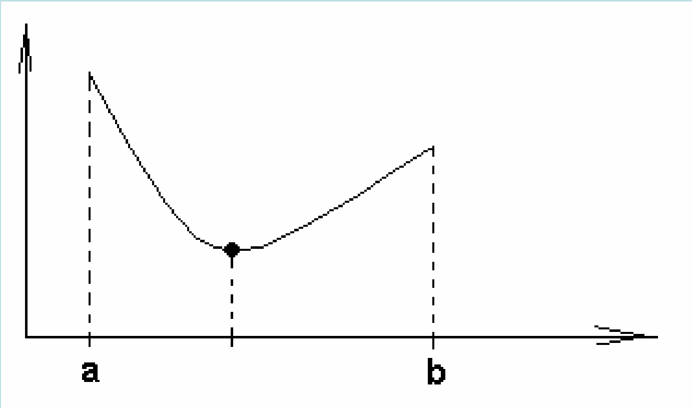
$$\mathrm{grad}(f) = \frac{\partial f}{\partial x_1}\vec{i} + \frac{\partial f}{\partial x_2}\vec{j} + \frac{\partial f}{\partial x_3}\vec{k}$$

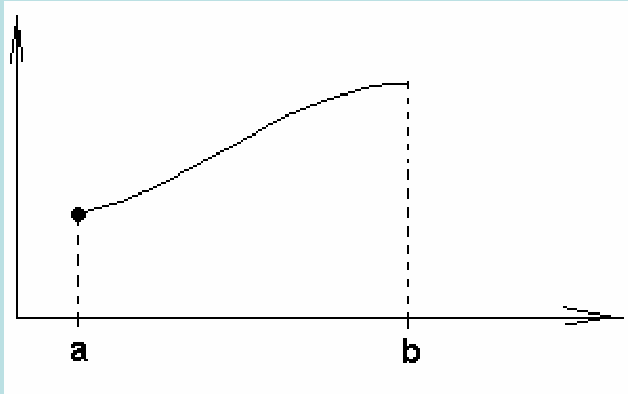
Теорема Вейерштрасса: Всякая функция f(x) непрерывная на отрезке [a,b] принимает на нем свое наибольшее и наименьшее значение.

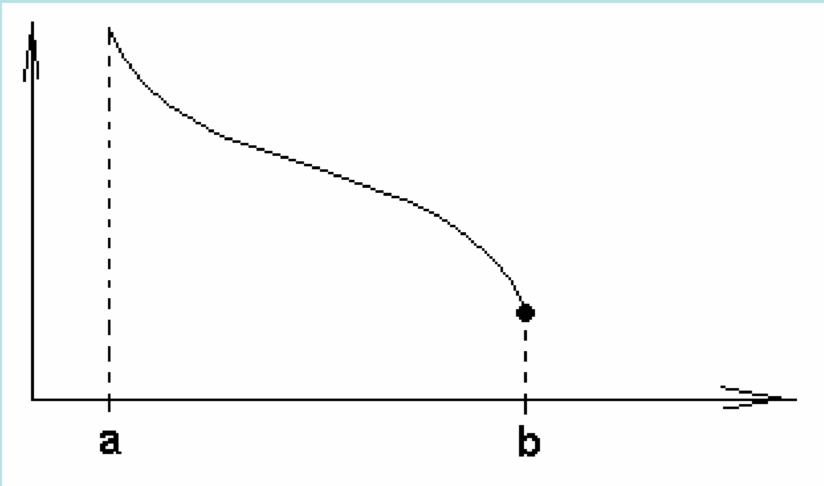
точки на которых $f'(x)=0$ называют стационарными.

$$f(x_1) \le f(x) \le f(x_2)$$

Унимодальная функция – функция имеющая один минимум на заданном отрезке (интервале).







Метод постоянного шага

Это простейший вариант одномерной опримизации.
Весь интервал разбивается на k-фрагментов.

