

Методы локальной оптимизации функций многих переменных

- метод наискорейшего спуска
- метод сопряженных градиентов
- метод Ньютона-Рафсона, квази-Ньютоновские методы

Постановка

- имеем начальную точку \mathbf{x}_0 , **необходимо** построить алгоритм движения к минимуму функции $\Phi(\mathbf{x})$, т.е. последовательность точек $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ в которых функция убывает и сходится к точке минимума

Решение:

в окрестности \mathbf{x}_k представим функцию квадратичной формой

$$\Phi(\mathbf{x}) = \Phi(\mathbf{x}_k) + \nabla \Phi(\mathbf{x}_k)(\mathbf{x} - \mathbf{x}_k) + \frac{1}{2} (\mathbf{x} - \mathbf{x}_k)^T \mathbf{H}(\mathbf{x}_k) (\mathbf{x} - \mathbf{x}_k) + \dots \quad (6.1)$$

условие минимума:

$$\nabla \Phi(\mathbf{x}_m) = 0 \quad \text{или} \quad \left(\frac{\partial \Phi}{\partial x_i} \right)_{\mathbf{x}=\mathbf{x}_m} = 0 \quad \text{для всех } i \quad (6.2)$$

$\mathbf{H}(\mathbf{x}_m)$ – матрица вторых производных

$$\frac{\partial^2 \Phi}{\partial x_i \partial x_j} = \mathbf{H} \quad \text{матрица гессиан,} \quad \text{положительно определена}$$

Ищем минимум квадратичной формы (6.1)

Метод наискорейшего спуска

1) имеем точку \mathbf{x}_k

вычисляем вектор направления спуска, направление антиградиента

$$\begin{aligned} \mathbf{s}_k &= -\mathbf{g}_k \\ \mathbf{g}_k &= \nabla \Phi(\mathbf{x})_{\mathbf{x}=\mathbf{x}_k} \end{aligned} \quad (6.3)$$

2) следующая точка

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{s}_k \quad (6.4)$$

λ - длина шага спуска

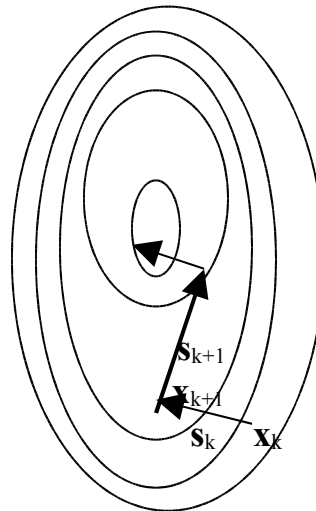
реализует минимум функции $\Phi(\mathbf{x}_k + \lambda \mathbf{s}_k)$ по направлению

$$\frac{\partial \Phi(\mathbf{x}_k + \lambda \mathbf{s}_k)}{\partial \lambda} = 0,$$

$$\lambda = \lambda_k$$

3) $\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{s}_k$

возврат к 1)



- метод удовлетворительно работает

- эффективен вдали от точки минимума, быстро сходится к окрестности минимума

- не эффективен в окрестности точки минимума, особенно для функции овражного типа

Метод сопряженных градиентов

- направление спуска выбирается с учетом предшествующего направления

$$\mathbf{s}_k = -\mathbf{g}_k + \gamma_k \mathbf{s}_{k-1} \quad (6.5)$$

скаляр γ_k

$$\gamma_k = \frac{\mathbf{g}_k \mathbf{g}_k}{\mathbf{g}_k - \mathbf{g}_{k-1}}$$

- направления спуска метода сопряженных градиентов выбирается более эффективно в области оврагов, чем в методе наискорейшего спуска,
- быстрее ведет к точке минимума,
- работает для овражных функций,
- хорошо работает в окрестности точки минимума

Метод Ньютона-Рафсона, квази Ньютоновские методы

- наиболее точный метод расчета направления спуска использует расчет матрицы Гессиана,

$$\mathbf{s}_k = -\mathbf{g}_k \mathbf{H} \quad (6.6)$$

минимум определяется точно, если $\Phi(\mathbf{x})$ – квадратичная форма (6.1)

$$\mathbf{x}_{\min} = \mathbf{x}_k + \mathbf{s}_k$$

для произвольной функции

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda \mathbf{s}_k$$

с величиной шага λ реализующий минимум функции по направлению \mathbf{s}_k

- быстро сходится, несколько итераций
- точен вблизи минимума
- вычислительно трудоемкий
- редко используется для макромолекул

квази Ньютоновские методы –

приближенный расчет матрицы гессиана по ходу итераций,

- метод Давидона-Флетчера,
- метод Флетчера-Пауэла.

