БЕЗУСЛОВНАЯ МНОГОМЕРНАЯ ОПТИМИЗАЦИЯ

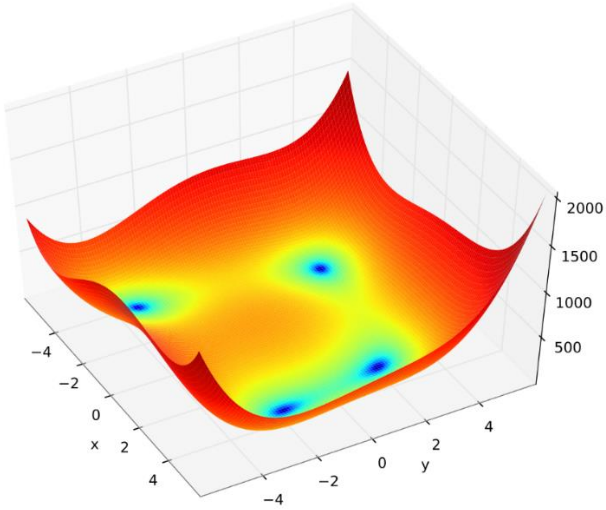
Рассматривается следующая многомерная задача безусловной оптимизации (точнее говоря, задача многомерной локальной безусловной оптимизации): ***найти минимум функции $\Phi(X)$, определенной в n-мерном евклидовом пространстве R^n***

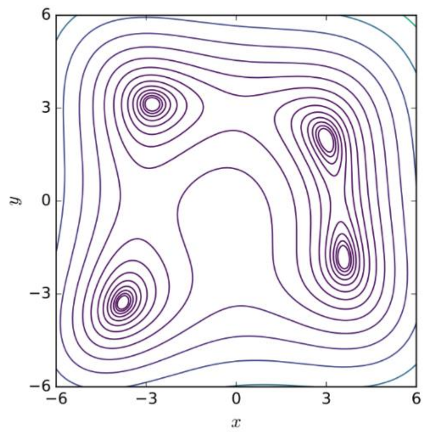
Найти $\min \Phi(\overline{X})$, $\overline{X} \in R^n$***, дано начальное приближение решения*** $\overline{X}^0$ ***и точность eps.***

Примеры для иллюстрации методов

1. Функция Химмельблау

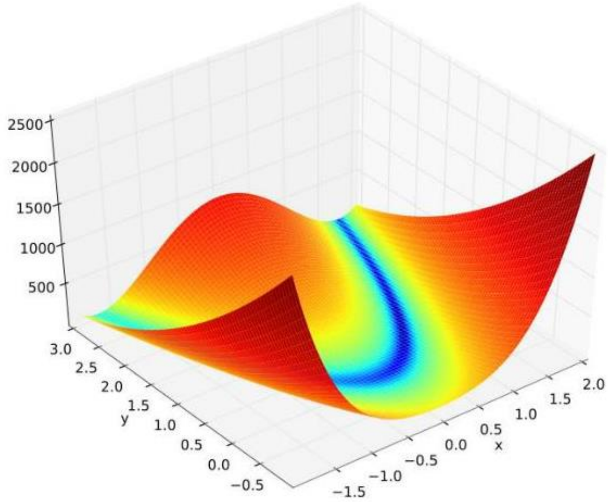
$$f(x,y) = \left(x^2 + y - 11\right)^2 + \left(x + y^2 - 7\right)^2$$

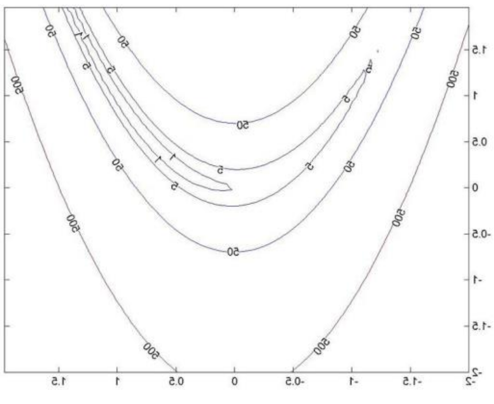




2. Функция Розенброка

$$f(x,y) = (1-x)^2 + 100\left(y + x^2\right)^2$$





1 Метод Гаусса-Зейделя (покоординатного спуска или цикличного покоорднатного спуска)

