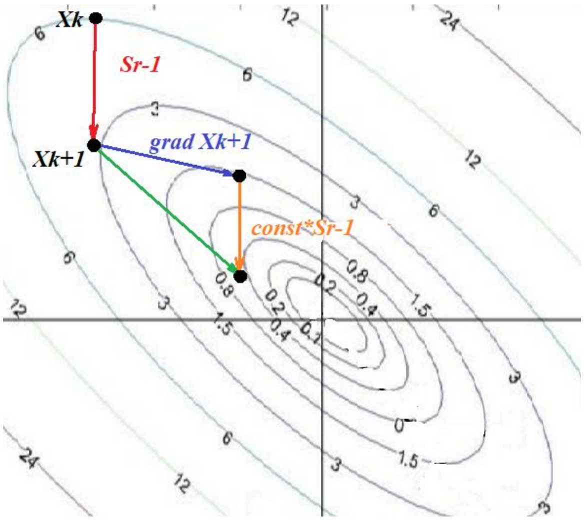
9 Метод сопряженных градиентов

Найти минимум Ф

$$\min_{\bar{X}\in R^n} \Phi(\bar{X}) = \Phi(\bar{X}^*) = \Phi^*$$



$\bar{X}^0$ - начальное приближение решения

$$\bar{X}^{r+1} = \bar{X}^r - \lambda^r \bar{S}^r,$$

где $\bar{S}^0 = \nabla\Phi(\bar{X}^0)$

$$\bar{S}^r = \nabla\Phi(\bar{X}^r) + \beta^r \bar{S}^{r-1}, \quad r = 1,2,\ldots$$

$$\beta^r = \frac{\left\|\nabla\Phi^r\right\|^2}{\left\|\nabla\Phi^{r-1}\right\|^2} = \frac{\sum_{i=1}^{n}\left(\frac{\partial\Phi(\bar{X}^r)}{\partial x_i}\right)^2}{\sum_{i=1}^{n}\left(\frac{\partial\Phi(\bar{X}^{r-1})}{\partial x_i}\right)^2}$$

Таким образом, метод сопряженных градиентов отличается от метода наискорейшего спуска только выбором направления спуска. При этом направление определяется не только антиградиентом, но и направление спуска на предыдущем шаге. Это позволяет более полно, чем в рассмотренных выше градиентных методах, учитывать особенности минимизируемой функции при построении последовательных приближений к ее точке минимума.

Пример 1

Найти минимум функции методом сопряженных градиентов.

$$f(x,y) = x^2 + 2y^2, \quad x, y \in R$$

Начальное приближение $x^0=y^0=1$.

Критерий окончания итераций $|\partial f/\partial x| \le 0.1; |\partial f/\partial y| \le 0.1$

Решение

Шаг 0.

Вычислим градиент $gradf = \begin{pmatrix} \partial f/\partial x \\ \partial f/\partial y \end{pmatrix} = \begin{pmatrix} 2x \\ 4y \end{pmatrix}$

Шаг вдоль антиградиента на λ:

$$\begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} x^0 \\ y^0 \end{pmatrix} - \lambda\, grad(x^0, y^0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \lambda \begin{pmatrix} 2\cdot 1 \\ 4\cdot 1 \end{pmatrix} = \begin{pmatrix} 1-2\lambda \\ 1-4\lambda \end{pmatrix}$$

Значение функции в точке $\begin{pmatrix} x^1 \\ y^1 \end{pmatrix}$

$$f(1-2\lambda, 1-4\lambda) = (1-2\lambda)^2 + 4(1-4\lambda)^2$$

Найдем λ, при которой f достигает минимума: $\lambda = \dfrac{5}{18}$.

Отсюда $\begin{pmatrix} x^1 \\ y^1 \end{pmatrix} = \begin{pmatrix} 1-2\lambda \\ 1-4\lambda \end{pmatrix} = \begin{pmatrix} 4/9 \\ -1/9 \end{pmatrix}$

Проверим критерий

$gradf(x^1, y^1) = \begin{pmatrix} 2x^1 \\ 4y^1 \end{pmatrix} = \begin{pmatrix} 8/9 \\ -4/9 \end{pmatrix}$ - критерий не выполняется ни по одной координате

Шаг 1.

$$\bar{S}^r = grad\ f^r + \beta^r \bar{S}^{r-1}, \quad где\ \beta^r = \frac{\left\| grad\ f^r \right\|^2}{\left\| grad\ f^{r-1} \right\|^2}$$

$$\beta^1 = \frac{\left\| grad\ f^1 \right\|^2}{\left\| grad\ f^0 \right\|^2} = \frac{\left\| \begin{pmatrix} 8/9 \\ -4/9 \end{pmatrix} \right\|^2}{\left\| \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\|^2} = \frac{64+16}{81 \cdot (4+16)} = \frac{4}{81}$$

$$\bar{S}^1 = grad\ f^1 + \beta^1 \bar{S}^0 = \begin{pmatrix} 8/9 \\ -4/9 \end{pmatrix} + \frac{4}{81}\begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 80/81 \\ -20/81 \end{pmatrix}$$

Шаг вдоль $\bar{S}^1$ на λ:

$$\begin{pmatrix} x^2 \\ y^2 \end{pmatrix} = \begin{pmatrix} x^1 \\ y^1 \end{pmatrix} - \lambda \bar{S}^1 = \begin{pmatrix} 4/9 \\ -1/9 \end{pmatrix} - \lambda \begin{pmatrix} 80/81 \\ -20/81 \end{pmatrix} = \begin{pmatrix} 4/9 - \lambda\, 80/81 \\ -1/9 + \lambda\, 20/81 \end{pmatrix}$$

Найдем λ, при которой *f* достигает минимума в точке $\begin{pmatrix} x^2 \\ y^2 \end{pmatrix}$: $\lambda = \dfrac{9}{20}$.

Отсюда $\begin{pmatrix} x^2 \\ y^2 \end{pmatrix} == \begin{pmatrix} \frac{4}{9} - \frac{9}{20}\frac{80}{81} \\ -\frac{1}{9} + \frac{9}{20}\frac{20}{81} \end{pmatrix} = \begin{pmatrix} \frac{4}{9} - \frac{9}{20}\frac{80}{81} \\ -\frac{1}{9} + \frac{9}{20}\frac{20}{81} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.

Проверяем критерий – выполняется.

Ответ: min f=f(0,0)=0.

Пример 2

Найти минимум функции методом сопряженных градиентов.

$$f(x,y) = x^2 + 2y^2 + xy - 7x - 7y, \qquad x, y \in R$$

Начальное приближение $x^0=y^0=0$.

Критерий окончания итераций $|\partial f/\partial x| \le 0.1; |\partial f/\partial y| \le 0.1$

$$\nabla f(x^{(0)}) = (-7,\ -7),$$

$$p^{(0)} = \nabla f(x^{(0)}) = (-7,\ -7),$$

$$g^{(0)}(\alpha) = f(0+7\alpha,\ 0+7\alpha) = 98(2\alpha^2 - \alpha).$$

Для нахождения точки минимума $\alpha^{(0)}$ функции $g^{(0)}(\alpha)$ используем условие $\left.\frac{dg^{(0)}(\alpha)}{d\alpha}\right|_{\alpha=\alpha^{(0)}} = 0$, то есть $\alpha^{(0)} = \frac{1}{4}$.

Откуда получим $x^{(1)} = (0,0) - \frac{1}{4}\cdot(-7,\ -7) = \left(\frac{7}{4},\ \frac{7}{4}\right)$.

И т.д.