Длина каждого шага $$h = \frac{b - a}{k}$$

Проводится последовательное вычисление (или получено из эксперимента) значений функции в точках:

$$f(a), f(a+h), f(a+2h), ..., f(b-h), f(b)$$

Метод деления отрезка пополам

Середина интервала $c = \frac{a + b}{2}$

Находим две точки $u_1 = \frac{a + b - \delta}{2}$ $u_2 = \frac{a + b + \delta}{2}$

δ – параметр метода $0 < \delta < b - a$

Если $f(u_1) \le f(u_2)$

то минимум находится в интервале: $[a, u_2]$

Если $f(u_1) \ge f(u_2)$

то минимум находится в интервале $[u_1, b]$

Метод золотого сечения

$$\frac{\text{длина отрезка}}{\text{длина большей части}} = \frac{\text{длина большей части}}{\text{длина меньшей части}}$$

$$x = 1 - y = \frac{3 - \sqrt{5}}{2} = 0.381966$$

Многомерная оптимизация

$$f(x_1, x_2, \ldots, x_N) = f(\vec{x})$$

Метод покоординатного спуска

метод представляет собой систему последовательных применений одномерной оптимизации.

Для этого выбираем начальное приближение

$$\vec{x}^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \ldots, x_N^{(0)})$$

Для нее проводим оптимизацю

$$\vec{x} = (x_1^{(1)}, x_2^{(0)}, x_3^{(0)}, \ldots, x_N^{(0)})$$

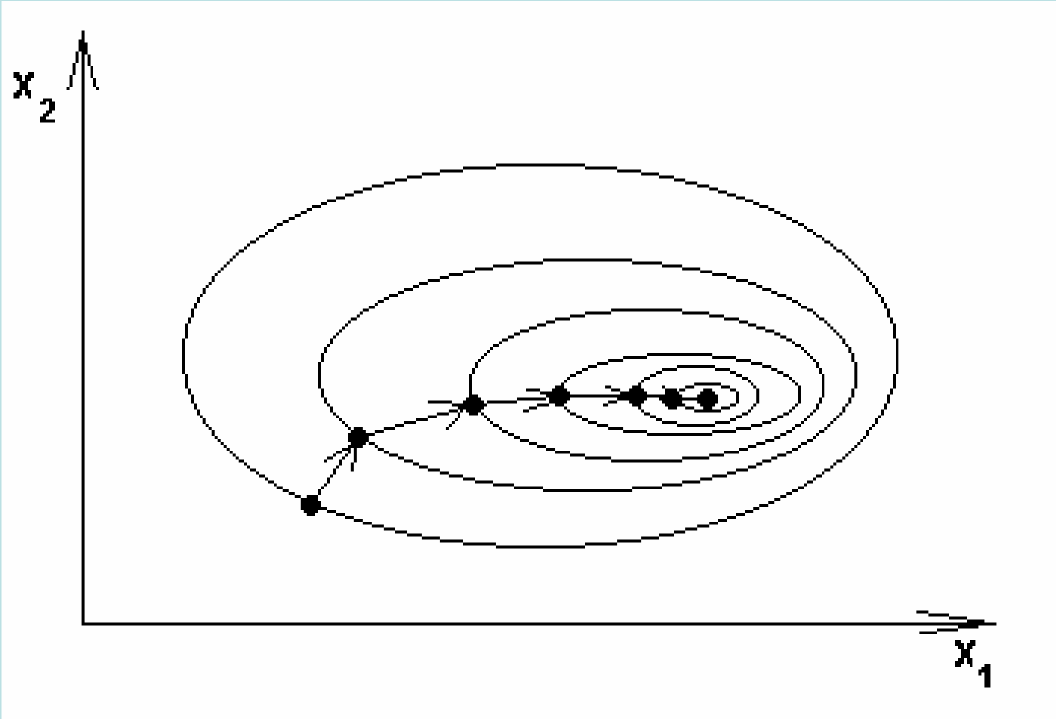
Последовательная оптимизация всех компонент порождает приближение к точке минимума

$$\vec{x}^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \ldots, x_N^{(1)})$$

После прохождения всех N-координат, циклы оптимизации компонент векторов продолжаются

$$\vec{\mathbf{x}}^{(0)}, \vec{\mathbf{x}}^{(1)}, \vec{\mathbf{x}}^{(2)}, \ldots, \vec{\mathbf{x}}^{(k)}$$

Метод наискорейшего (градиентного) спуска
(Steepest Descent)



Коши (1845 г.)

