

Методы минимизации функции многих переменных

Классификация методов

$$f(\vec{x}) = f(x_1, x_2, \dots, x_n)$$

градиент $\vec{g}(\vec{x}) = \begin{bmatrix} \frac{\partial f(\vec{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\vec{x})}{\partial x_n} \end{bmatrix}$

Матрица Гессе $G_{ij} = \frac{\partial^2 f(\vec{x})}{\partial x_i \partial x_j} \quad i, j = \overline{1, n}$

Необходимое и достаточное условие существования минимума $f(\vec{x})$ в $\vec{x}_0$

1. $\vec{g}(\vec{x}_0) = 0$ т.е. $\frac{\partial f(\vec{x}_0)}{\partial x_i} = 0 \quad i = \overline{1, n}$

2. $G(\vec{x}_0)$ - положительно определена

т.е. $\frac{\partial^2 f(\vec{x}_0)}{\partial x_i^2} > 0$



Регулярные методы:

1. Покоординатные (нулевого порядка)

в $\vec{x}_k \quad f(\vec{x}_k)$

2. Градиентные (первого порядка)

в $\vec{x}_k \quad f(\vec{x}_k) \text{ и } \vec{g}(\vec{x}_k)$

3. Второго порядка

в $\vec{x}_k \quad f(\vec{x}_k) \quad \vec{g}(\vec{x}_k) \quad G_k = \left[\frac{\partial^2 f(\vec{x}_k)}{\partial x_i \partial x_j} \right] \quad i, j = \overline{1, n}$

Покоординатные методы

Идея методов: $f(\bar{x}) \rightarrow \min$ $\bar{x} = \begin{bmatrix} x_1 \\ \dots \\ x_n \end{bmatrix}$

Начальное приближение $\bar{x}_0 = \begin{bmatrix} x_{10} \\ x_{20} \\ \dots \\ x_{n0} \end{bmatrix}$

1. $x_{10} = \text{var}$ $x_{20}, x_{30}, \dots, x_{n0} = \text{const}$

$$f(x_{10}) \rightarrow \min \quad \bar{x}_1 = \begin{bmatrix} x_{11} \\ x_{20} \\ \dots \\ x_{n0} \end{bmatrix}$$

2. $x_{20} = \text{var}$ $x_{11}, x_{30}, \dots, x_{n0} = \text{const}$

$$f(x_{20}) \rightarrow \min \quad \bar{x}_2 = \begin{bmatrix} x_{11} \\ x_{21} \\ \dots \\ x_{n0} \end{bmatrix} \quad \text{и т.д.}$$

Получают последовательность приближений
 $\bar{x}_0 \quad \bar{x}_1 \quad \bar{x}_2 \quad \dots$

$$\text{с} \quad f(\bar{x}_0) \geq f(\bar{x}_1) \geq f(\bar{x}_2) \geq \dots$$

покоординатные методы:

1. Метод Гаусса-Зейделя

2. Метод Розенброка

3. Метод ДСК

4. Метод Пауэлла

5. Симплексный метод Нельдера-Мида

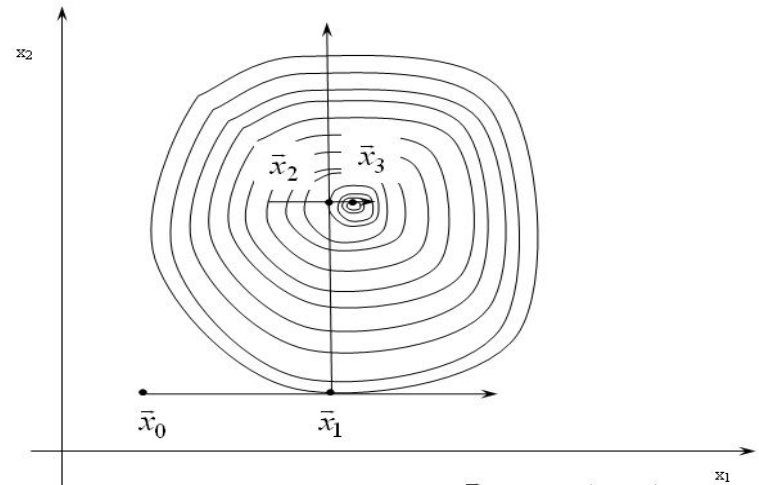
Метод Гаусса-Зейделя

Поочередно по каждой переменной x_i -
поиск частного минимума $f(\bar{x})$

т.е. решение уравнения $\frac{\partial f(\bar{x})}{\partial x_i} = 0$

Учет ограничений на переменные:

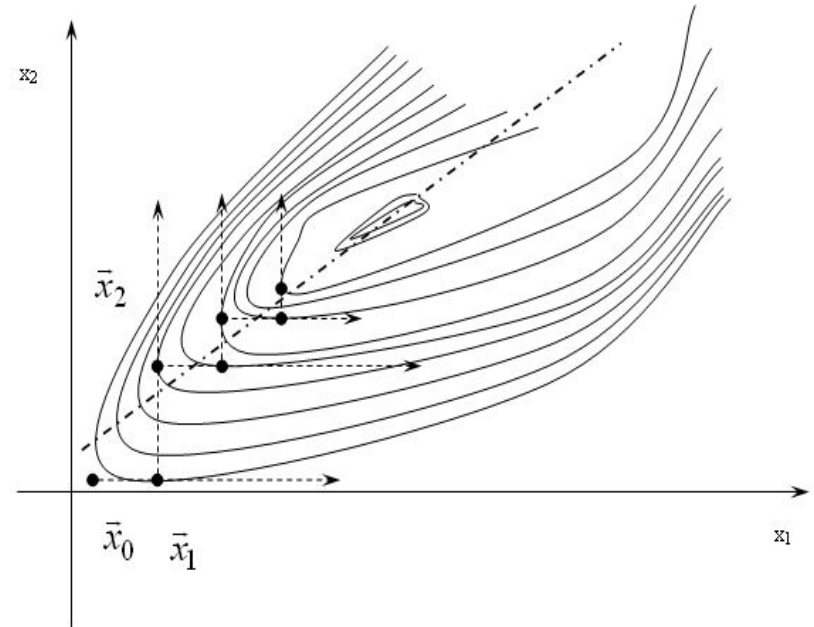
1. движение по x_i только внутри заданной области σ
2. при достижении границы области σ поиск частного минимума по данной x_i прекращается



$$\bar{x}_k = \bar{x}_{k-1} + \alpha_k \bar{P}_k$$

$$\bar{P}_1 = \bar{e}_1 = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$\bar{P}_2 = \bar{e}_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}$$