Общая итерационная схема метода покоординатного спуска:

$$\overline{X}_1^{r+1} = \overline{X}^{r+1} + \lambda_1^r \overline{L}_1$$
$$\overline{X}_2^{r+1} = \overline{X}_1^{r+1} + \lambda_2^r \overline{L}_2$$
$$\ldots$$
$$\overline{X}_n^{r+1} = \overline{X}_{n-1}^{r+1} + \lambda_n^r \overline{L}_n = \overline{X}^{r+1}$$

где *n*-размерность задачи, направление $\overline{L}_i$ выбирается вдоль i-ой координатной оси, т.е.

$$\overline{L}_i = \left(l_{i1}, \ldots, l_{in}\right)$$
$$l_{ij} = \begin{cases} 1 \; если \;\; j = i \\ 0 \; иначе \end{cases}$$

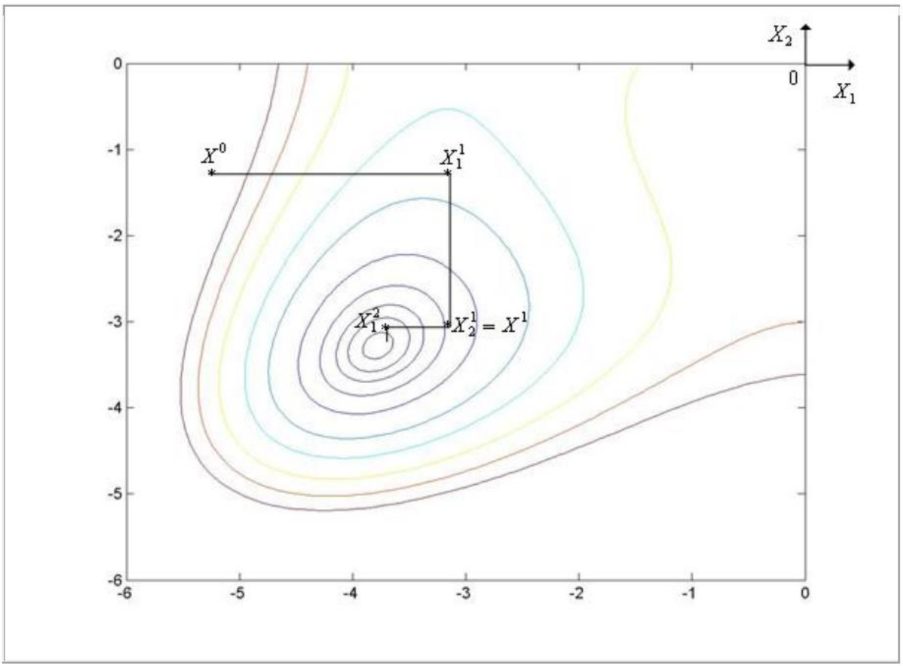
Величина шага λ_i^r вдоль направления на каждом шаге *i* выбирается таким образом, чтобы минимизировать Φ(X) в этом направлении.
Т.е., например, на первом шаге (i=1)

$$\min_{\lambda \in (-\infty; +\infty)} \Phi\left(\overline{X}^r + \lambda \overline{L}_1\right) == \overline{X}^{r+1} + \lambda_1^r \overline{L}_1 = \Phi\left(\overline{X}_1^{r+1}\right)$$

Каждая итерация состоит из *n* шагов, следующая итерация начинается в точке, полученной на последнем шаге предыдущей итерации.
Определение величины шага – задача одномерной оптимизации.

