Московский государственный университет имени М.В.Ломоносова
Факультет вычислительной математики и кибернетики
Кафедра математических методов прогнозирования

Отчёт по теме
Продвинутые методы безусловной оптимизации

Выполнил студент 517 группы
Камалов Руслан Рамилевич

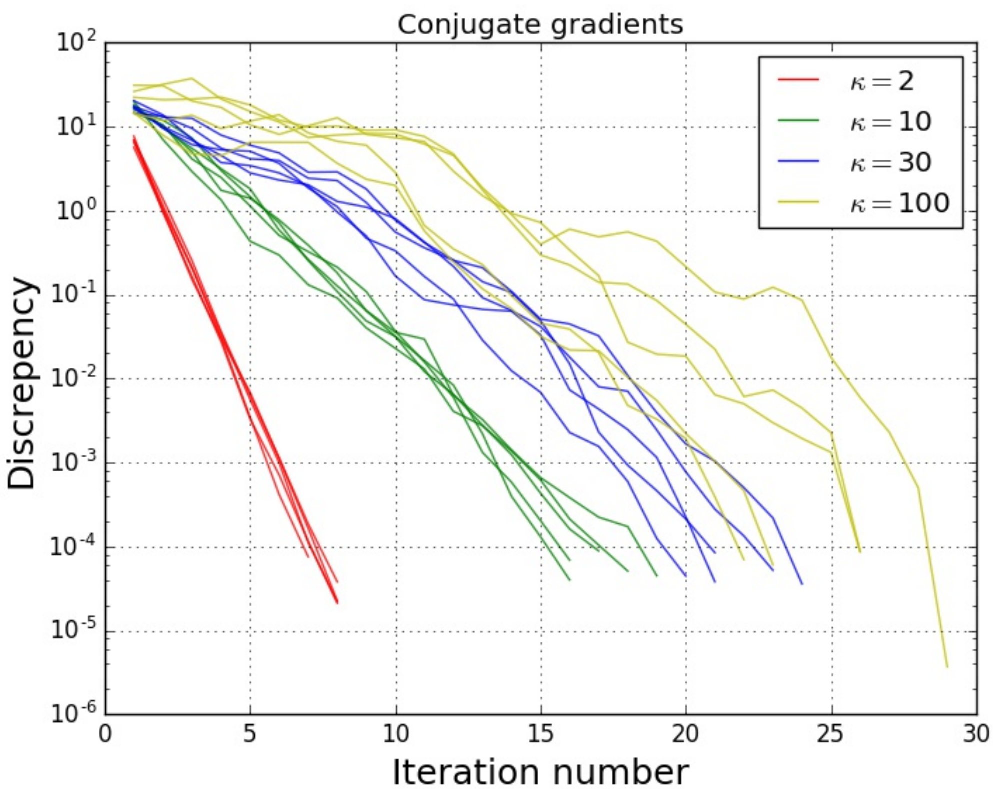
Москва
2016

1 Метод сопряженных градиентов

Эксперимент:

1. реализован метод сопряженных градиентов
2. выбраны значения для числа обусловленности системы $\kappa \in \{2, 10, 30, 100\}$
3. для каждого κ сгенерированы 5 положительно определенных матриц $A \in R^{n \times n}, n = 30$, с числом обусловленности $\kappa(A) = \kappa$, а также вектора правых частей $b \in R^{n \times 1}$
4. к каждой из полученных систем $Ax = b$, применен метод сопряженных градиентов

Результат эксперимента:



Вывод:

Результат эксперимента согласуется с теоретическим результатом относительно скорости сходимости метода сопряженных градиентов

$$\|x_k - x_*\| \leq \frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \|x_0 - x_*\|$$

Действительно, из результатов эксперимента видно, что метод имеет **линейную** скорость сходимости, а при увеличении числа обусловленности κ, скорость сходимости замедляется, оставаясь при этом линейной.

Также можно отметить сходимость метода за число шагов меньшее n. Собственные числа матрицы A выбирались из равномерного распределения $R[1, \kappa]$, и группировались в кластеры, за счет этого метод сходился за число шагов меньшее размера системы. Чем больше значение κ, тем на больше количество кластеров мог "расслоиться"случайный вектор размера n

2 Функция потерь логистической регрессии

Эксперимент:

1. реализованы процедуры подсчета значения, градиента, а также умножения произвольного вектора на гессиан функции потерь логистической регрессии

2. реализована процедура разностного подсчета значения градиента и произведения произвольного вектора на гессиан функции потерь логистической регрессии

3. Сгенерирована случайная выборка $X \in R^{200\times100}$, случайный вектор ответов $y \in \{-1, +1\}^{200}$

4. Для четырех случайных пар точек (w, v) проверена корректность численного подсчета градиента и процедуры умножения гессиана на произвольный вектор

Результат:

	(w_1, x_1)	(w_2, x_2)	(w_3, x_3)	(w_4, x_4)
$\|\nabla f - (\nabla f)^*\|_\infty$	0.00015	0.00014	0.00014	0.00005
$\|\nabla fv - (\nabla f)^*v\|_\infty$	0.02166	0.02773	0.02930	0.01990

Вывод:

Численный подсчет градиента и гессиана функции потерь логистической регрессии реализованы корректно.