Метод Розенброка

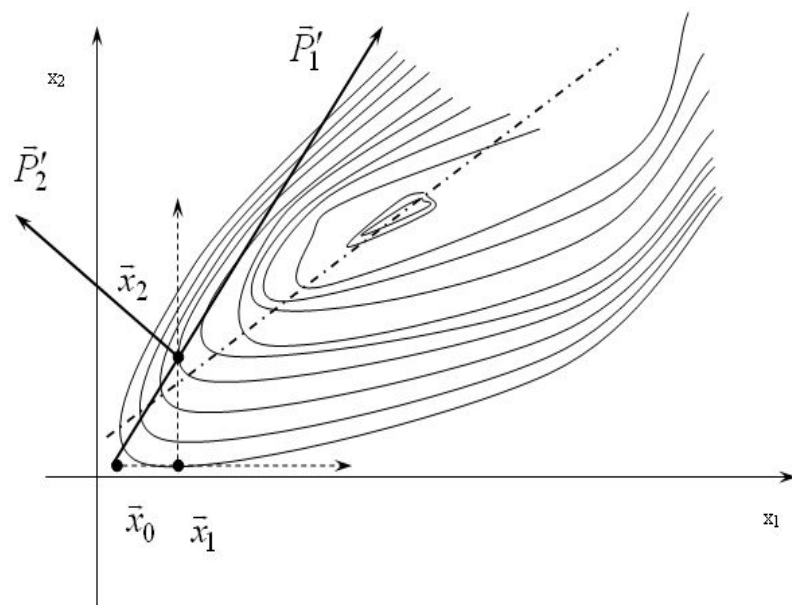
Идея: смена системы координат с

$x_1, x_2, \dots, x_n$ на $x'_1, x'_2, \dots, x'_n$

$\bar{x}'_1$ - направление полного перемещения на первой итерации (часто близка к оси оврага)

$\bar{P}_j$ - направление перемещения на предыдущей итерации

$\bar{P}'_j$ - направление перемещения на следующей итерации



α_j - продвижение по каждому направлению $\bar{P}_j$

$$\bar{a}_i = \sum_{j=1}^n \alpha_j \bar{P}_j \quad i = \overline{1, n}$$

$$\bar{b}_1 = \bar{a}_1 \quad \bar{b}_i = \bar{a}_i - \sum_{j=1}^{i-1} (\bar{a}_i \bar{P}'_j) \bar{P}'_j \quad i = \overline{2, n}$$

процедура ортогонализации
Грамма-Шмидта

$$\bar{P}'_1 = \frac{\bar{b}_1}{\|\bar{b}_1\|} \quad \bar{P}'_j = \frac{\bar{b}_j}{\|\bar{b}_j\|} \quad j = \overline{2, n}$$

$\bar{P}'_1$ - направление наибыстрейшего продвижения на предыдущей итерации

$\bar{P}'_2$ - наилучшее направление, ортогональное $\bar{P}'_1$

$\bar{P}'_3$ - наилучшее направление, ортогональное $\bar{P}'_1$ и $\bar{P}'_2$

и т.д.

Для квадратичной ЦФ в $\bar{x}^*$ система направлений метода Розенброка соответствует направлениям осей функции $\rightarrow$ возможен поиск *глобального минимума*

Метод Пауэлла

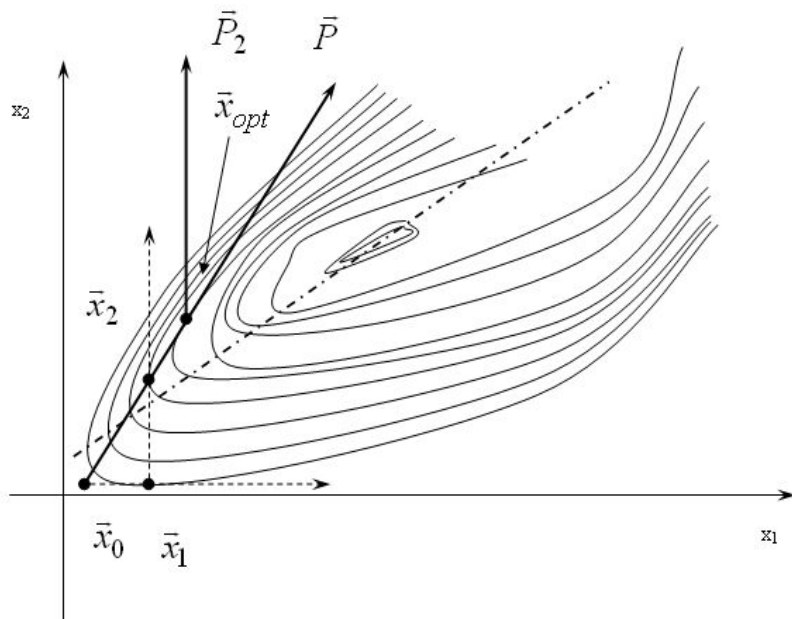
Алгоритм метода:

1. по координатный спуск методом Гаусса-Зейделя по выбранным ортогональным направлениям $\bar{P}_1, \bar{P}_2, \dots, \bar{P}_n$ (переход от $\bar{x}_0$ к $\bar{x}_1, \dots, \bar{x}_n$)

2. Вычисление нового направления

$$\bar{P} = \bar{x}_n - \bar{x}_0 = \sum_{j=1}^n \alpha_j \bar{P}_j$$

3. Поиск в направлении $\bar{P}$ частного минимума ЦФ $f(\bar{x})$ $\bar{x}_{n+1} = \bar{x}_{opt}$