ППЭ многоатомных биомолекул очень сложная –

- овражного типа (следствие корреляций положений атомов), направления движения атомов резко не эквивалентны на ППЭ,
- множество локальных минимумов (в том числе артефакты силового поля),
- последовательность точек \mathbf{x}_k алгоритмов локальной оптимизации сходится медленно, т.к. ППЭ аппроксимируется квадратичной формой только в малой окрестности точки \mathbf{x}_k .

Псевдо-глобальная оптимизация

Генетический алгоритм оптимизации

[Goldberg, 1989; Judson P.S. *J.Comp.Chem.* 1993, 14,p.1407-1414]

Реализует процесс биологической эволюции для определения оптимальных состояний системы –

метод псевдоглобальной оптимизации функции многих переменных

$\Phi(x)$ –

x_i - вектор оптимизируемых параметров,

ГА работает с **дискретными** значениями параметров,

Подготовка системы

1. определение генов

1.1 задаем область определения каждой переменной x_i

1.2 назначаем переменным множество дискретных пронумерованных значений определения – n_i – число дискретных значений переменной i

1.3 кодируем n_i – строкой бит в двоичном коде

например углы вращения $0 < \phi < 360$, разбиваем на 32 дискретных значения = 0,12,24,...,348 = 0,1,2,...,32 =
=00000,00001,00010,00011,...,11111

каждый параметр кодируется строкой из 5-ти элементов 0 1

1.4. **каждый параметр – ген**

1.5. **строка параметров = строка генов = хромосома**

Примеры генов – параметров и хромосом

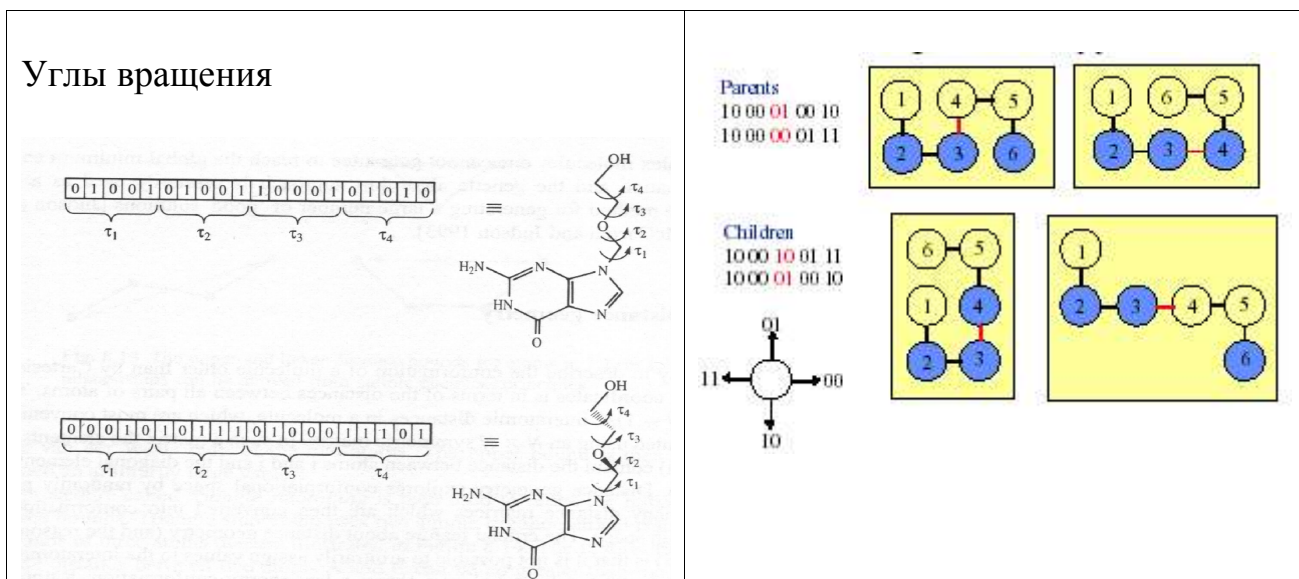


Рис. 6-2.