3 Сравнение реализованных методов

Эксперимент:

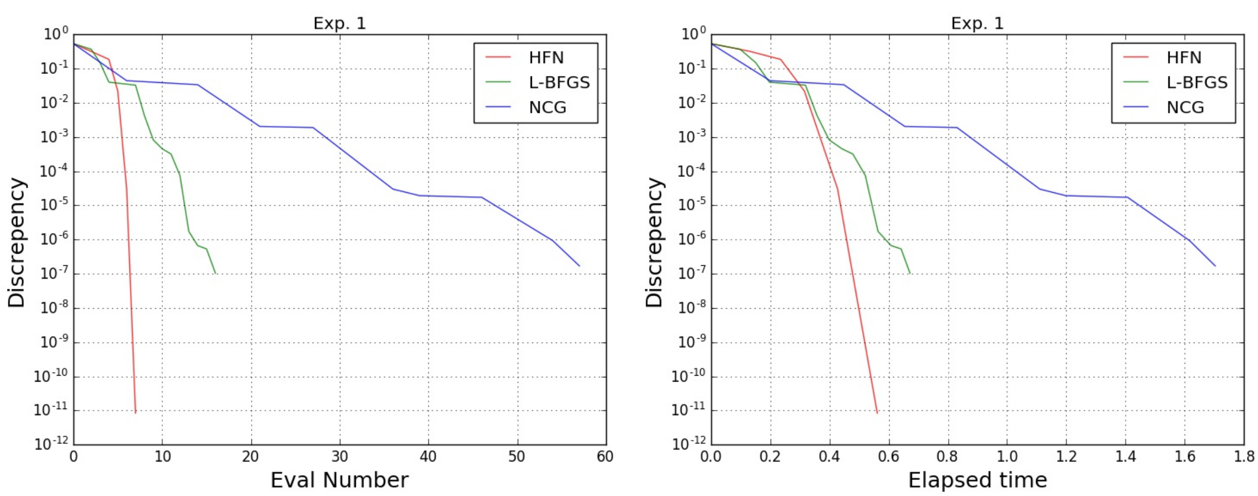
1. реализованы три метода оптимизации: L-BFGS,HFN, NCG

2. выбраны четыре датасета с размерами: $heart : (n = 240, d = 13); german.numer : (n = 1000, d = 24); duke\ breast - cancer : (n = 44, d = 7129); leukemia : (n = 38, d = 7129)$

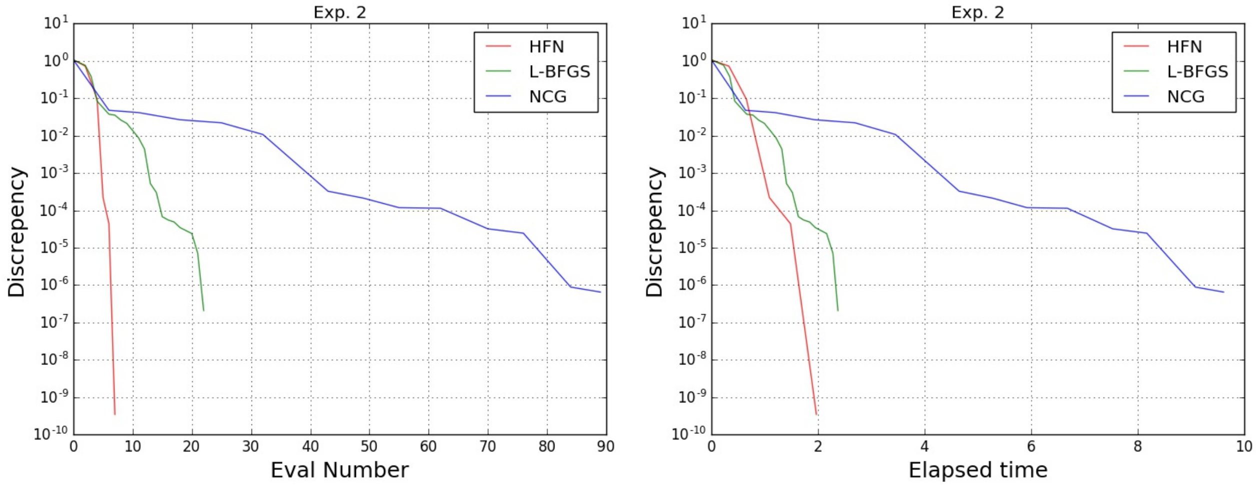
3. для каждого датасета протестированы три метода для оптимизации логистической функции потерь