4. Цикл минимизации от $\bar{x}_{opt}$ по направлениям

$$\bar{P}_2, \bar{P}_3, \dots, \bar{P}_n, \bar{P}$$

Направления поиска – «веер» вдоль оси оврага

Правило для исключения образования в процессе замены линейно-зависимой системы векторов $\bar{P}_j$:

замена $\bar{P}_1$ на $\bar{P}$ запрещена при невыполнении условий одновременно

1. $f(\bar{x}_0) > f(\bar{x}^1)$
2. $\left[f(\bar{x}^1) - 2f(\bar{x}_{opt}) + f(\bar{x}_0) \right] \left[f(\bar{x}^1) - f(\bar{x}_{opt}) - \Delta \right]^2 > \frac{1}{2} \Delta \left[f(\bar{x}^1) - f(\bar{x}_0) \right]^2$

где $\bar{x}^1 = 2\bar{x}_{opt} - \bar{x}_0$, Δ - максимальное уменьшение значения ЦФ по направлению за цикл минимизации ($\bar{x}_0$ и $\bar{x}_{opt}$ - начальная и конечная точки цикла)

Симплексный метод Нельдера-Мида

Идея: построение нерегулярного симплекса

метод эффективен при $n \leq 6$

Алгоритм метода:

1. выбор вершин симплекса

$$\bar{x}_0 \quad \bar{x}_1 \quad \bar{x}_2 \quad \dots \quad \bar{x}_n$$

2. вычисление значений ЦФ в

вершинах симплекса

$$f_0 \quad f_1 \quad f_2 \quad \dots \quad f_n$$

3. определение

$$f_h = \max_i f_i = f(\bar{x}_h)$$

$$f_l = \min_i f_i = f(\bar{x}_l)$$

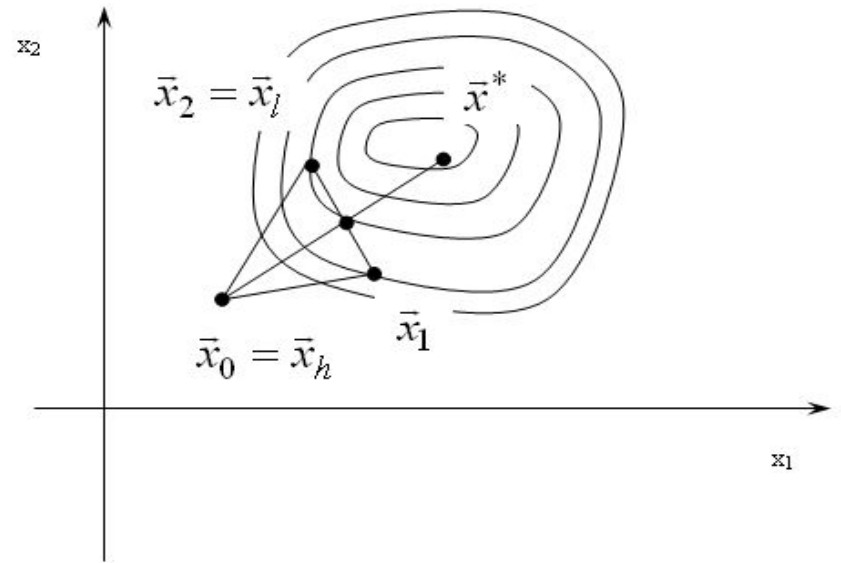
4. расчет центра тяжести точек (кроме $\bar{x}_h$)

$$\bar{\bar{x}} = \frac{1}{n} \sum_{i \neq h} \bar{x}_i \quad \bar{f} = f(\bar{\bar{x}})$$

5. перемещение $\bar{x}_h$

а) операция отражения $\bar{x}_h \rightarrow \bar{x}^*$

$$\bar{x}^* = (1 + \alpha)\bar{\bar{x}} - \alpha\bar{x}_h \text{ и } f^* = f(\bar{x}^*)$$



где $\alpha = \frac{|\bar{x}^* - \bar{\bar{x}}|}{|\bar{\bar{x}} - \bar{x}_h|} > 0$ - коэффициент отражения

Возможные ситуации:

1) $f_l < f^* < f_h$ - замена $\bar{x}_h$ на $\bar{x}^*$ и построение нового симплекса (выбор f_h f_l $\bar{f}$, отражение)

Симплексный метод Нельдера-Мида

2) $f^* < f_l$ - операция растяжения

$$\bar{x}^{**} = \gamma \bar{x}^* + (1 - \gamma) \bar{x} \quad f^{**} = f(\bar{x}^{**})$$

где $\gamma = \frac{|\bar{x}^{**} - \bar{x}|}{|\bar{x}^* - \bar{x}|}$ - коэффициент растяжения