2. Генерация исходных особей – множества генов и хромосом

- некоторое количество **N хромосом** случайно покрывающих всю область определения
- вычисляется качество хромосомы, величина функции $\varphi_i = \Phi(\chi_i)$

Организация эволюционного цикла

Для организации эволюции генов определяются 4 операции

1. **выживание** – уничтожение особей (хромосом) плохого качества
2. **скрещивание** – создание хромосомы потомка из пары родителей
3. **мутация** - изменение гена в хромосоме потомка

Основной эволюционный цикл

1. **выживание**, расчет качества особей популяции P_k и выживание 1/2 лучших особей из популяции
2. **скрещивание** пар хромосом родителей, **генерация N/2 детей**
 - **случайный** выбор пар родителей
 - **случайный** выбор позиции точки кроссовера вдоль гена, обмен частями последовательности в паре **хромосом**

111110000010101
родители -----> 111110010001110
100011110001110

3. **мутация случайного** гена в хромосоме потомства – с вероятностью $gen_{mut} \ll 1$ малая величина

111110010001110 -----> 111110011001110

4. восстановление популяции до N особей = P_{k+1}
5. сравнение популяций P_{k+1} и P_k
6. возврат на 1. если P_{k+1} и P_k различны
7. В идеале, все особи P_m одинаковы и имеют максимально-возможное качество – глобальный минимум

В реальности – ГА алгоритм не гарантирует достижение глобального минимума, но генерирует популяцию хорошего качества

Метод Монте Карло

- генерация микроканонического ансамбля конформаций при заданной температуре T

энергия системы

$$\Phi(\mathbf{x}) = \Phi(\mathbf{r}_1, \dots, \mathbf{r}_N)$$

Система принимает состояния \mathbf{x}_α , $\alpha=1,2,\dots,K \rightarrow \infty$

вероятность найти систему в состоянии \mathbf{x}_α - распределение Гиббса

$$\rho_\alpha = \exp(-\Phi(\mathbf{x}_\alpha)/kT)/Z$$

вероятность найти систему в состоянии \mathbf{x}_β

$$\rho_\beta = \exp(-\Phi(\mathbf{x}_\beta)/kT)/Z$$

$\pi_{\alpha\beta}$ и $\pi_{\beta\alpha}$ **вероятности переходов** между состояниями

в состоянии равновесия, должен соблюдаться

-принцип микроскопической обратимости переходов между состояниями

-принцип детального равновесия между состояниями

$$\rho_\alpha \pi_{\alpha\beta} = \rho_\beta \pi_{\beta\alpha}$$

либо

$$\pi_{\alpha\beta}/\pi_{\beta\alpha} = \exp(-[\Phi(\mathbf{x}_\beta) - \Phi(\mathbf{x}_\alpha)]/kT)$$

откуда следует что вероятности переходов

$$\begin{aligned} \pi_{\alpha\beta} &= \exp(-[\Phi(\mathbf{x}_\beta) - \Phi(\mathbf{x}_\alpha)]/kT), \text{ если } \Phi(\mathbf{x}_\beta) > \Phi(\mathbf{x}_\alpha) \\ &= 1 \qquad \qquad \qquad \text{если } \Phi(\mathbf{x}_\beta) < \Phi(\mathbf{x}_\alpha) \end{aligned} \quad (6.7)$$

- это метод **Метрополиса** – метод существенной выборки

Алгоритм метода Монте Карло

- 1) имеется некоторое начальное состояние \mathbf{x}_α
- 2) выберем **случайно** состояние \mathbf{x}_β
- 3) **перейдем** в новое состояние \mathbf{x}_β согласно критерию Метрополиса