Ответ: *min f=f(3;1)=-14.*

10 Метод Ньютона

Пусть Ф всюду дважды дифференцируема в n-мерном евклидовом пространстве R^n.

Найти минимум Ф

$$\min_{\overline{X}\in R^n} \Phi(\overline{X}) = \Phi(\overline{X}^*) = \Phi^*$$

$\overline{X}^0$ - начальное приближение решения

Рассмотрим первые три члена разложения функции Ф в ряд Тейлора в окрестности точки $\overline{X}^r$

$$\Phi(\overline{X}) \approx \tilde{\Phi}^r(\overline{X}) = \Phi(\overline{X}^r) + (\nabla\Phi(\overline{X}^r), (\overline{X} - \overline{X}^r)) + \frac{1}{2}(H(\overline{X}^r)(\overline{X} - \overline{X}^r), (\overline{X} - \overline{X}^r))$$

где $H(\overline{X}^r)$ -- матрица Гессе функции Ф.

Отсюда

$$\nabla\tilde{\Phi}^r(\overline{X}) = \nabla\Phi(\overline{X}^r) + H(\overline{X}^r)(\overline{X} - \overline{X}^r)$$

Если $H(\overline{X}^r)$ положительно определена, то $\tilde{\Phi}^r$ достигает минимума в точке, где ее градиент - нулевой вектор, т.е.

$$\nabla\Phi(\overline{X}^r) + H(\overline{X}^r)(\overline{X} - \overline{X}^r) = 0$$

Или

$$\nabla\Phi(\overline{X}^r) = -H(\overline{X}^r)(\overline{X} - \overline{X}^r)$$

$$H^{-1}(\overline{X}^r)\nabla\Phi(\overline{X}^r) = -H^{-1}(\overline{X}^r)H(\overline{X}^r)(\overline{X} - \overline{X}^r)$$

откуда

$$\overline{X}^{r+1} = \overline{X}^r - H^{-1}(\overline{X}^r)\nabla\Phi(\overline{X}^r) = \overline{X}^r + \Delta^r$$

Т.е. получили фактически метод Ньютона (касательных).

Замечание: для того, чтобы избежать обращения матрицы H, на практике обычно находят Δ^r из решения СЛАУ

$$H(\overline{X}^r)\Delta^r = -\nabla\Phi(\overline{X}^r)$$

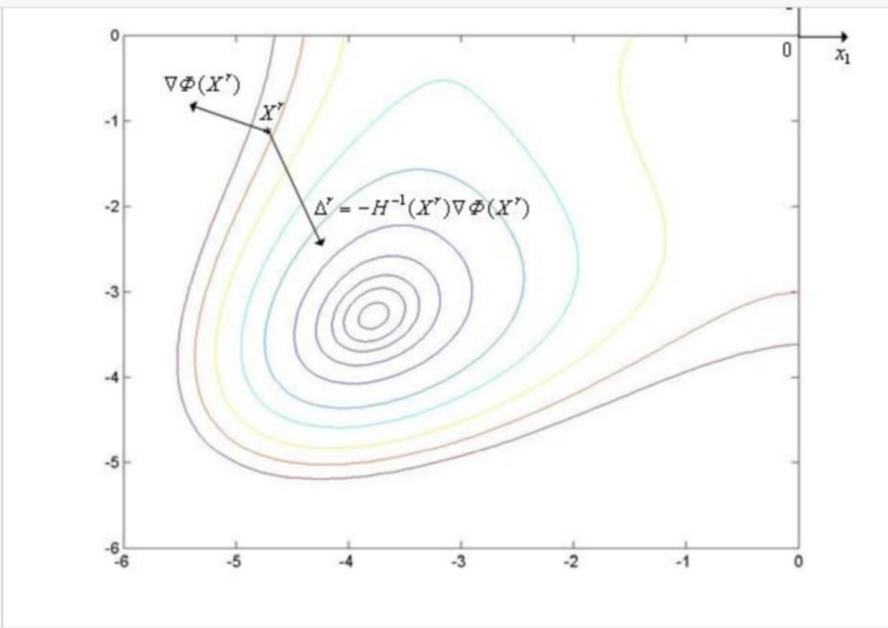
Свойства метода Ньютона

1. Метод **эффективен при минимизации овражных функций**.
   Найдем скалярное произведение

$$\left(\nabla\Phi\left(\overline{X}^r\right),\Delta^r\right)=\left(\nabla\Phi\left(\overline{X}^r\right),H^{-1}\left(\overline{X}^r\right)\nabla\Phi\left(\overline{X}^r\right)\right)<0$$

в силу положительной определенности матрицы Гессе в точке $\overline{X}^r$.

Т.е. вектор поиска Δ^r образует тупой угол с градиентом Ф (см рис.) => вектор поиска может составлять с осью оврага меньший угол, чем антиградиент.



2. Если размерность задачи велика, то **обращение матрицы Гессе** может потребовать значительных вычислительных ресурсов.
   Для того, чтобы избежать обращения матрицы H, на практике обычно находят Δ^r из решения СЛАУ

$$H\left(\overline{X}^r\right)\Delta^r=-\nabla\Phi\left(\overline{X}^r\right)$$

3. Значение Ф в точке $\overline{X}^{r+1}$ может оказаться больше, чем значение в точке $\overline{X}^r$ вследствие большой величины шага (**«перескок» минимума** Ф вдоль направления Δ^r)
   Величина шага в направлении Δ^r , которая приводит к убыванию функции Ф, может быть обеспечена путем добавления в итерационную формулу коэффициента λ^r , т.е. итерационная формула имеет вид

$$\overline{X}^{r+1}=\overline{X}^r-\lambda^r H^{-1}\left(\overline{X}^r\right)\nabla\Phi\left(\overline{X}^r\right)=\overline{X}^r+\lambda^r\Delta^r$$

где коэффициент λ^r выбирают тем или иным способом так, чтобы обеспечить условие

$$\Phi\left(\overline{X}^{r+1}\right) < \Phi\left(\overline{X}^{r}\right)$$

4. **Направление Δ^r ведет к убыванию Ф только при положительной определенности Н,** т.е. на каждой итерации необходимо проверять обусловленность Н. Более того, матрица Н может быть вырожденной (т.е. не иметь обратной).

   Для того, чтобы направление спуска независимо от определенности матрицы Н вело к убыванию функции Ф, в качестве вектора Δ^r можно использовать вектор

   $$\Delta^r = -\left(\mu^r I + H\left(\overline{X}^{r}\right)\right)^{-1} \nabla\Phi\left(\overline{X}^{r}\right)$$

   где I– единичная матрица, а μ^r >0 – параметр, выбираемый так, чтобы матрица $\mu^r I + H\left(\overline{X}^{r}\right)$ являлась положительно определенной.