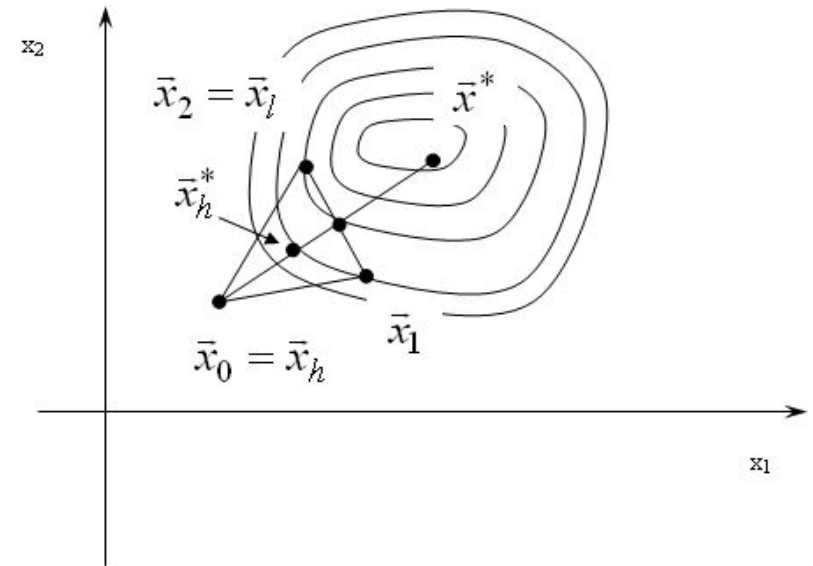
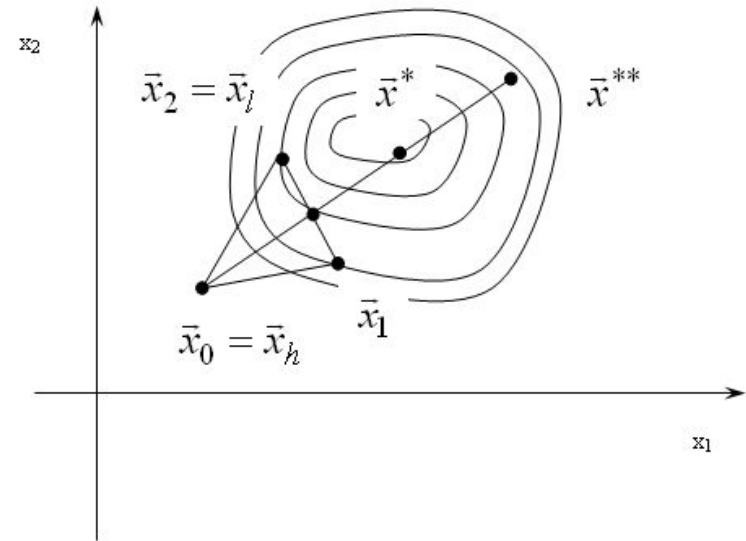
При $f^{**} < f_l$ - замена $\bar{x}_h$ на $\bar{x}^{**}$ и исследование $\bar{x}^{**}$ на сходимость к минимуму, если сходимость не достигнута – построение нового симплекса (выбор $f_h f_l \bar{f}$, отражение)

При $f^{**} \geq f_l$ - $\bar{x}^{**}$ отбрасывают, заменяют $\bar{x}_h$ на $\bar{x}^*$, проверяют $\bar{x}^*$ на сходимость к минимуму, если сходимость не достигнута – построение нового симплекса (выбор $f_h f_l \bar{f}$, отражение)

3) $f^* > f_h$ - операция сжатия

$$\bar{x}_h^* = \beta \bar{x}_h + (1 - \beta) \bar{x} \quad f_h^* = f(\bar{x}_h^*)$$

где $\beta = \frac{|\bar{x} - \bar{x}_h^*|}{|\bar{x} - \bar{x}_h|}$ $0 < \beta < 1$ коэффициент сжатия



Симплексный метод Нельдера-Мида

При $f_h^* < f_h$ - замена $\bar{x}_h$ на $\bar{x}_h^*$ и построение нового симплекса (выбор $f_h f_l \bar{f}$, отражение)

При $f_h^* > f_h$ и $f_h^* > f^*$ - уменьшение размера симплекса $\bar{x}_i = \frac{1}{2}(\bar{x}_i + \bar{x}_l)$, расчет $f(\bar{x}_i)$, проверка на сходимость $\bar{x}_l$ к минимуму, если сходимость не достигнута, то построение нового симплекса (выбор $f_h f_l \bar{f}$, отражение)

Проверка сходимости

$$\sigma^2 = \sum_{i=0}^n \frac{(f_i - \bar{f})^2}{n+1}, \quad \text{где } \bar{f} = \sum_{i=0}^n \frac{f_i}{n+1}$$

При погрешности определения минимума ε :
если $\sigma < \varepsilon$

то $\bar{x}_l$ - точка минимума

Рекомендации Нельдера и Мида по выбору коэффициентов отражения, растяжения и сжатия

$$\alpha = 1 \quad \gamma = 2 \quad \beta = 0.5$$

Градиентные методы

- Идея: выбор «наилучшего» направления спуска
- Градиент – направление наискорейшего возрастания ЦФ
- Антиградиент – «наилучшее» направление к точке минимума
- Градиентные методы
- 1. Метод наискорейшего спуска
- 2. Метод сопряженных градиентов
- 3. Семейство методов с переменной метрикой
- 4. Метод оврагов

Метод наискорейшего спуска

Основные операции:

1. определение направления градиента ЦФ в $\bar{x}_k$

$$\bar{x}_k = \begin{bmatrix} x_{k1} \\ x_{k2} \\ \dots \\ x_{kn} \end{bmatrix} \quad \bar{g}_k = g(\bar{x}_k) = \begin{bmatrix} \partial f(\bar{x}_k) / \partial x_{k1} \\ \partial f(\bar{x}_k) / \partial x_{k2} \\ \dots \\ \partial f(\bar{x}_k) / \partial x_{kn} \end{bmatrix}$$

2. Перемещение в направлении антиградиента

$$\begin{cases} \bar{P}_k = -\bar{g}_k \\ \bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{P}_k \end{cases}$$

Расчет шага α_k в направлении антиградиента по условию частного минимума

$$f(\bar{x}_k + \alpha_k \bar{P}_k) = \min_{\alpha} f(\bar{x}_k + \alpha \bar{P}_k)$$

Способ расчета α_k с минимальными вычислительными затратами – условие

минимума по α типа

$$\varphi(\alpha) = \frac{\partial f(\bar{x}_k + \alpha \bar{P}_k)}{\partial \alpha} = 0$$

т.е. α_k - корень уравнения $\varphi(\alpha) = 0$

Способы решения $\varphi(\alpha) = 0$ - на основе квадратичного приближения $f(\bar{x})$ вблизи $\bar{x}_k$

Метод наискорейшего спуска

1. Способ сочетания метода секущих и квадратичной аппроксимации

а) поиск α_1 и α_3 по условию

$$\varphi_1 = \varphi(\alpha_1) < 0 \quad (\text{спуск})$$

$$\varphi_3 = \varphi(\alpha_3) > 0 \quad (\text{подъем})$$

б) линеаризация $\varphi(\alpha)$ - построение $L(\alpha)$

$$\begin{cases} \varphi_1 = k\alpha_1 + b \\ \varphi_3 = k\alpha_3 + b \end{cases}$$