движение вдоль линии анти–градиента согласно уравнению

$$x^{(k)} = x^{(k-1)} - \alpha^{(k-1)} g^{(k-1)}$$

Пример:

$$f(x) = x^2 + x$$

$$g(x) = \frac{\partial f}{\partial x} = 2x + 1$$

$$x = x - \alpha g$$

$$f(\alpha) = (x - \alpha g)^2 + x - \alpha g$$

$$\frac{\partial f}{\partial \alpha} = -2xg + 2\alpha g^2 - g = 0$$

$$\frac{\partial f}{\partial \alpha} = -2xg + 2\alpha g^2 - g = 0$$

$$-2x + 2\alpha g - 1 = 0$$

Поскольку

$$-2x - 1 = -g$$

$$\alpha = \frac{1}{2}$$

$$x^{(0)} = 1$$

$$x^{(1)} = x^{(0)} - \alpha g^{(0)} = 1 - \frac{1}{2}(2 \cdot 1 + 1) = 1 - \frac{3}{2} = -\frac{1}{2}$$

решение

Другое начальное приближение

$$x^{(0)} = 11$$

$$x^{(1)} = 11 - \frac{1}{2}(2 \cdot 11 + 1) = 11 - \frac{23}{2} = \frac{22}{2} - \frac{23}{2} = -\frac{1}{2}$$

Применяем метод SD для решения систем линейных уравнений

$$A\vec{x} = \vec{b}$$

$$J(x) = x^{+}Ax - 2x^{+}b$$

$$J(x) = x^{+}Ax - x^{+}b - b^{+}x$$

$$g(x) = \frac{\partial J(x)}{\partial x^{+}} = Ax - b$$

$$g^{+} = g(x^{+}) = \frac{\partial J(x)}{\partial x} = x^{+}A - b^{+}$$

$$J(\alpha) = (x^+ - \alpha g^+)A(x - \alpha g) - (x^+ - \alpha g^+)b - b^+(x - \alpha g)$$

$$\frac{\partial J(\alpha)}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left[-\alpha g^+ Ax - \alpha x^+ Ag + \alpha^2 g^+ Ag + \alpha g^+ b + \alpha b^+ g\right]$$

$$\frac{\partial J(\alpha)}{\partial \alpha} = -g^+ Ax - x^+ Ag + 2\alpha g^+ Ag + g^+ b + b^+ g$$

$$\boxed{g^+ b - g^+ Ax = -g^+(Ax - b) = -g^+ g}$$

$$\frac{\partial J(\alpha)}{\partial \alpha} = 2\alpha g^+ Ag + 2g^+ g = 0$$

$$\frac{\partial J(\alpha)}{\partial \alpha} = 2\alpha g^{+} A g + 2 g^{+} g = 0$$

$$\alpha = \frac{g^{+} g}{g^{+} A g}$$

$$\boxed{x^{(k+1)} = x^{(k+1)} - \frac{g^{+} g}{g^{+} A g} g^{(k)}}$$

Пример:

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \qquad b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$x^{(0)} = b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad g = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

$$g^+ \cdot g = \begin{pmatrix} 2 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = 4 + 4 = 8$$

$$g^+ \cdot A \cdot g = \begin{pmatrix} 2 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 & 2 \end{pmatrix}\begin{pmatrix} 6 \\ 6 \end{pmatrix} = 12 + 12 = 24$$

$$\alpha = \frac{g^+ \cdot g}{g^+ \cdot A \cdot g} = \frac{8}{24} = \frac{1}{3}$$

$$x = x - \alpha\, g = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$g = Ax - y = \frac{1}{3}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 \\ 3 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0$$

решение $$x = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.333333 \\ 0.333333 \end{pmatrix}$$

Предположим, что начальное приближение выбирается в виде

$$x^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$g = Ax - y = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\alpha = \frac{1}{2} \qquad x^{(1)} = x^{(0)} - \alpha\, g^{(0)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