Схема метода

1. Задаем начальную точку $\overline{X}^0$, начальную величину шага λ^0=1, коэффициент дробления шага $\nu \in (0,1]$. Счетчик количества итераций r=0.
2. Вычисляем в точке $\overline{X}^r$ вектор градиента $\nabla\Phi\left(\overline{X}^{r}\right)$ и матрицу Гессе $H\left(\overline{X}^{r}\right)$
3. Находим вектор Δ^r из решения СЛАУ

   $$H\left(\overline{X}^{r}\right)\Delta^r = -\nabla\Phi\left(\overline{X}^{r}\right)$$

4. Вычисляем компоненты вектора $\overline{X}^{r+1}$

   $$\overline{X}^{r+1} = \overline{X}^{r} + \Delta^r$$

5. Вычисляем значение функции в найденной точке $\Phi\left(\overline{X}^{r+1}\right)$
6. Проверяем условие окончания поиска можно использовать одно из стандартных условий:

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

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

   $$\left\|\nabla\Phi^r\right\| \le eps\_grad$$

Если условие окончания поиска выполнено, то $\overline{X}* \approx \overline{X}^{r+1}$, конец.
Если нет – п.7

7. Если условие

   $\Phi\left(\overline{X}^{r+1}\right) < \Phi\left(\overline{X}^{r}\right)$, то полагаем r=r+1 и переход на п.2 иначе фиксированное количество раз полагаем $\lambda^r = \nu\,\lambda^r$ и переходим на п.4

Методы, использующие случайный поиск

10 Метод с возвратом при неудачном шаге

Найти минимум Ф

$$\min_{\overline{X}\in R^n} \Phi(\overline{X}) = \Phi(\overline{X}^*) = \Phi^*$$

$\overline{X}^0$ - начальное приближение решения

Итерационная формула:

$$\overline{X}^{r+1} = \overline{X}^r + \lambda^r \overline{s}^r,$$

где $\overline{s}^r$ - единичный случайный вектор $\overline{S}^r = \dfrac{\overline{\psi}^r}{\left\|\overline{\psi}^r\right\|}$, $\overline{\psi}^r = \left(\psi_1^r, \psi_2^r, \ldots, \psi_n^r\right)$

ψ_i^r – независимые случайные величины, равномерно распределенные на интервале [-1,1].

Схема метода:

1. Задаем:

   $\overline{X}^0$ - начальное приближение , λ^0 - начальная длина шага

   r = 0

   K – предельное количество неудачных попыток (рекомендуется К=3n)

2. k=1 начальное значение счетчика числа неудачных попыток

3. Получаем реализацию случайных чисел $\psi_1^r, \psi_2^r, \ldots, \psi_n^r$

находим пробную точку

$$\overline{X}^{r+1} = \overline{X}^r + \lambda^r \frac{\overline{\psi}^r}{\left\|\overline{\psi}^r\right\|}$$

4. Вычисляем значение $\Phi(\overline{X}^{r+1})$.

5. Если $\Phi(\overline{X}^{r+1}) < \Phi(\overline{X}^r)$, то полагаем r = r+1 и переходим к п.3.

   Иначе – переходим к п.6.

6. Полагаем k = k+1. Если k <K , то переходим к п.3. Иначе – переходим к п.7.

7. Проверяем условие окончания поиска (см. ниже). Если выполнено, то полагаем X* ≈ Xr+1 и завершаем итерации.

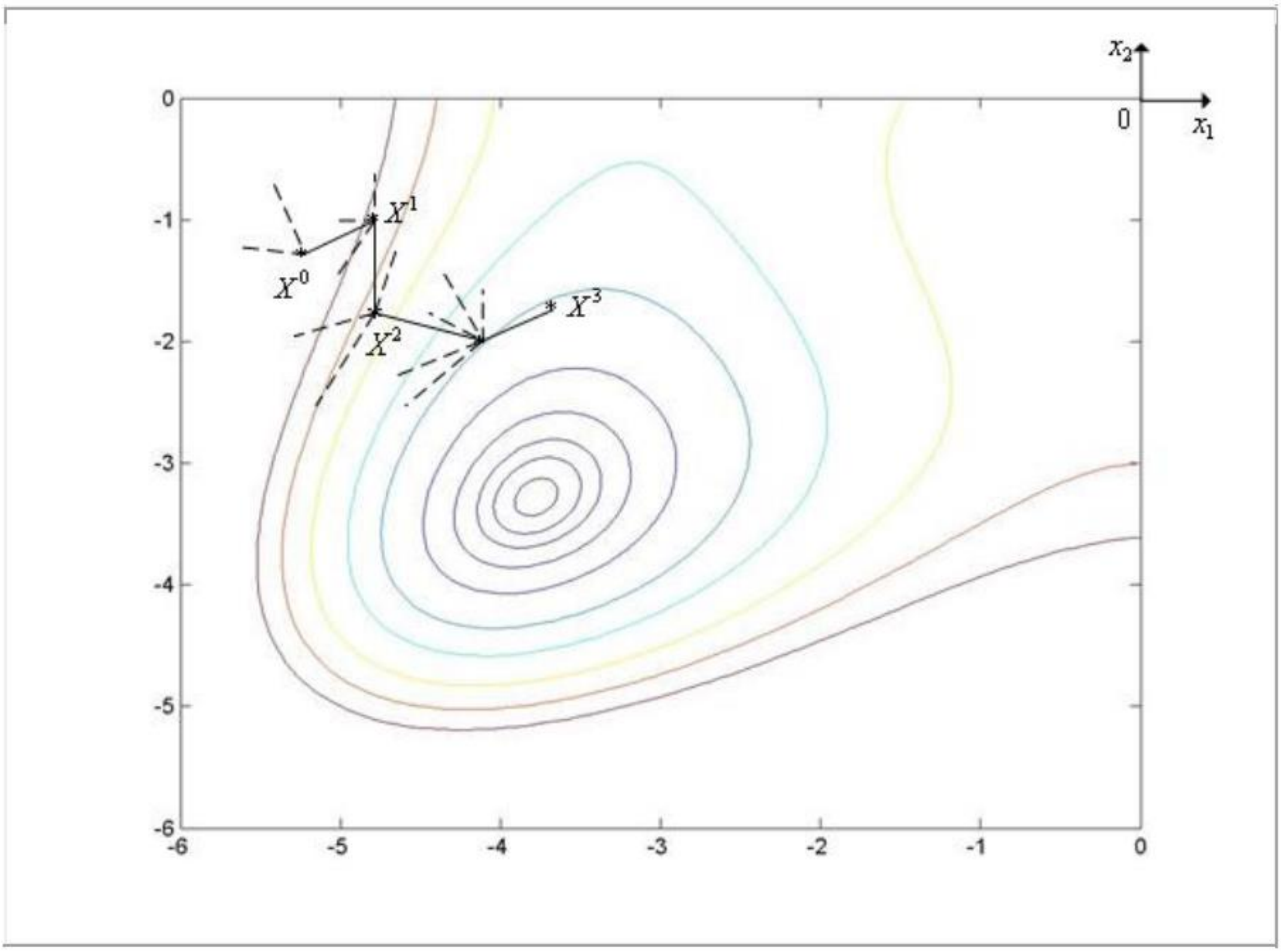
Иначе – полагаем r = r+1, $\lambda^r = \alpha\, \lambda^r$, $\alpha \in (0{,}1)$ и переходим к п.2.

Здесь α – коэффициент уменьшения шага (свободный параметр метода).

В качестве условия окончания поиска можно использоваться одно из стандартных условий окончания итераций:

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

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



10(а) Метод наилучшей пробы - модификация метода с возвратом при неудачном шаге (также одношаговый метод оптимизации, на каждой итерации разыгрывается несколько случайных направлений и выбирается лучшее).

Схема метода наилучшей пробы

1. Задаем:

   $\overline{X}^0$ - начальное приближение, λ^0 - начальная длина шага

   r = 0

   M – количество пробных направлений в одной точке

2. Генерируем M случайных векторов $\overline{\psi}_1^r, \overline{\psi}_2^r, ..., \overline{\psi}_M^r$

находим M пробных точек

$$\overline{X}_i^{r+1} = \overline{X}_i^r + \lambda^r{}_i \frac{\overline{\psi}_i^r}{\left\|\overline{\psi}_i^r\right\|}$$

2. Вычисляем значения функции Ф(X) в пробных точках, находим минимальное из этих значений

$$\Phi\left(\overline{X}^{r+1}\right) = \Phi\left(\overline{X}_k^{r+1}\right) = \min \Phi\left(\overline{X}_i^{r+1}\right)$$

4. Если $\Phi\left(\overline{X}^{r+1}\right) < \Phi\left(\overline{X}^{r1}\right)$, то полагаем r = r +1 и переходим к п.2. Иначе – переходим к п.5.