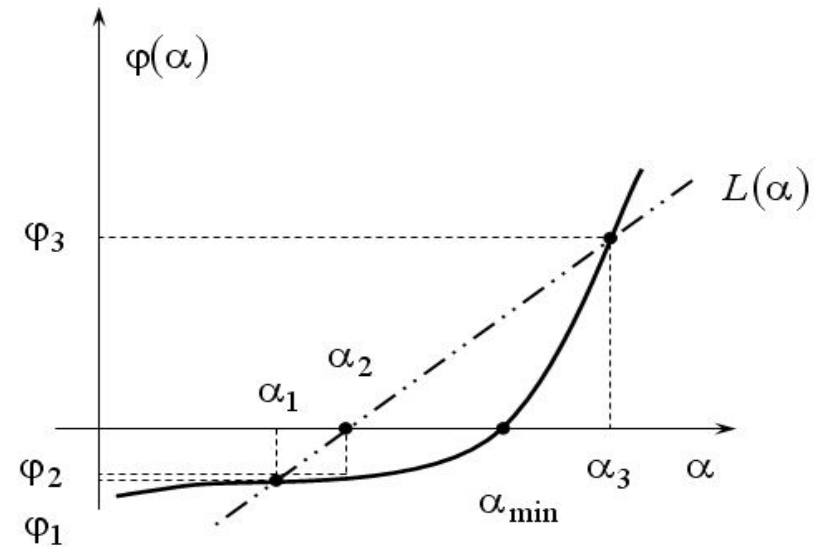
$$\text{т.е.} \quad L(\alpha) = \frac{\varphi_3 - \varphi_1}{\alpha_3 - \alpha_1} \alpha + \left(\varphi_1 - \frac{\varphi_3 - \varphi_1}{\alpha_3 - \alpha_1} \alpha_1 \right)$$

$$L(\alpha_2) = 0 \Rightarrow \alpha_2 = \frac{\varphi_3 \alpha_1 - \varphi_1 \alpha_3}{\varphi_3 - \varphi_1} \quad \varphi_2 = \varphi(\alpha_2)$$

в) по α_1 , α_2 и α_3 построение квадратичной аппроксимации функции $\varphi(\alpha)$ и определение корня $\alpha_{\min} = \alpha_k \quad (\alpha_k \in [\alpha_1, \alpha_3])$

$$\varphi(\alpha) \approx \psi(\alpha) = a\alpha^2 + b\alpha + c$$

$$\psi(\alpha_1) = \varphi_1 \quad \psi(\alpha_2) = \varphi_2 \quad \psi(\alpha_3) = \varphi_3$$



$$a = \frac{1}{\alpha_3 - \alpha_1} \left(\frac{\varphi_3 - \varphi_1}{\alpha_3 - \alpha_1} - \frac{\varphi_2 - \varphi_1}{\alpha_2 - \alpha_1} \right)$$

$$b = \frac{\varphi_2 - \varphi_1}{\alpha_2 - \alpha_1} - a(\alpha_2 + \alpha_1)$$

$$c = \varphi_1 - a\alpha_1^2 - b\alpha_1$$

$$\text{Из } \psi(\alpha) = 0 \quad - \quad \alpha_k = \alpha_{\min} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Метод наискорейшего спуска

2. Метод определения α_k по двум значениям $\varphi(\alpha)$

выбор α_1 и α_3 по условию

$$\varphi_1 = \varphi(\alpha_1) < 0 \quad (\text{спуск})$$

$$\varphi_3 = \varphi(\alpha_3) > 0 \quad (\text{подъем})$$

замечание: возможна ситуация $\alpha_1 = 0$

$$\alpha_k = \alpha_{\min} = \alpha_3 - \left(\frac{\varphi_3 + \varphi - z}{\varphi_3 - \varphi_1 - 2\varphi} \right) (\varphi_3 - \varphi_1)$$

где

$$z = 3 \frac{f(\alpha_1) - f(\alpha_3)}{\alpha_3 - \alpha_1} + \varphi_1 + \varphi_3$$
$$\varphi = \left(z^2 - \varphi_1 \varphi_3 \right)^{1/2}$$

3. Метод квадратичной интерполяции (полином Лагранжа)