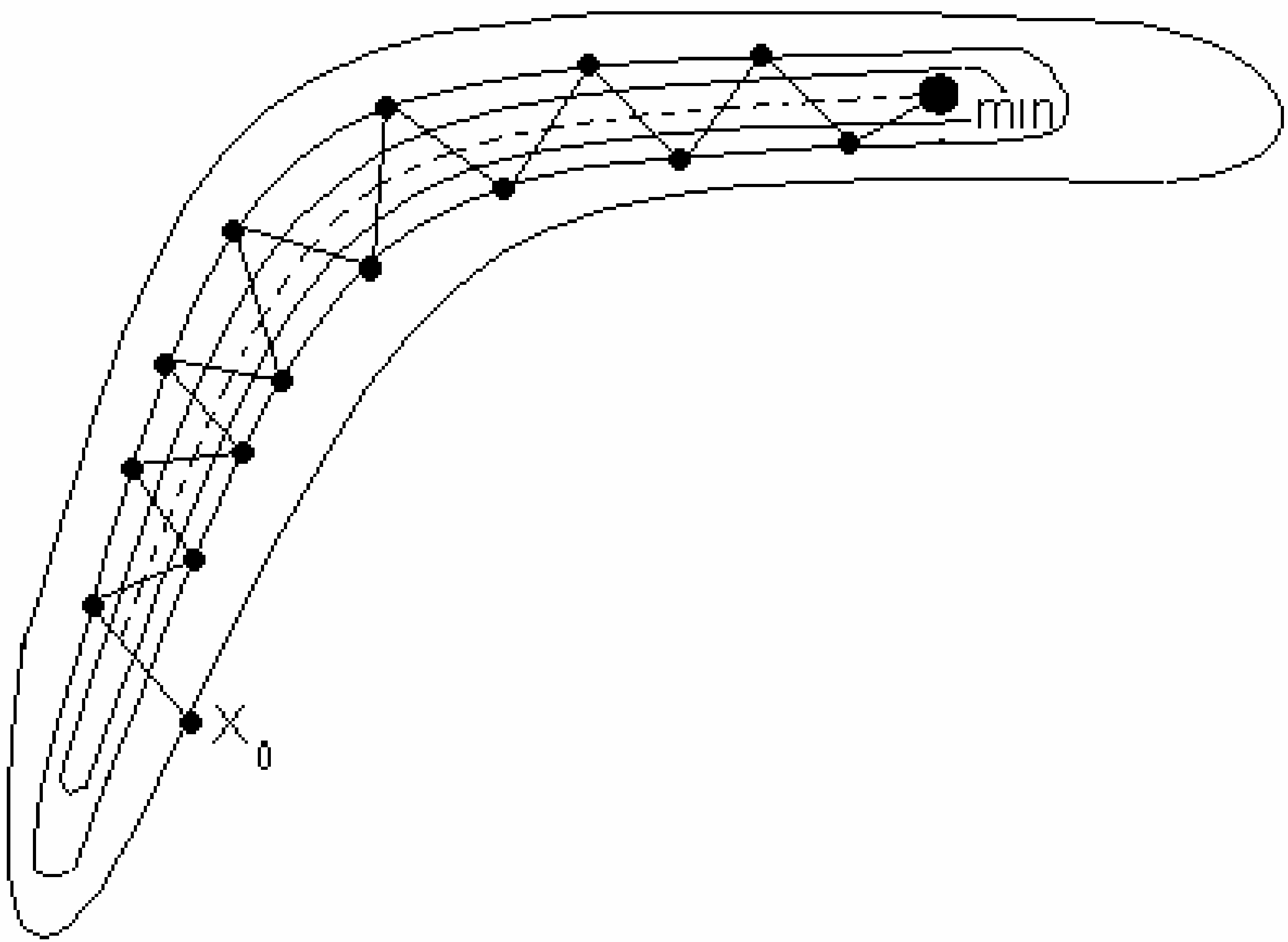
$$g = Ax - y = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -0.5 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

$$x^{(2)} = x^{(1)} - \alpha\, g^{(1)} = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} - \frac{1}{2} \cdot \frac{1}{2}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0.25 \\ 0 \end{pmatrix}$$