5. Проверяем условие окончания поиска. Если выполнено, то полагаем X* ≈ Xr+1 и завершаем итерации.

Иначе – полагаем r = r+1, $\lambda^r = \alpha\, \lambda^r$, $\alpha \in (0,1)$ и переходим к п.2.

Здесь α – коэффициент уменьшения шага (свободный параметр метода).

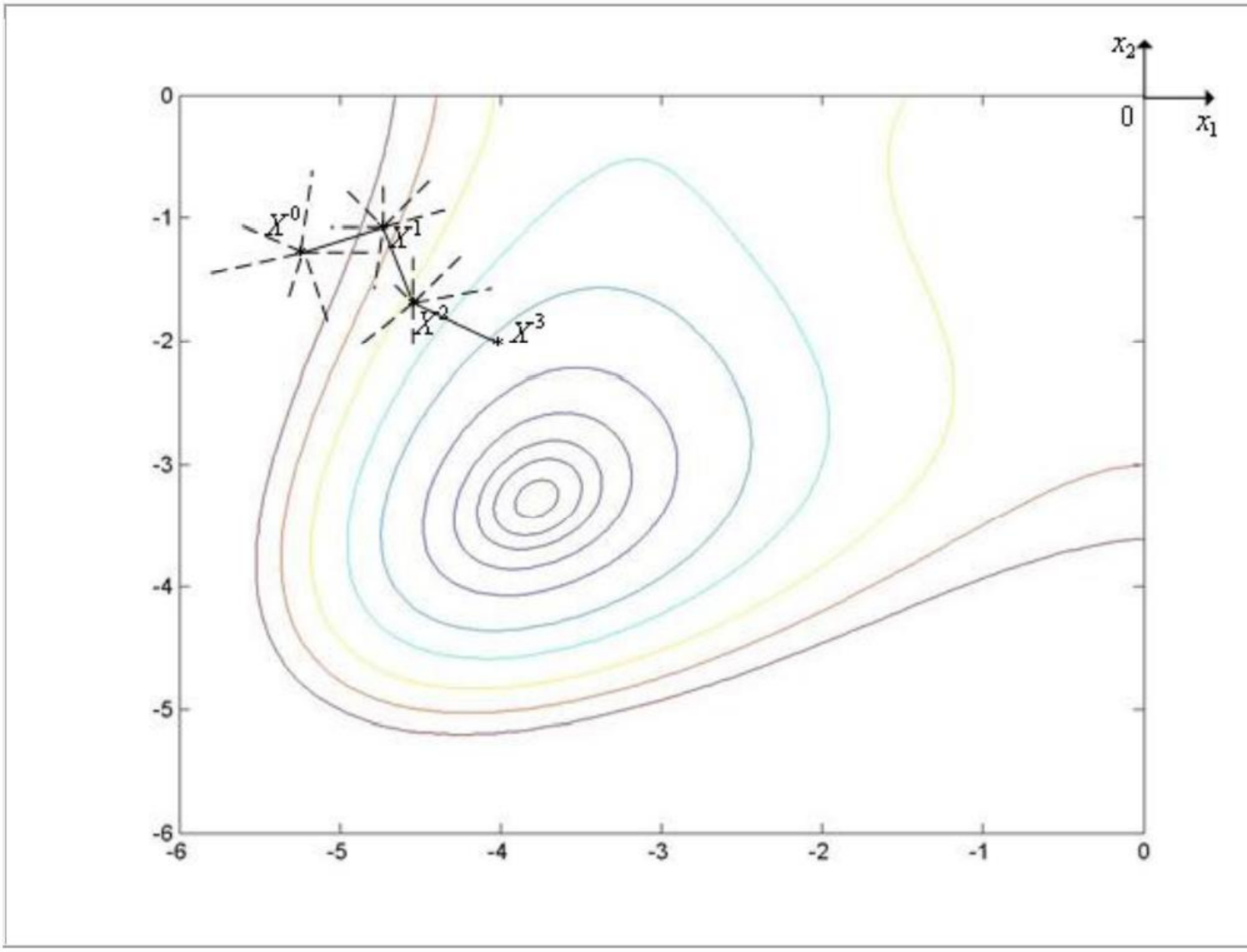


Иллюстрация метода наилучшей пробы на примере функции Химмельблау

11 Метод комплексов

Комплексом называется многогранник с N > n+1 вершинами (не обязательно выпуклый). Рекомендуется N=2n.

Вообще говоря, комплексом в комбинаторной топологии называется геометрическая фигура, которая может быть разбита на более элементарные фигуры. В нашем случае такими элементарными фигурами являются симплексы (**симплициальные комплексы**).

При решении методом комплексов используются следующие операции:

- генерация случайного комплекса;
- отражение вершины комплекса с растяжением;
- сжатие комплекса.

Генерация случайного комплекса

В пространстве Rn координаты вершин случайного комплекса с N вершинами могут быть найдены по формуле

$$\overline{X}_i = \overline{X}_0 + l\,\frac{\overline{\psi}_i}{\left\|\overline{\psi}_i\right\|}, \quad i \in [1,\ldots,N]$$

Где $\overline{X}_0$ -- произвольная начальная точка

i – номер вершины комплекса

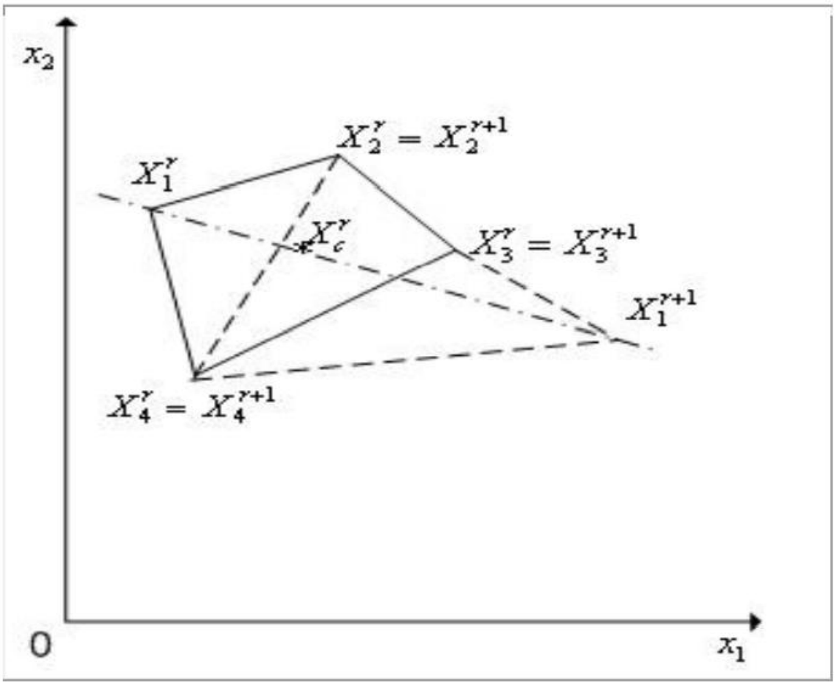
l – скаляр, определяющий размеры комплекса

$\overline{\psi}_i = (\psi_1, \psi_2, \ldots, \psi_n)$ - случайный вектор, компоненты которого независимые случайные величины, равномерно распределенные на интервале [-1,1].