Достоинство метода: не требует вычисления $\varphi(\alpha) = f'(\alpha)$

а) выбор α_1, α_2 и α_3 из условия

$$f_1 = f(\alpha_1) \quad f_2 = f(\alpha_2) \quad f_3 = f(\alpha_3)$$

$$f_1 > f_2 \quad \text{и} \quad f_2 < f_3$$

б) построение полинома Лагранжа

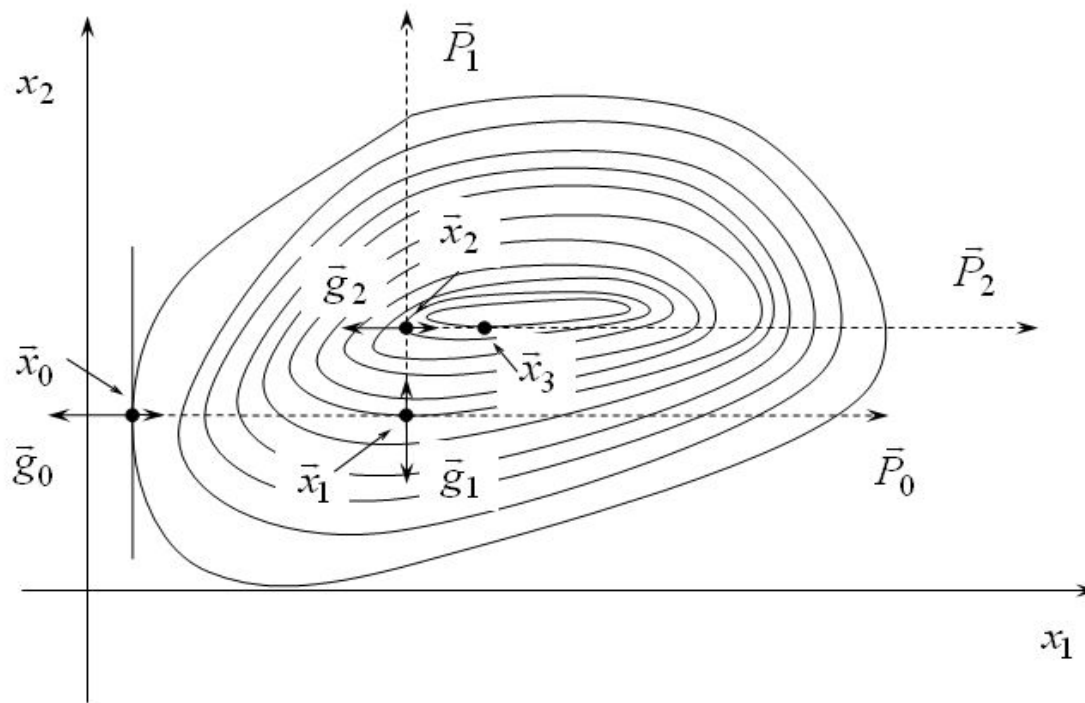
$$f(\alpha) \approx \bar{f}_1(\alpha - \alpha_2)(\alpha - \alpha_3) + \bar{f}_2(\alpha - \alpha_1)(\alpha - \alpha_3) + \bar{f}_3(\alpha - \alpha_1)(\alpha - \alpha_2)$$

где $\bar{f}_i = \frac{f_i}{(\alpha_i - \alpha_j)(\alpha_i - \alpha_k)} \quad k \neq i, j \neq i, k \neq j$

в) из условия $f'_\alpha(\alpha) = 0$

$$\alpha_k = \alpha_{\min} = \frac{\sum_{i=1}^3 \bar{f}_i (\alpha_j + \alpha_k)}{2 \sum_{i=1}^3 \bar{f}_i}$$

Геометрическая иллюстрация метода циклического спуска



- Для функций с «овражным» рельефом – спуск к оси оврага функции (как в методе Гаусса-Зейделя)
- Улучшение движения по антиградиенту – коррекция направления с учетом предыдущих направлений

Метод сопряженных градиентов

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{P}_k$$

где
$$\bar{P}_k = -\bar{g}_k + \beta_k \bar{P}_{k-1}$$

Величина β_k ?

1. $\bar{x}_0 \quad \bar{P}_0 = -\bar{g}_0 \rightarrow \bar{x}_1 = \bar{x}_0 + \alpha_0 \bar{P}_0$

2. $\bar{x}_1 \quad \bar{P}_{1cn} = -\bar{g}_1 + \beta_1 \bar{P}_0 \rightarrow \bar{x}_2 = \bar{x}_1 + \alpha_1 \bar{P}_{1cn}$
 ($\bar{P}_{1cn}$ сопряжено $\bar{P}_0$)

3. $\bar{x}_2 \quad \bar{P}_{2cn} = -\bar{g}_2 + \beta_2 \bar{P}_{1cn} + \beta_0 \bar{P}_0 \rightarrow \bar{x}_3 = \bar{x}_2 + \alpha_2 \bar{P}_{2cn}$
 ($\bar{P}_{2cn}$ взаимно сопряженное $\bar{P}_{1cn}$ и $\bar{P}_0$)

...

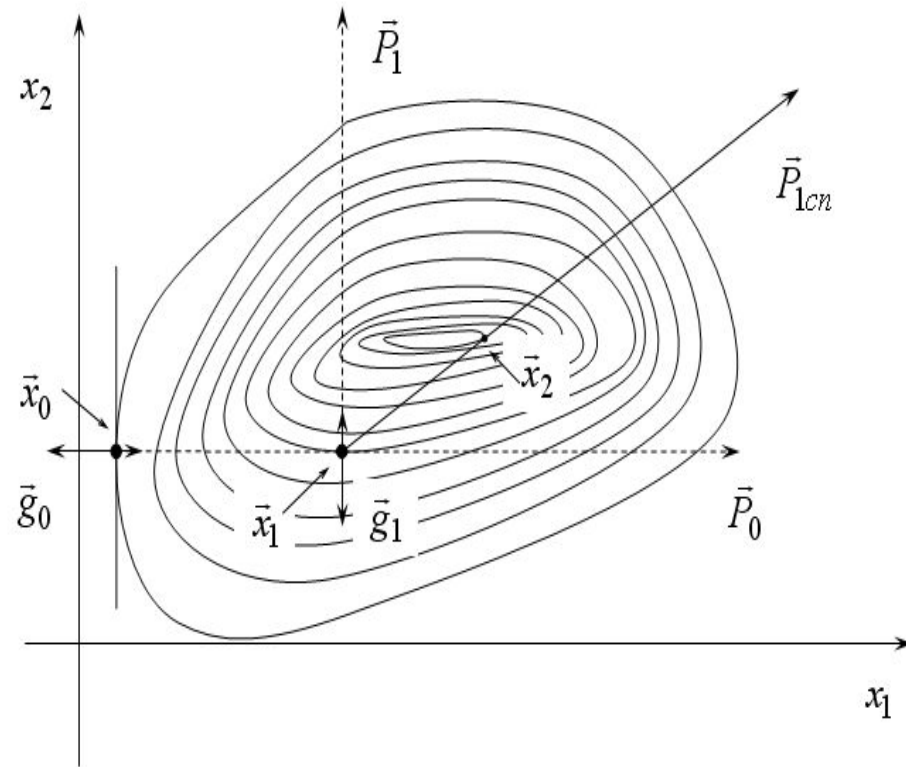
(к-1). $\bar{x}_k \quad \bar{P}_{kcn} = -\bar{g}_k + \sum_{i=1}^k \beta_i \bar{P}_{(i-1)cn} \rightarrow \bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{P}_{kcn}$