Результат:

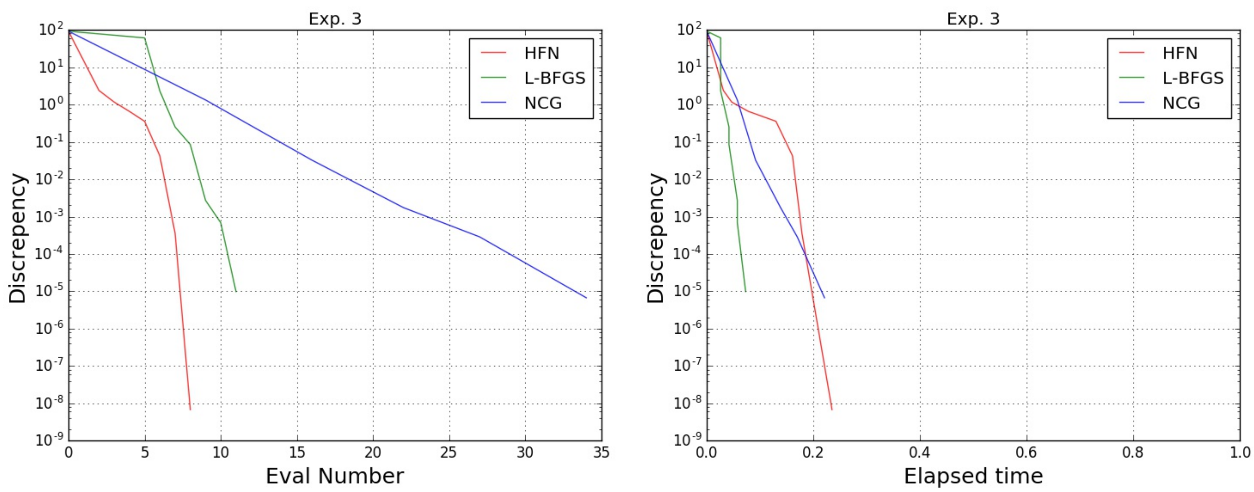
HEART



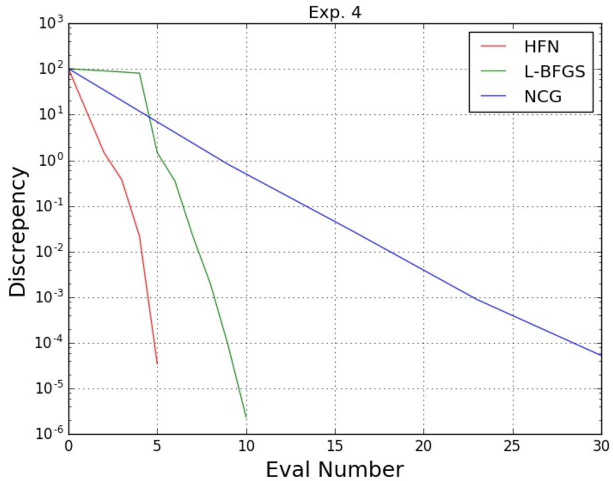
GERMAN.NUMER

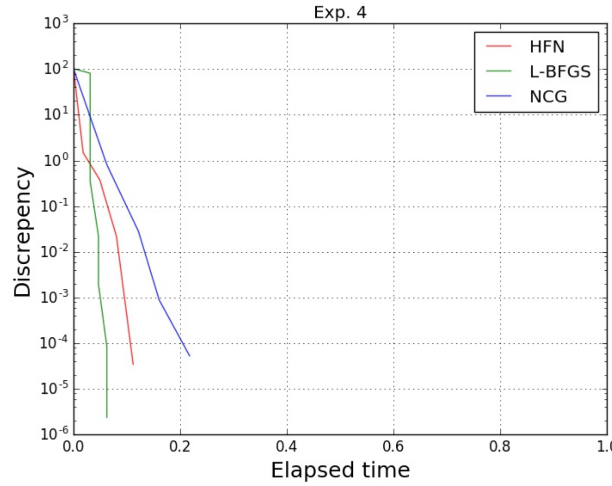


DUKE



LEUKEMIA





Вывод:

Глядя на графики сходимости методов можно сказать, что HFN и L-BFGS имеют квадратичную скорость сходимости, а NCG линейную.