Отражение вершины комплекса с растяжением

Пусть в пространстве Rn задан комплекс Cr с N вершинами и его вершину Xk необходимо отразить через центр тяжести комплекса с растяжением.

В новом комплексе Cr+1 все вершины, кроме k -ой, совпадают с соответствующими вершинами исходного комплекса Cr , а k -я вершина находится на прямой, проходящей через центр тяжести этого комплекса и его вершину Xk.



Тогда координаты вершин нового комплекса

$$\overline{X}_i^{r+1} = \overline{X}_i^r \ , \quad i \in [0,\ldots,N], \quad i \neq k$$

$$\overline{X}_k^{r+1} = \overline{X}_c^r + \alpha\left(\overline{X}_c^r - \overline{X}_k^r\right)$$

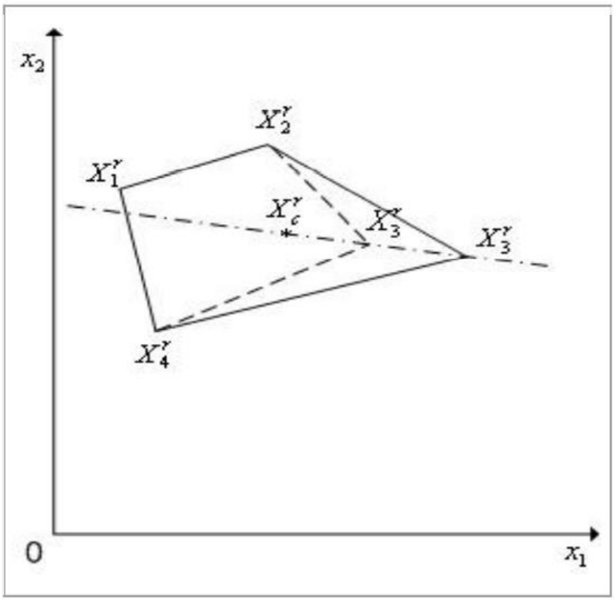
Где α - коэффициент растяжения комплекса, $\overline{X}_c^r$ - координаты центра тяжести комплекса Cr

$$\overline{X}_c^r = \frac{1}{N}\sum_{i=1}^{N}\overline{X}_i^r$$

Сжатие комплекса

Пусть в пространстве Rn задан комплекс Cr с N вершинами, и его вершину $\overline{X}_k^r$ необходимо переместить ближе к центру тяжести.

В новом комплексе Cr+1 все вершины, кроме k -ой, совпадают с соответствующими вершинами исходного комплекса Cr , а k -я вершина находится на прямой, проходящей через центр тяжести этого комплекса и его вершину $\overline{X}_k^r$.



Координаты вершин нового комплекса:

$$\overline{X}_i^{r+1} = \overline{X}_i^r \ , \quad i \in [0,\ldots,N], \quad i \neq k$$
$$\overline{X}_k^{r+1} = \overline{X}_c^r + \beta\left(\overline{X}_k^r - \overline{X}_c^r\right)$$

Где β - коэффициент растяжения комплекса (рекомендуемое значение $\beta = 2$), $\overline{X}_c^r$ - координаты центра тяжести комплекса Cr

Схема простейшего варианта метода комплексов

1. Задаем начальную точку $\overline{X}_0^0$, исходя из которой должен быть построен комплекс C0 , начальное значение величины $l = l_0$ и полагаем счетчик числа итераций r =0.

2. Генерируем N случайных векторов $\overline{\psi}_i^{\,r}$, i=1,…N и находим координаты вершин комплекса Cr

$$.\ \overline{X}_i = \overline{X}_0 + l\,\frac{\overline{\psi}_i}{\left\|\overline{\psi}_i\right\|}\ , \quad i \in [1,\ldots,N]$$

3. Вычисляем значения Ф(Xi) во всех вершинах комплекса Cr .

4. Находим максимальное из значений Ф(X) в вершинах комплекса Cr

$$\Phi\left(\overline{X}_k^r\right) = \max \Phi\left(\overline{X}_i^{r1}\right)$$

5. Выполняем отражение вершины комплекса $\overline{X}_k^r$ с растяжением r k X – получаем вершину $\overline{X}_k^{r+1}$ и новый комплекс Cr+1.

6. Вычисляем $\Phi\left(\overline{X}_k^{r+1}\right)$

7. Если $\Phi\left(\overline{X}_k^{r+1}\right) < \Phi\left(\overline{X}_k^{r}\right)$, то полагаем r =r +1 и переходим к п.8.

Иначе – выполняем сжатие симплекса Cr+1 в направлении $\overline{X}_k^{r+1} - \overline{X}_c^{r+1}$, полагаем r =r+2 и переходим к п.6.

8. Проверяем условие окончания поиска.

Если выполнено, то в качестве точки X∗ полагаем вершину комплекса Cr , к которой функция Ф(X) имеет наименьшее значение и завершаем итерации.

Иначе – переходим к п. 4.

В качестве критерия окончания поиска может использоваться одно из условий:

максимальная длина ребра комплекса Cr не превышает *epsx* – требуемую точность решения по X .

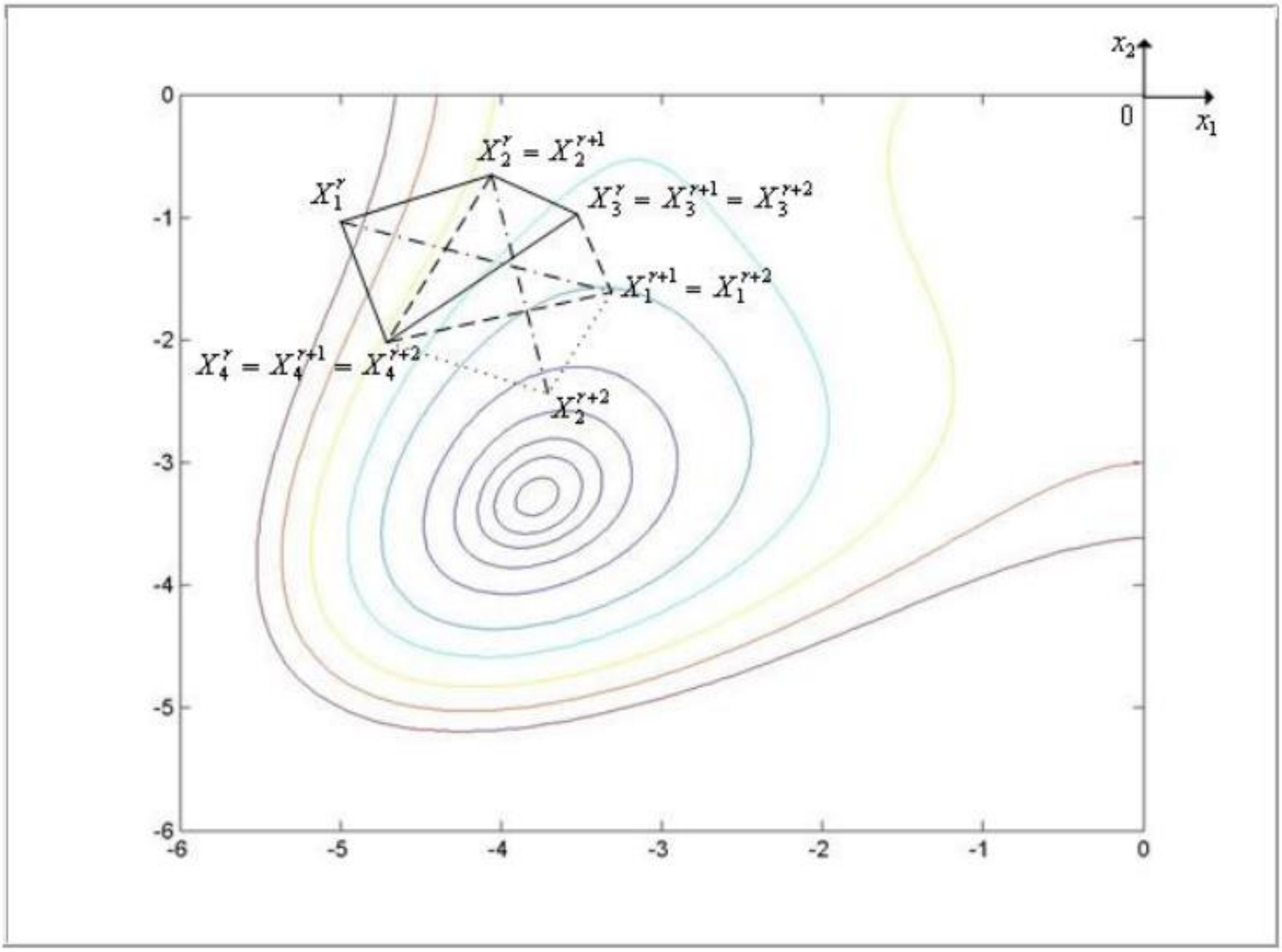
максимальная разность значений функции Ф(X) в двух вершинах комплекса Cr не превышает epsФ – требуемую точность решения по Ф.