($\bar{P}_{kcn}$ взаимно сопряженное $\bar{P}_{(k-1)cn}, \dots, \bar{P}_0$)

$$\beta_k = \frac{|\bar{g}_k|}{|\bar{g}_{k-1}|}$$

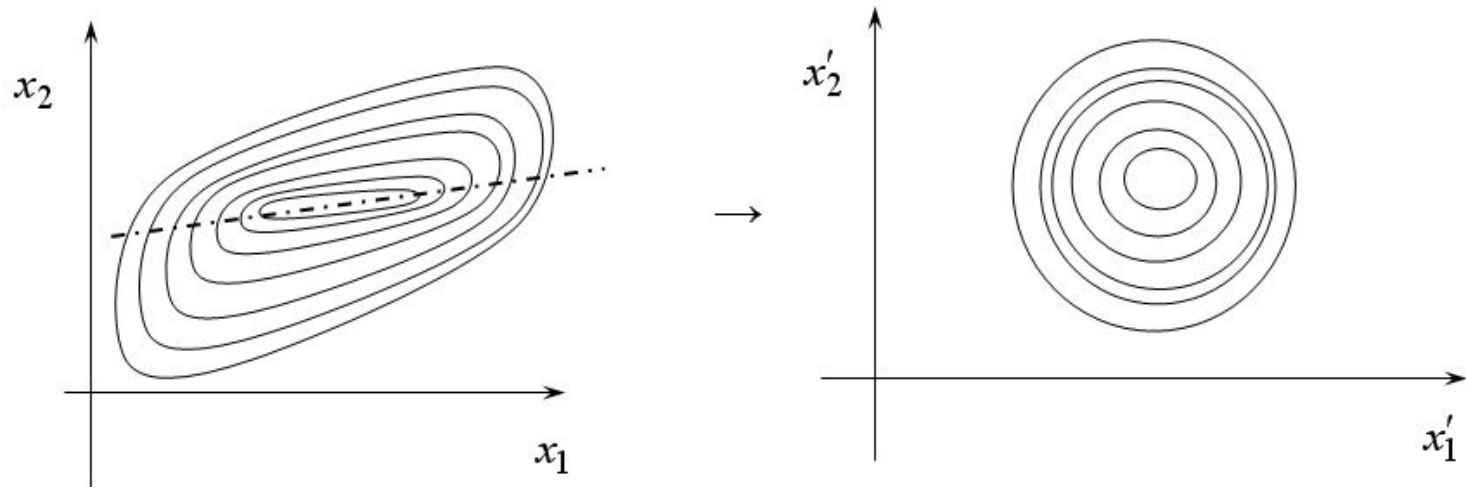
Минимум квадратичной функции n переменных
 достигим за n шагов

После n спусков - $(n+1)$ шаг с $\bar{P}_{n+1} = -\bar{g}_{n+1}$



Методы с переменной метрикой

Основа – преобразование системы координат



При известном преобразовании координат из любой точки пространства указывается направление минимума функции

Класс методов с переменной метрикой

$$\bar{x}_{k+1} = \bar{x}_k + \alpha_k \bar{P}_k$$

где $\bar{P}_k = -H_k \bar{g}_k$ $\bar{g}_k$ - градиент в $\bar{x}_k$

H_k - положительно определенная матрица размером $n \times n$

$H_{k+1} = H_k + E_k$ где E_k - корректирующая матрица

Какой должна быть H_k при минимизации конкретной функции?

Методы с переменной метрикой

- Условия: 1. $f(\bar{x})$ - квадратичная функция ($G_k = G_{k+1} = \text{const} = G$)
2. $\bar{x}_{k+1}$ - точка минимума $f(\bar{x})$ ($\bar{g}_{k+1} = 0$)
3. $\bar{x}_k \neq \bar{x}_{k+1}$

Определение направления, по которому достигается минимум:

1. разложение $f(\bar{x})$ в ряд Тейлора

$$f(\bar{x}) = f(\bar{x}_k) + [\bar{x} - \bar{x}_k]^T \bar{g}_k + [\bar{x} - \bar{x}_k]^T G_k(\bar{x}_k) [\bar{x} - \bar{x}_k] / 2$$

$$f(\bar{x}) = f(\bar{x}_{k+1}) + [\bar{x} - \bar{x}_{k+1}]^T \bar{g}_{k+1} + [\bar{x} - \bar{x}_{k+1}]^T G_{k+1}(\bar{x}_{k+1}) [\bar{x} - \bar{x}_{k+1}] / 2$$

2. $f(\bar{x}_{k+1}) = f(\bar{x}_k) + [\bar{x}_{k+1} - \bar{x}_k]^T \bar{g}_k + [\bar{x}_{k+1} - \bar{x}_k]^T G_k(\bar{x}_k) [\bar{x}_{k+1} - \bar{x}_k] / 2$
«+»

$$f(\bar{x}_k) = f(\bar{x}_{k+1}) + [\bar{x}_k - \bar{x}_{k+1}]^T \bar{g}_{k+1} + [\bar{x}_k - \bar{x}_{k+1}]^T G_{k+1}(\bar{x}_{k+1}) [\bar{x}_k - \bar{x}_{k+1}] / 2$$

3. $[\bar{x}_{k+1} - \bar{x}_k]^T [\bar{g}_k + G[\bar{x}_{k+1} - \bar{x}_k]] = 0$

$$\bar{g}_k + G[\bar{x}_{k+1} - \bar{x}_k] = 0 \quad \left| * G^{-1} \right.$$

Обобщенная формула Ньютона

$$\bar{x}_{k+1} = \bar{x}_k - G^{-1} \bar{g}_k$$

указывает точное перемещение из любой $\bar{x}_k$

в точку минимума (для квадратичных функций)

Методы с переменной метрикой

Для не квадратичных функций – направление с переменной метрикой $\bar{P}_k = -H_k \bar{g}_k$

Принцип методов с переменной метрикой – построение матрицы H_k в процессе поиска минимума, приближающейся в G^{-1}

Метод Ньютона – второго порядка (эффективнее методов с переменной метрикой), обладает недостатком: расчет на каждой итерации G_k

Методы с переменной метрикой – первого порядка (расчет $\bar{g}_k$, $H_k \rightarrow G^{-1}$ через n итераций)