$\pi_{\alpha\beta} = 1$, если $\Phi(\mathbf{x}_\beta) < \Phi(\mathbf{x}_\alpha)$ – энергия понижается

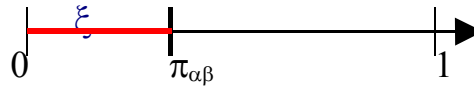
$\pi_{\alpha\beta} = \exp(-[\Phi(\mathbf{x}_\beta) - \Phi(\mathbf{x}_\alpha)]/kT)$, если $\Phi(\mathbf{x}_\beta) > \Phi(\mathbf{x}_\alpha)$ – энергия
повышается

как принять состояние с вероятностью $\pi_{\alpha\beta} < 1$?

-генерируем случайное число ξ ,

равномерно распределенное на интервале (0,1)

-если $\xi < \pi_{\alpha\beta} = \exp(-[\Phi(\mathbf{x}_\beta) - \Phi(\mathbf{x}_\alpha)]/kT)$, то состояние \mathbf{x}_β **принимается**

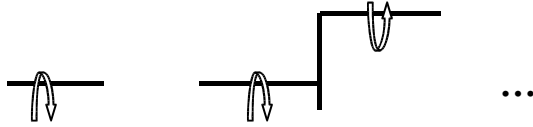


- 4) возврат к пункту 1)

Оптимальная работа алгоритма

достигается тогда, когда состояния α , β выбираются **случайно**, но достаточно **близкими**

- переход $\alpha \rightarrow \beta$ состоит из случайной комбинации малых «стандартных движений» для вращений вокруг связей



- малые вращения вокруг последовательности связей

- Свойства ансамбля состояний метода МК

- Алгоритм генерирует цепочку состояний α и число посещений nv_α этого состояния

$(\alpha, nv_\alpha), (\beta, nv_\beta), (\gamma, nv_\gamma), \dots$ составляющих Марковскую цепь =

- каждое состояние зависит только от одного предыдущего

- состояния с низкой энергией достигаются с большей вероятностью

- при достаточно долгой работе алгоритма все состояния системы будут достигнуты независимо от стартового состояния

Расчет средних характеристик по ансамблю состояний метода МК

среднее значение энергии системы

$$\langle U \rangle = \frac{1}{\sum_{\alpha} n_{\alpha}} \sum_{\alpha} n_{\alpha} U_{\alpha} \quad (6.8)$$

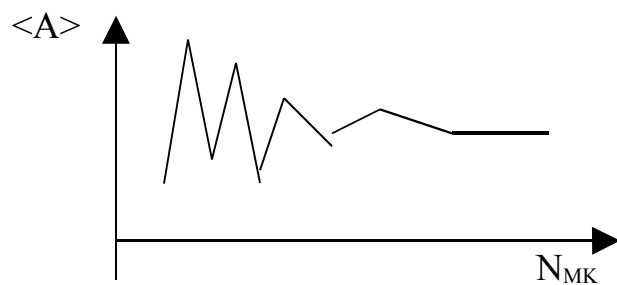
средне-квадратичная флуктуация энергии

$$\langle \Delta U^2 \rangle^{1/2} = \left[\frac{1}{\sum_{\alpha} n_{\alpha}} \sum_{\alpha} n_{\alpha} (U_{\alpha} - \langle U \rangle)^2 \right]^{1/2} \quad (6.9)$$

Метод МК сходится к равновесному ансамблю когда среднее

$$\langle A \rangle(N_{МК}) \sim \text{const}$$

- независимо от числа генерированных МК
состояний



Глобальная оптимизация методом МК - моделирование «отжига» системы

МК метод сопряженный с постепенным уменьшением температуры по
По специальному расписанию
 $T \rightarrow T - \Delta T$ через каждые N шагов

- **высокие T**

$$\Delta\Phi/RT \ll 1$$

- все состояния достижимы, возможны переходы через барьеры с высокой вероятностью,
плохая начальная точка передвигается в область низкой энергии

- **плавное понижение T**

- траектория МК способна преодолевать более низкие барьеры и передвигать в область низких энергий на ППЭ