Могут использоваться также более сложные условия окончания поиска:

$$\frac{1}{N}\sqrt{\sum_{i=1}^{N}\left\|\overline{X}_c^r - \overline{X}_i^r\right\|} \le epsx$$

$$\frac{1}{N}\sqrt{\sum_{i=1}^{N}\left|\Phi_{cp}^r - \Phi\left(\overline{X}_i^r\right)\right|} \le eps\Phi$$

Где $\Phi_{cp}^r = \dfrac{1}{N}\sum_{i=1}^{N}\Phi\left(\overline{X}_i^r\right)$

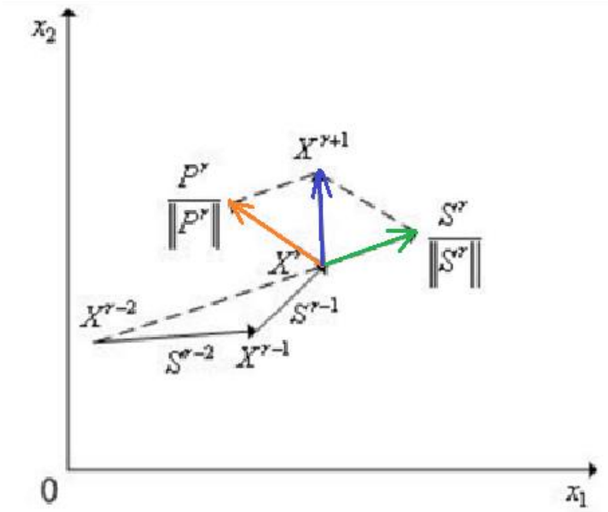


Известно множество модификаций рассмотренного метода комплексов, направленных, в частности, на преодоление «уплощения» комплекса в процессе поиска. С этой целью через фиксированное количество итераций находятся максимальная и минимальная диагонали комплекса и, если их отношение превышает заданное, то по рассмотренной схеме производится построение нового комплекса.

12 Метод повторяющегося случайного поиска (3-шаговый метод)

Найти минимум Ф

$$\min_{\overline{X}\in R^n} \Phi(\overline{X}) = \Phi(\overline{X}^*) = \Phi^*$$



Итерационная формула:

$\overline{X}^0$ - начальное приближение решения

$$\overline{X}^{r+1} = \overline{X}^r + \lambda^r \overline{\Delta}^r,$$

где $\overline{\Delta}^r = \underbrace{\beta \frac{\overline{S}^r}{\|\overline{S}^r\|}}_{\text{«историческая» компонента}} + \underbrace{(1-\beta)\frac{\overline{P}^r}{\|\overline{P}^r\|}}_{\text{случайная компонента}}$

$\overline{P}^r$ - вектор, компоненты которого – случайные числа, равномерно распределенные на [0,1]

$\overline{S}^r = \gamma \overline{S}^{r-1} + (1-\gamma)\overline{S}^{r-2}$ - вектор «предыстории» (средневзвешенное направление поиска на двух предыдущих шагах), $\gamma \in [0,1]$

β - относительный вес детерминированной и случайной компонент направления

Например, при $\gamma = \beta = 0.5$ и $\lambda^r = 2$

$$\overline{S}^r = 0.5\left(\overline{S}^{r-1} + \overline{S}^{r-2}\right)$$

И соотв. итерационная схема

$$\overline{X}^{r+1} = \overline{X}^{r+1} + \left(\frac{\overline{S}^r}{\|\overline{S}^r\|} + \frac{\overline{P}^r}{\|\overline{P}^r\|}\right)$$

Упрощенная схема метода повторяющегося случайного поиска

1. Задаем:

начальную точку $\overline{X}^0$,

Полагаем счетчик числа итераций r = 2, задаем начальный шаг λ^r, значения коэффициентов β,γ

2. «Разгон»: тем или иным способом, например, с помощью одношагового метода наилучшей пробы определяем точки $\overline{X}^1$ и $\overline{X}^2$.

3. Генерируем n-мерный случайный вектор $\overline{P}^r$ и вычисляем

$$\overline{X}^{r+1} = \overline{X}^r + \lambda^r \overline{\Delta}^r,$$

где $\overline{\Delta}^r = \beta \frac{\overline{S}^r}{\left\|\overline{S}^r\right\|} + (1-\beta)\frac{\overline{P}^r}{\left\|\overline{P}^r\right\|}$

$$\overline{S}^r = \gamma \overline{S}^{r-1} + (1-\gamma)\overline{S}^{r-2}$$

4. Если $\Phi\left(\overline{X}^{r+1}\right) < \Phi\left(\overline{X}^r\right)$, то проверяем условие окончания итераций (см. ниже). Если выполнено, то полагаем X ∗ ≈ Xr+1 и завершаем итерации.

Если условие окончания итераций не выполнено, то некоторому правилу увеличиваем длину шага (например, $\lambda^{r+1} = 2\lambda^r$), принимаем r =r +1 и переходим к п.3.

Если $\Phi\left(\overline{X}^{r+1}\right) \geq \Phi\left(\overline{X}^r\right)$, то переходим к п. 5.