Цикл вычислений в методах с переменной метрикой:

1. расчет $\bar{g}_k$
2. расчет $H_k = H_{k-1} + E_{k-1}$
3. расчет направления движения $\bar{P}_k = -H_k \bar{g}_k$
4. поиск минимума вдоль направления $\bar{P}_k$ $\bar{\sigma}_k = \alpha_k \bar{P}_k = \bar{x}_{k+1} - \bar{x}_k$
5. повторение цикла 1.-4. вычислений для $\bar{x}_{k+1}$

Методы с переменной метрикой

1. Метод Давидона-Флетчера-Пауэлла

$$\begin{aligned}\bar{\sigma}_k &= \bar{x}_{k+1} - \bar{x}_k \\ \bar{y}_k &= \bar{g}_{k+1} - \bar{g}_k\end{aligned}\quad H_1 = I \quad E_k = \frac{\bar{\sigma}_k \bar{\sigma}_k^T}{\bar{\sigma}_k^T \bar{y}_k} - \frac{H_k \bar{y}_k \bar{y}_k^T H_k}{\bar{y}_k^T H_k \bar{y}_k}$$

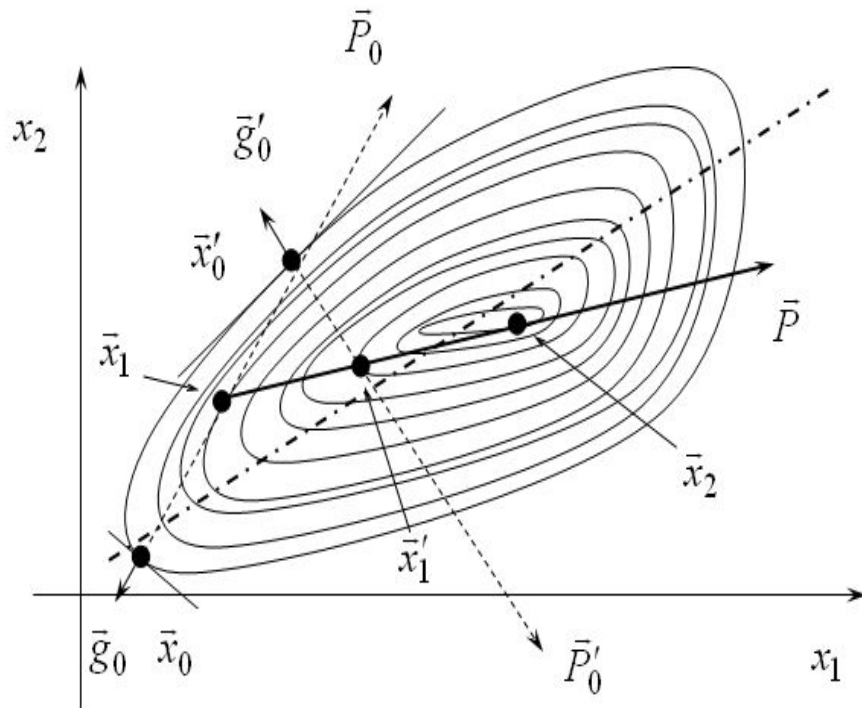
2. Семейство методов Гринстанда

требование наименьшей коррекции E_k в смысле ее нормы (удержание элементов H_k от сильного увеличения)

Метод Гольдфарба

$$E_k = \frac{1}{\bar{y}_k^T \bar{\sigma}_k} \left\{ -\bar{\sigma}_k \bar{y}_k^T H_k - H_k \bar{y}_k \bar{\sigma}_k^T + \left(1 + \frac{\bar{y}_k^T H_k \bar{y}_k}{\bar{y}_k^T \bar{\sigma}_k} \right) \bar{\sigma}_k \bar{\sigma}_k^T \right\}$$

Метод оврагов



1. методом наискорейшего спуска переход от $\bar{x}_0$ к $\bar{x}_1$
 2. методом наискорейшего спуска переход от $\bar{x}'_0$ к $\bar{x}'_1$
 3. при $f(\bar{x}'_1) < f(\bar{x}_1)$ - овражный шаг к $\bar{x}_2$ в направлении

$$\bar{P} = \bar{x}'_1 - \bar{x}_1$$
 4. методом наискорейшего спуска переход от $\bar{x}_2$ к $\bar{x}_3$
 5. при $f(\bar{x}_3) < f(\bar{x}'_1)$ - овражный шаг к $\bar{x}_4$ в направлении

$$\bar{P}_1 = \bar{x}_3 - \bar{x}'_1$$
- и т.д.

Невыполнение условий – неудачный выбор начальных приближений $\bar{x}_0$ и $\bar{x}'_0$

Недостатки метода: нет свойства универсальности,
от величины овражного шага
зависит попадание в окрестность
точки минимума

Методы второго порядка

$$\bar{x}_{k+1} = \bar{x}_k - \alpha_k G^{-1}(\bar{x}_k) \bar{g}_k$$

где $\bar{g}_k$ - градиент в $\bar{x}_k$

$G(\bar{x}_k)$ - матрица Гессе в $\bar{x}_k$

при $\alpha_k = 1$ - формула Ньютона

при α_k из условия частного минимума ЦФ – алгоритм Ньютона-Рафсона

при $G(\bar{x}_k)$ не положительно определенной – расходимость метода

Замечание: $G(\bar{x}_k)$ исследуют на положительную определенность, в случае невыполнения условий – дополнительная задача по выбору направлений спуска

Сравнение методов

1. Методы нулевого порядка

гарантирована сходимость к минимуму – метод Розенброка и
метод Нельдера-Мида

хорошее быстроедействие на функциях с простым рельефом – метод Пауэлла

2. Методы первого порядка

устойчивость метода – метод Гольдфарба

эффективность (быстроедействие) – метод Давидона-Флетчера-Пауэлла

3. Методы второго порядка

редкое применение из-за необходимости расчета $G(\bar{x}_k)$ (применение только в
случаях аналитически заданной функции $f(\bar{x})$)