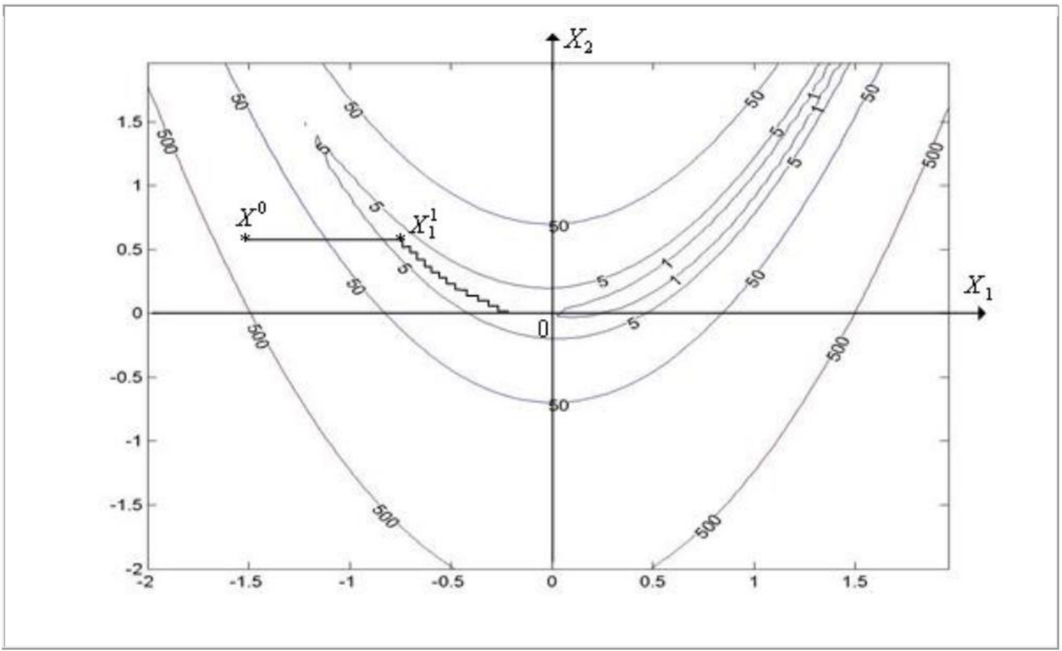
Итерационный процесс завершается при выполнении одного из условий

$$\left\| \overline{X}^{r+1} - \overline{X}^r \right\| \le epsx$$
$$\left| \Phi\left(\overline{X}^{r+1}\right) - \Phi\left(\overline{X}^r\right) \right| \le eps\Phi$$



Траектория поиска минимума **неовражной** функции Химмельблау методом Гаусса-Зейделя

Метод Гаусса-Зейделя медленно сходится на овражных функциях, в которых овраг не ориентирован в направлении какой-либо из координатных осей.



Траектория поиска минимума овражной функции Розенброка методом Гаусса-Зейделя.

2 Метод Хука-Дживса (метод конфигураций, метод пробных шагов)

Итерационные формулы аналогичны формулам, используемым в методе Гаусса-Зейделя

Общая идея – производится последовательность «пробных» шагов в выбранном направлении до тех пор, пока за удачей не последует неуспех.

Общая итерационная схема:

$\overline{X}_0^{r+1} = \overline{X}^r$ - начальное приближение экстремума для i-ой итерации

$\overline{X}_i^{r+1} = \overline{X}_{i-1}^{r+1} + \lambda_1^r \overline{L}_i$

$\overline{X}_n^{r+1} = \overline{X}^{r+1}$ - начальное приближение для следующей итерации (последняя точка текущей)

где направление $\overline{L}_i$ выбирается вдоль i-ой координатной оси, т.е.

$$\overline{L}_i = (l_{i1}, \ldots, l_{in})$$

$$l_{ij} = \begin{cases} 1\, если\;\; j = i \\ 0\, иначе \end{cases}$$

Величина шага λ_i^r вдоль координатной оси определяется из условий

$$\lambda_i^r = \begin{cases} \Delta_i^r\; если\; \Phi\left(\overline{X}_{i-1}^{r+1} + \Delta_i^r \overline{L}_i\right) < \Phi\left(\overline{X}_{i-1}^{r+1}\right), \quad т.е.\, пробный\, шаг\, привел\, к\;\, уменьшению \Phi \\ -\Delta_i^r\; если\; \Phi\left(\overline{X}_{i-1}^{r+1} - \Delta_i^r \overline{L}_i\right) < \Phi\left(\overline{X}_{i-1}^{r+1}\right) < \Phi\left(\overline{X}_{i-1}^{r+1} + \Delta_i^r \overline{L}_i\right), \quad т.е.\, пробный\, шаг\, в\, обратном\, направлени \\ \qquad\qquad\qquad\qquad\qquad\qquad привел\, к\;\, уменьшению \Phi \\ 0,\; если\; \min\left(\Phi\left(\overline{X}_{i-1}^{r+1} - \Delta_i^r \overline{L}_i\right), \Phi\left(\overline{X}_{i-1}^{r+1} + \Delta_i^r \overline{L}_i\right)\right) > \Phi\left(\overline{X}_{i-1}^{r+1}\right) \end{cases}$$