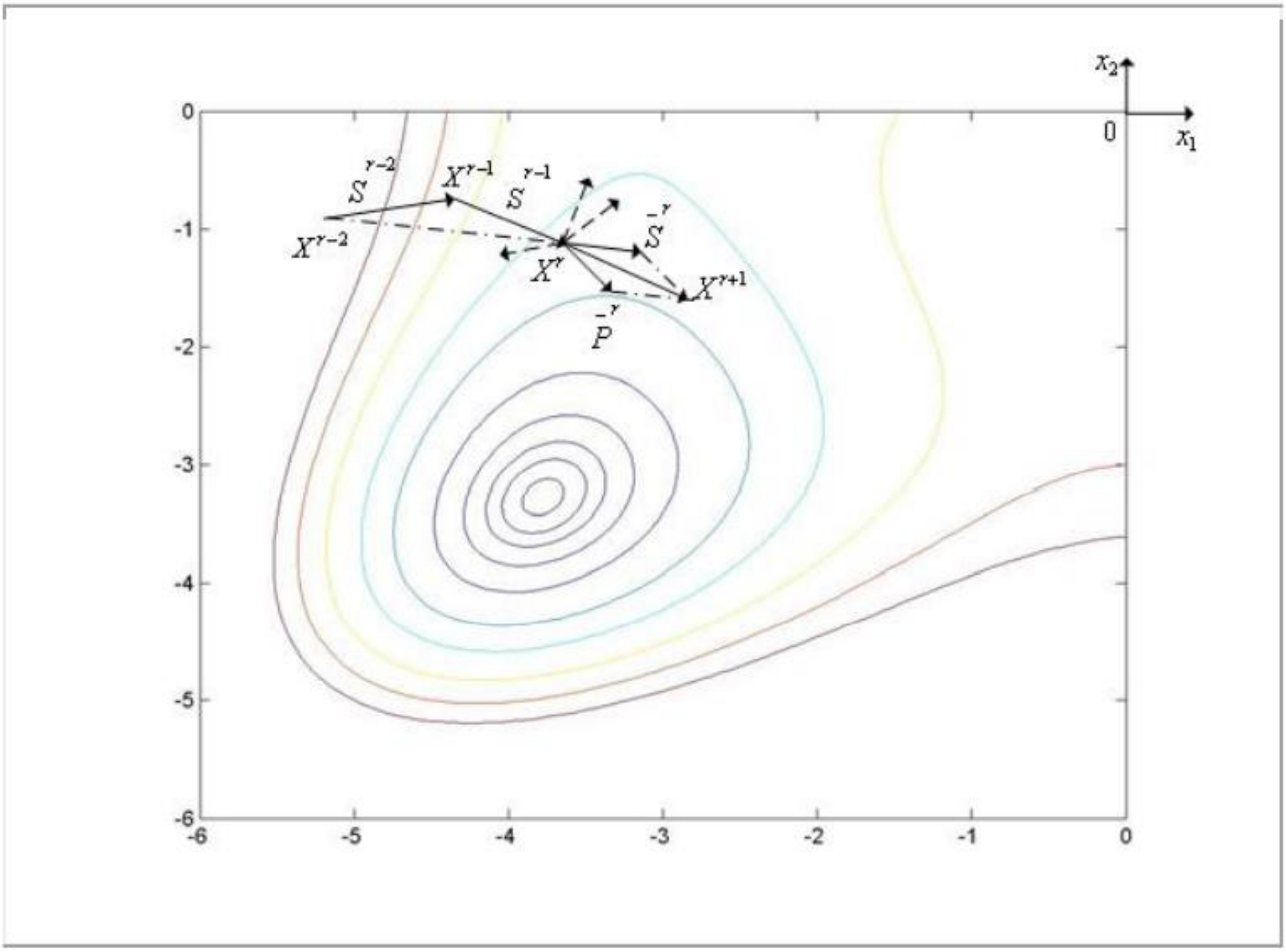
5. Некоторое фиксированное количество раз делаем попытку, исходя из той же точки Xr, не меняя длины шага λr , добиться уменьшения значения функции Φ(X) путем только изменения вектора Pr, т.е., не меняя Xr и λr , переходим на п. 3.

Если это фиксированное количество попыток не привело к успеху, то, исходя из той же точки Xr, по некоторому правилу уменьшаем длину шага λr (например, $\lambda^{r+1} = 0{,}5\lambda^r$), , и переходим к п.3.

В качестве условия окончания поиска можно использоваться одно из стандартных условий окончания итераций:

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

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



Известно множество модификаций рассмотренной простейшей схемы метода повторяющегося случайного поиска. Например, в процессе поиска могут изменяться по некоторым правилам не только длина шага λ , но и коэффициенты β,γ .

13 Метод случайного поиска с постоянным радиусом поиска и случайными направлениями

Метод случайного поиска с постоянным радиусом поиска и случайными направлениями использует процедуру генерации случайных точек, равномерно распределенных по поверхности гиперсферы в пространстве Rn .

Пусть

$\overline{X}^c = \left(x_1^c, x_2^c, \ldots, x_n^2\right)$– вектор координат центра гиперсферы,

ρ – радиус гиперсферы,

R – вектор с началом в точке $\overline{X}^c$ и концом в искомой точке на поверхности гиперсферы,

α_i , $i \in [1,..., n]$– углы между вектором R и ортами координатных осей

Генерация случайных точек, равномерно распределенных по поверхности гиперсферы радиуса ρ:

- генерируем *n* случайных чисел, равномерно распределенных в интервале [0,2π);
- вычисляем направляющие косинусы $\cos\alpha_i$, $i \in [1,..., n]$вектора R;
- находим координаты искомой точки $x_i = x_i^c + \rho\cos\alpha_i$, $i \in [1,\ldots,n]$.

Упрощенная схема метода случайного поиска с постоянным радиусом поиска и случайными направлениями

1. Задаем начальную точку X0 , начальный радиус гиперсферы ρ0 , и полагаем счетчик числа итераций r =0.

2. Генерируем N случайных точек $\overline{X}_i^{\,r}$, $i \in [1,\ldots,N]$, равномерно распределенных по поверхности гиперсферы радиуса ρr с центром в точке Xr.

Здесь N – количество точек – свободный параметр метода.

3. Вычисляем значения Φ(X) в полученных точках и находим точку $\overline{X}_k^{\,r}$, в которой достигается минимальное значение функции Φ

4. Каким-либо из рассмотренных методов одномерной оптимизации находим минимум функции Φ в направлении $\left(\overline{X}_k^{\,r} - \overline{X}^{\,r}\right)$:

$$\Phi\left(\overline{X}^{\,r+1}\right) = \min_{\lambda} \Phi\left(\overline{X}^{\,r} + \lambda\left(\overline{X}_k^{\,r} - \overline{X}^{\,r}\right)\right)$$

5. Проверяем условие окончания итераций (см. ниже). Если выполнено, то полагаем X∗≈ Xr+1 ≈ и завершаем итерации.

Иначе – принимаем r = r +1 и переходим к п.2.

Условия окончания поиска:

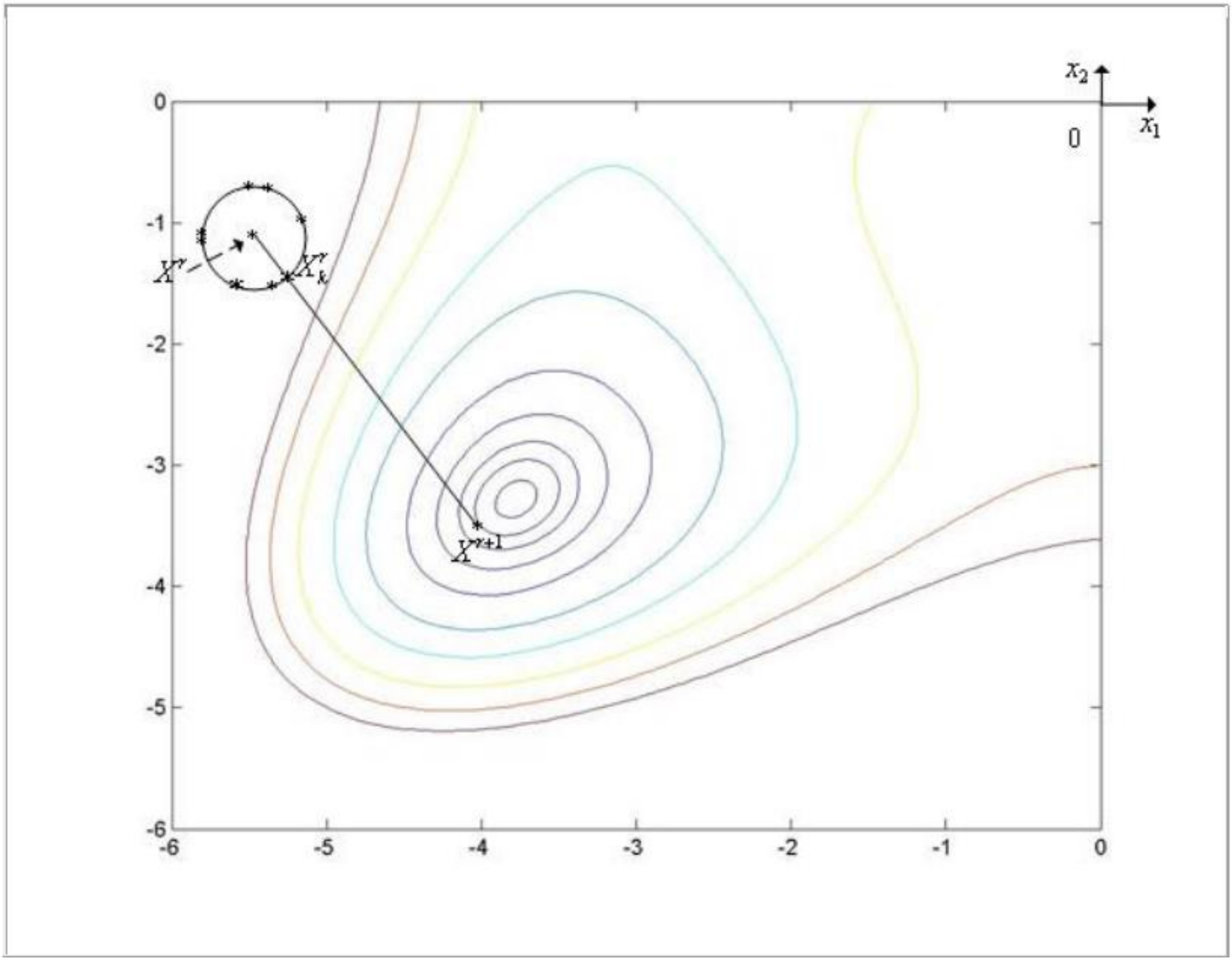
$$\left\|\overline{X}^{r+1}-\overline{X}^{r}\right\| \le epsx$$

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

Могут быть использованы также другие критерии окончания поиска, например, условие не превышения текущим радиусом гиперсферы величины *epsx* :

$$\rho^{r} \le epsx$$

В процессе поиска радиус гиперсферы может меняться, увеличиваясь при удачных шагах (вдали от точки X*) и уменьшаясь при неудачных шагах (вблизи от точки X*).
Поиск может быть ускорен, если точки на гиперсфере выбирать (случайным образом) в некотором секторе по отношению к предыдущему направлению. Угол раскрыва этого сектора может меняться в процессе поиска.



Примечание

Одна итерация по методу случайного поиска с постоянным радиусом поиска и случайными направлениями может привести к уменьшению минимизируемой функции в большей степени, чем один шаг поиска в направлении антиградиента этой функции (см рис)

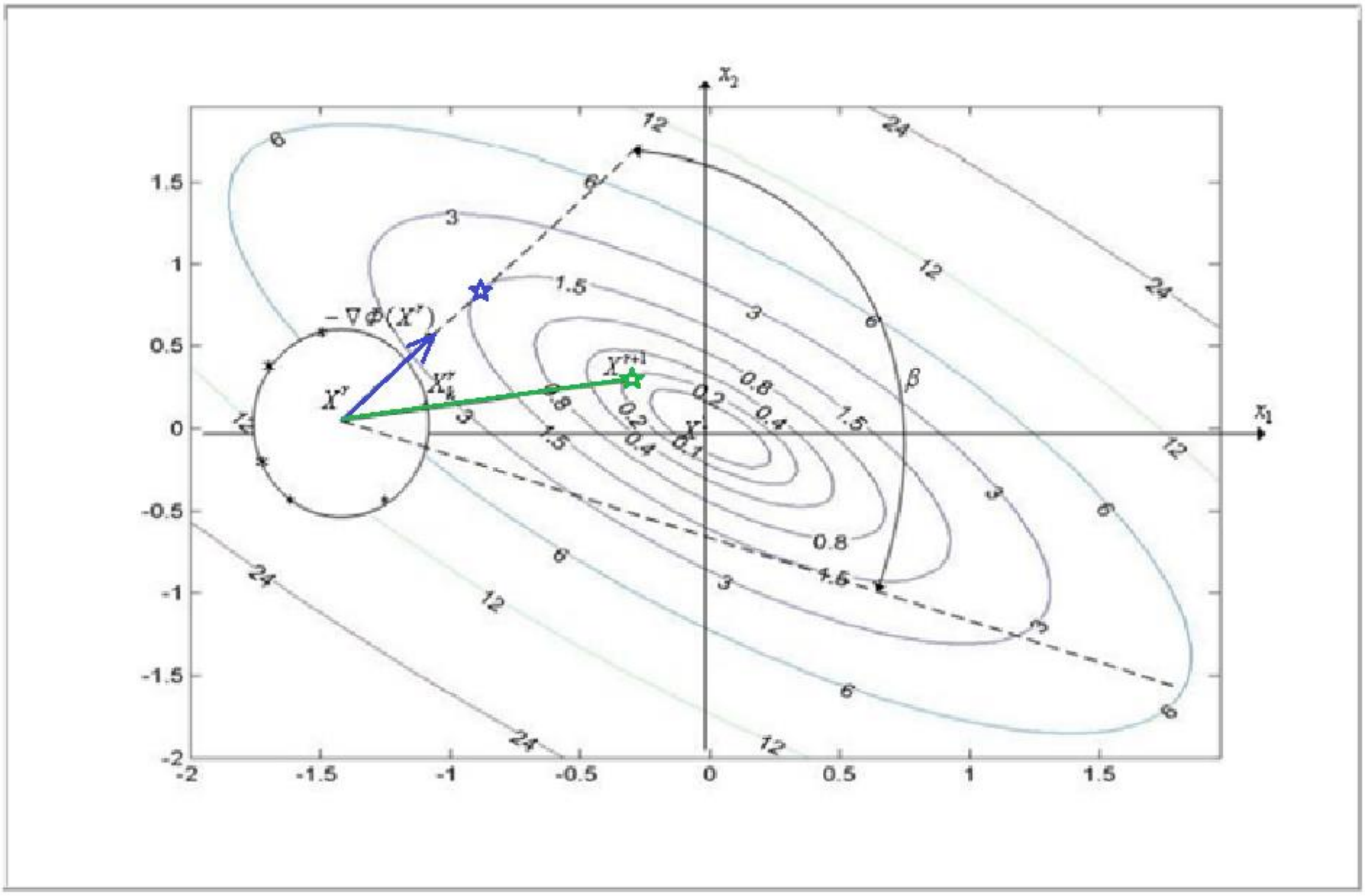


Иллюстрация: любое направление поиска в секторе β лучше, чем направление антиградиента