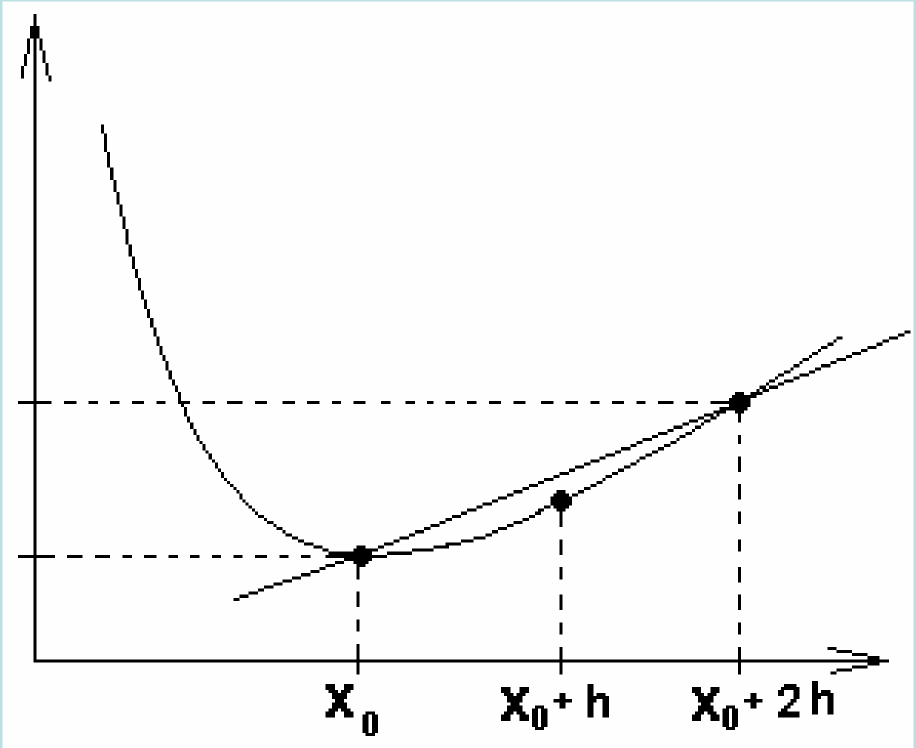
$$x^{(3)} = \begin{pmatrix} 0.421 \\ 0.257 \end{pmatrix} \qquad x^{(4)} = \begin{pmatrix} 0.375 \\ 0.313 \end{pmatrix} \qquad x^{(8)} = \begin{pmatrix} 0.3334 \\ 0.3336 \end{pmatrix} \approx \frac{1}{3}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Скорость достижения решения (кол-во шагов итераций) существенно зависит от выбора начального приближения !