После завершения n шагов выполняется спуск в направлении $\left(\overline{X}^{r+1} - \overline{X}^r\right)$:

$\overline{X}^{r+1} = \overline{X}^r + \alpha^r \left(\overline{X}^{r+1} - \overline{X}^r\right)$, где α^r - т.н. ускоряющий множитель

Модификации метода Хука-Дживса в зависимости от способа выбора α^r:

1. $\alpha^r = const$ (например, α^r =2)
2. α^r выбирается из условия $\Phi\left(\overline{X}^{r+1}\right) < \Phi\left(\overline{X}^r\right)$
3. α^r находится в результате решения одномерной задачи оптимизации Ф в направлении $\left(\overline{X}^{r+1} - \overline{X}^r\right)$

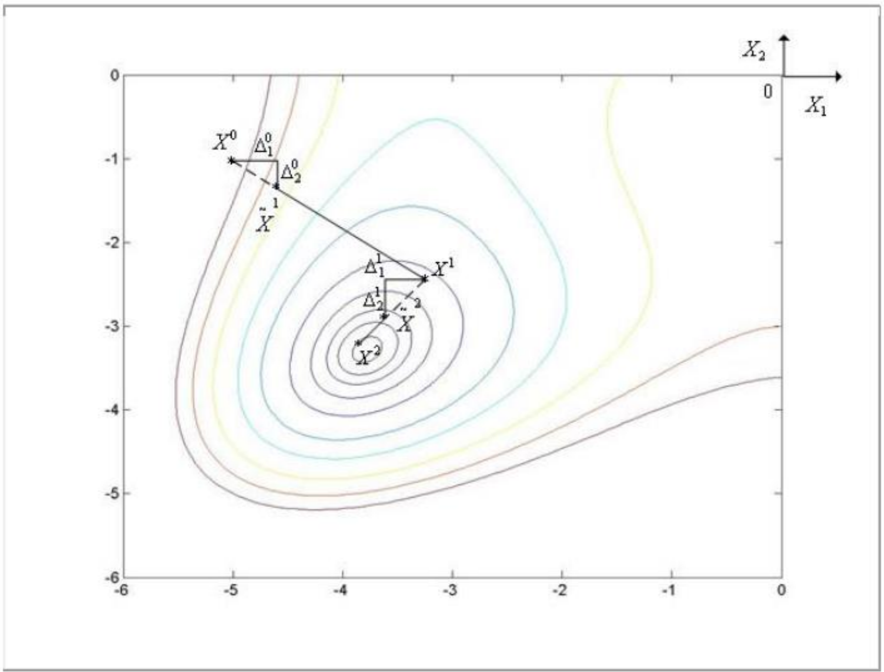
Итерационный процесс завершается при выполнении одного из условий