Проблема Оврагов

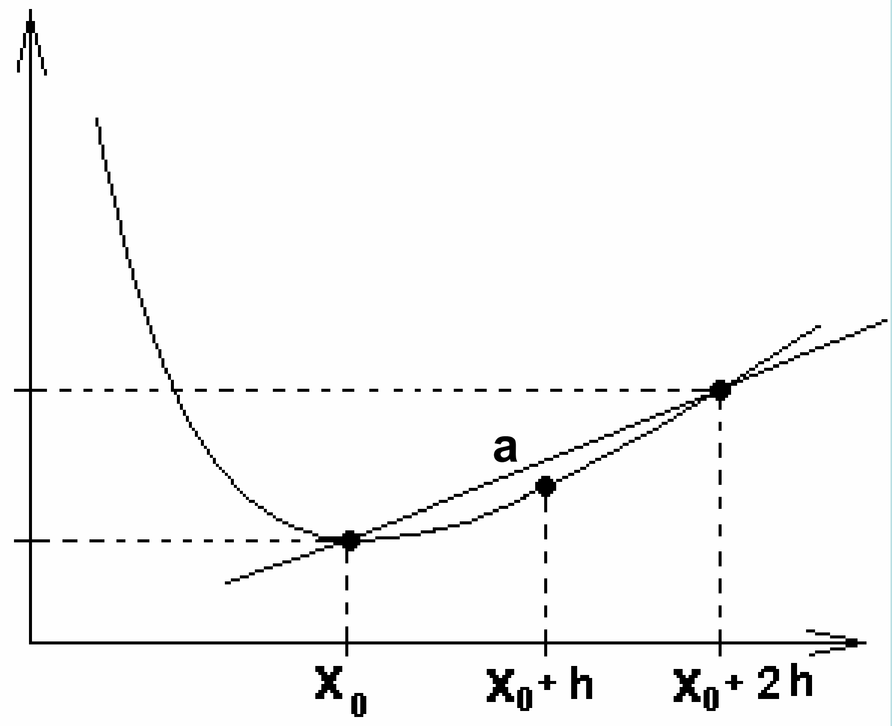


Метод Ньютона



$$a = x_0 + h$$

$$f(a) = f(x_0) + \left(\frac{\partial f}{\partial x}\right)_{x=x_0} h + \tfrac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)_{x=x_0} h^2 + \cdots$$



$$\left(\frac{\partial f}{\partial x}\right)_{x=x_0} = 0$$

$$f(a) = f(x_0) + \tfrac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)_{x=x_0} h^2$$

$$f(x_0 + 2h) = f(x_0) + \tfrac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)_{x=x_0} 4h^2$$

$$f'(x)_{x=a} = \frac{f(x_0 + 2h) - f(x_0)}{2h}$$

$$f'(x)_{x=a} = \frac{1}{2h}\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)4h^2 = \left(\frac{\partial^2 f}{\partial x^2}\right)h = \left(\frac{\partial^2 f}{\partial x^2}\right)(a - x_0)$$

$$x_0 = a - \left(\frac{\partial^2 f}{\partial x^2} \right)^{-1} f'(x)_{x=a}$$

$$x^{(k)} = x^{(k-1)} - \alpha \left(\frac{\partial^2 f}{\partial x^2} \right)^{-1}_{x=x^{(k-1)}} f'(x)_{x=x^{(k-1)}}$$

Обобщение такого алгоритма на случай функции многих переменных

$$f(\vec{a}) = f(\vec{x}_0) + \sum_i \left(\frac{\partial f}{\partial x_i}\right)_{x=x_0} (a_i - x_{0i}) + \frac{1}{2}\sum_{i,j} \left(a_i - x_{0i}\right)\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{x=x_0} \left(a_j - x_{0j}\right) + \cdots$$

$$g_i = \left(\frac{\partial f}{\partial x_i}\right)_{x=x_0}$$

$$H_{i,j} = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)_{x=x_0}$$

компактное представление разложения имеет вид

$$f(\vec{a}) = f(\vec{x}_0) + \vec{g}\cdot(\vec{a} - \vec{x}_0) + \tfrac{1}{2}(\vec{a} - \vec{x}_0)H(\vec{a} - \vec{x}_0)$$

учитывая равенство нулю градиента

$$\vec{g} = 0$$

получаем приближенное выражение для градиента в точке

$$\vec{g} = H(\vec{a} - \vec{x}_0)$$

$$x^{(k)} = x^{(k-1)} - \alpha H^{-1} g^{(k-1)}$$

Пример:

$$f(x_1, x_2) = x_1^2 + x_2^2 + x_1 x_2 + a x_1 + x_2$$

$$\vec{g}(x_1, x_2) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 + a \\ 2x_2 + x_1 + 1 \end{pmatrix}$$

матрица Гесса

$$H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1} & \dfrac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2} \\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$$H^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$H^{-1}\vec{g} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2x_1 + x_2 + a \\ 2x_2 + x_1 + 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3x_1 + 2a - 1 \\ 3x_2 + 2 - a \end{pmatrix}$$

$$\vec{x}^{(0)} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 2a - 1 \\ 2 - a \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 2a - 1 \\ 2 - a \end{pmatrix}$$

вычислим теперь градиент в этой точке

$$g^{(1)} = -\frac{1}{3}\begin{pmatrix} -2a+1+2a-1 \\ -2+a+2-a \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Таким образом вектор

$$x^{(1)} = -\frac{1}{3}\begin{pmatrix} 2a-1 \\ 2-a \end{pmatrix}$$

Является решением задачи

Возьмем теперь такое начальное приближение

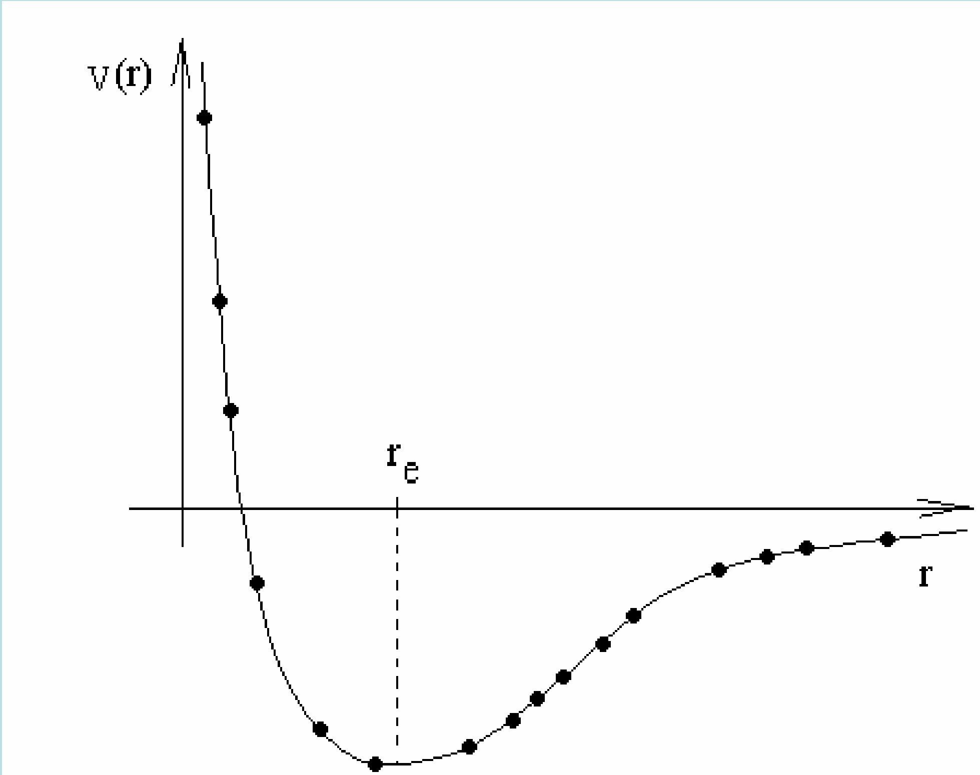
$$\vec{x}^{(0)} = \begin{pmatrix} 101 \\ -5 \end{pmatrix}$$

$$x^{(1)} = \begin{pmatrix} 101 \\ -5 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 302+2a \\ -13-a \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 303 \\ -15 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 302+2a \\ -13-a \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 1-2a \\ -2+a \end{pmatrix} = -\frac{1}{3}\begin{pmatrix} 2a-1 \\ 2-a \end{pmatrix}$$

Приближенные оценки гессиана приводят к так называемым *квазиньютоновым методам* (Давидон, 1959)

Нелінійний метод найменьших квадратів

$$V(r) = D \cdot (1 - e^{-\beta(r - r_e)})^2$$



$$V_k = V(r_k)$$

$$\Phi(D, \beta, r_e) = \sum_{k=1}^{N} \left(\hat{V}_k - V_k\right)^2$$

$$\frac{\partial \Phi(D, \beta, r_e)}{\partial D} = -2 \sum_{k} \left(\hat{V}_k - V_k\right) \frac{\partial V_k}{\partial D}$$

$$\frac{\partial \Phi(D, \beta, r_e)}{\partial \beta} = -2 \sum_{k} \left(\hat{V}_k - V_k\right) \frac{\partial V_k}{\partial \beta}$$

$$\frac{\partial \Phi(D, \beta, r_e)}{\partial r_e} = -2 \sum_{k} \left(\hat{V}_k - V_k\right) \frac{\partial V_k}{\partial r_e}$$

$$\frac{\partial V_k}{\partial x} = \left(\frac{\partial V(r)}{\partial x} \right)_{r=r_k}$$

$$\frac{\partial V(r)}{\partial D} = 1 - 2e^{-\beta(r-r_e)} + e^{-2\beta(r-r_e)}$$