$$\left\| \overline{X}^{r+1} - \overline{X}^r \right\| \le epsx$$

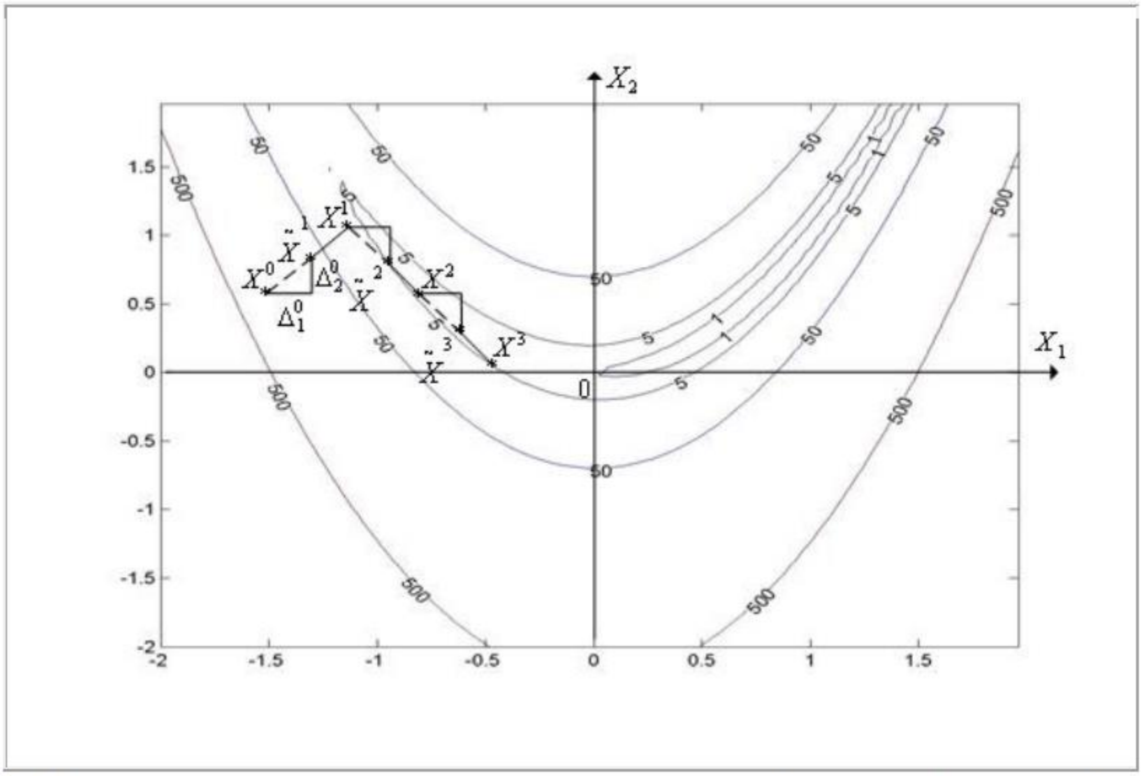
$$\left| \Phi\left(\overline{X}^{r+1}\right) - \Phi\left(\overline{X}^r\right) \right| \le eps\Phi$$

Схема метода Хука-Дживса

1. Задаем начальную точку $\overline{X}^0$ и вектор пробных шагов $\overline{\Delta}^0 = \left(\Delta^0{}_1,\ \Delta^0{}_2 \ldots \Delta^0{}_n\right)$.
   Полагаем r=0.
2. Последовательно находим точки $\overline{X}_1^{r+1}, \overline{X}_2^{r+1}, \ldots, \overline{X}_3^{r+1} = \overline{X}^{r+1}$
3. Если $\overline{X}^{r+1} \neq \overline{X}^r$ - переходим к п.4
   Иначе – уменьшаем длины пробных шагов и переходим к п.2
4. Если одно из условий окончание поиска выполнено - $\overline{X}^* \approx \overline{X}^{r+1}$, конец.
   Иначе – спуск в направлении $\left(\overline{X}^{r+1} - \overline{X}^r\right)$:
   $\overline{X}^{r+1} = \overline{X}^r + \alpha^r\left(\overline{X}^{r+1} - \overline{X}^r\right)$
   r=r+1, переход на п.2



Траектория поиска минимума не овражной функции Химмельблау методом Хука-Дживса



Траектория поиска минимума овражной функции Химмельблау методом Хука-Дживса. Ускоряющий множитель α=const=2

3 Метод Розенброка (метод вращающихся координат)

При решении задачи методом Розенброка на каждой итерации используется преобразование системы координат таким образом, чтобы в новой системе координат одна из осей совпадала с направлением предыдущего шага.
Остальные оси новой системы координат обычно находят с помощью процедуры ортогонализации (Грамма-Шмидта).

Ортогонализация Грамма-Шмидта
Опр. Набор векторов $e_1, e_2 \ldots e_n$ называется ортогональным, если для любых двух векторов из этого набора выполняется

$$(e_1, e_2) = \begin{cases} 0, \text{ если } i \neq j \\ не\, 0, \text{ если } i = j \end{cases}$$

Набор векторов ортонормирован, если они ортогональны и скалярное произведение любых двух равно 1:

$$(e_1, e_2) = \begin{cases} 0, \text{ если } i \neq j \\ 1, \text{ если } i = j \end{cases}$$

Рассмотрим произвольный набор линейно-независимых векторов $p_1,\ p_2 \cdots p_n$
Построим на основе этих векторов ортогональный набор векторов.