$$\frac{\partial V(r)}{\partial \beta} = 2D(r - r_e)\left(e^{-\beta(r-r_e)} - e^{-2\beta(r-r_e)}\right)$$

$$\frac{\partial V(r)}{\partial r_e} = 2D\beta\left(- e^{-\beta(r-r_e)} + e^{-2\beta(r-r_e)}\right)$$

$$
\begin{cases}
\dfrac{\partial \Phi(D, \beta, r_e)}{\partial D} = 0 \\
\dfrac{\partial \Phi(D, \beta, r_e)}{\partial \beta} = 0 \\
\dfrac{\partial \Phi(D, \beta, r_e)}{\partial r_e} = 0
\end{cases}
$$

$$
x^{(\ell)} = x^{(\ell-1)} - \alpha\, H^{-1} g^{(\ell-1)} \qquad x^{(\ell)} = \begin{pmatrix} D \\ \beta \\ r_e \end{pmatrix}^{(\ell)}
$$

$$H_{ij} = \frac{\partial^2 \Phi}{\partial x_i \partial x_j} = 2\sum_k \frac{\partial V_k}{\partial x_i}\frac{\partial V_k}{\partial x_j} - 2\sum_k \left(\hat{V}_k - V_k\right)\frac{\partial^2 V}{\partial x_i \partial x_j}$$

$$x_i, x_j = D, \beta, r_e$$

$$\frac{\partial^2 V_k}{\partial D^2} = 0$$

$$\frac{\partial^2 V_k}{\partial D \partial \beta} = 2\left(r - r_e\right)\left(e^{-\beta(r-r_e)} - e^{-2\beta(r-r_e)}\right)$$

$$\frac{\partial^2 V_k}{\partial D \partial r_e} = -2\beta\left(e^{-\beta(r-r_e)} - e^{-2\beta(r-r_e)}\right)$$

$$\frac{\partial^2 V_k}{\partial \beta^2} = -2D\left(r - r_e\right)^2\left(-e^{-\beta(r-r_e)} + 2e^{-2\beta(r-r_e)}\right)$$

$$\frac{\partial^2 V_k}{\partial \beta \partial r_e} = 2D\beta\left(r - r_e\right)\left(e^{-\beta(r-r_e)} - 2e^{-2\beta(r-r_e)}\right)$$

$$\frac{\partial^2 V_k}{\partial r_e^2} = -2D\beta^2\left(e^{-\beta(r-r_e)} - 2e^{-2\beta(r-r_e)}\right)$$

У найпростішому варіанті ітераційна процедура

$$\begin{pmatrix} D \\ \beta \\ r_e \end{pmatrix}^{(\ell+1)} = \begin{pmatrix} D \\ \beta \\ r_e \end{pmatrix}^{(\ell)} - \alpha \begin{pmatrix} \dfrac{\partial \Phi(D,\beta,r_e)}{\partial D} \\ \dfrac{\partial \Phi(D,\beta,r_e)}{\partial \beta} \\ \dfrac{\partial \Phi(D,\beta,r_e)}{\partial r_e} \end{pmatrix}$$

$$\vec{g} \equiv \begin{pmatrix} \dfrac{\partial \Phi(D,\beta,r_e)}{\partial D} \\ \dfrac{\partial \Phi(D,\beta,r_e)}{\partial \beta} \\ \dfrac{\partial \Phi(D,\beta,r_e)}{\partial r_e} \end{pmatrix} = -2 \cdot \begin{pmatrix} \sum\limits_k \left(\hat{V}_k - V_k\right) \dfrac{\partial V_k}{\partial D} \\ \sum\limits_k \left(\hat{V}_k - V_k\right) \dfrac{\partial V_k}{\partial \beta} \\ \sum\limits_k \left(\hat{V}_k - V_k\right) \dfrac{\partial V_k}{\partial r_e} \end{pmatrix}$$

$$\left\| \begin{matrix} \dfrac{\partial \Phi(D, \beta, r_e)}{\partial D} \\ \dfrac{\partial \Phi(D, \beta, r_e)}{\partial \beta} \\ \dfrac{\partial \Phi(D, \beta, r_e)}{\partial r_e} \end{matrix} \right\| \le \text{eps}$$