$$e_1 = \frac{p_1}{\|p_1\|} = \frac{p_1}{\sqrt{(p_1, p_1)}}$$

Остальные e_k найдем по формулам:

$$y_k = p_k - \sum_{j=1}^{k-1} \frac{(p_k, y_j)}{(y_j, y_j)} y_j \quad k = 2, \ldots, n$$

$$e_k = \frac{y_k}{\|y_k\|} \quad k = 2, \ldots, n$$

Каждая итерация метода Розенброка состоит из двух этапов.
Первый этап.
Выполняется поиск по направлениям, совпадающим с ортами координатных осей.
В зависимости от модификации метода первый этап может выполняться с использованием различных методов (например, Гаусса-Зейделя).
Второй этап.
Система векторов с использованием ортогонализации Грамма-Шмидта заменяется новой системой линейно независимых векторов

Схема метода Розенброка

1. Задаем начальную точку $\overline{X}^0$, полагаем r =0, i =1, и орты исходной системы координат обозначаем $e^0{}_1, e^0{}_2 \ldots e^0{}_n$
2. Исходя из точки $\overline{X}^r$ выполняем одну итерацию по методу Гаусса-Зейделя – получаем точку $\overline{X}^{r+1}$ и совокупность векторов $q^0{}_1, q^0{}_2 \ldots q^0{}_n$.
3. Если одно из стандартных условий окончания итераций

$\left\|\overline{X}^{r+1} - \overline{X}^r\right\| \leq epsx$

$\left|\Phi\left(\overline{X}^{r+1}\right) - \Phi\left(\overline{X}^r\right)\right| \leq eps\Phi$

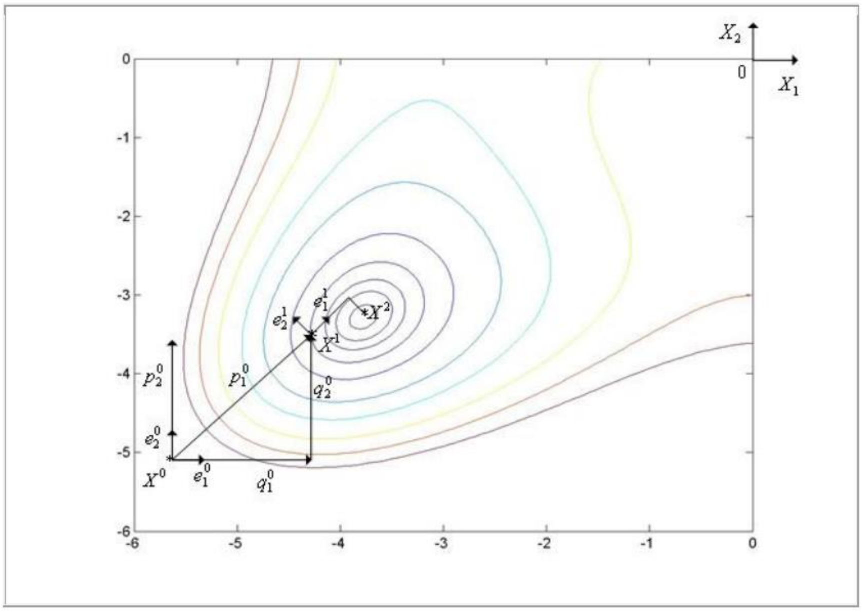
выполнено, то полагаем $\overline{X}* \approx \overline{X}^{r+1}$ и заканчиваем вычисления, иначе переходим к п.4.
4. На основе векторов $q^0{}_1, q^0{}_2 \ldots q^0{}_n$ находим векторы $p^r{}_1, p^r{}_2 \ldots p^r{}_n$:

$$p_i^r = \sum_{j=i}^{n} q_i^r \text{ , i=1,2,…,n}$$

5. С помощью процедуры ортогонализации Грамма-Шмидта выполняем переход от системы векторов $p^r{}_1, p^r{}_2 \ldots p^r{}_n$ к системе векторов $e^{r+1}{}_1, e^{r+1}{}_2 \ldots e^{r+1}{}_n$, полагаем r =r +1 и переходим к п. 2.

Заметим, что $p_n^r = q_n^r$

По сравнению с методом Гаусса-Зейделя и методом Хука-Дживса метод Розенброка имеет, как правило, более высокую эффективность на овражных функциях **с непрямолинейным оврагом.**



Траектория поиска минимума функции Химмельблау методом Розенброка (n=2)