Сравнение оптимизационных методов

Для подгонки параметров модели Барояна-Рвачёва необходимо максимизировать значение R^2 \in [0; 1] – схожесть смоделированной кривой и реальных данных о недельной заболеваемости во время эпидемии. Было проведено сравнение методов выпуклой оптимизации, реализованных в библиотеке numpy и реализованного собственными силами генетического алгоритма. Среди подходящих к задаче методов:

* SLSQP
* L-BFGS-B
* Truncated NC
* Nelder-Mead
* Genetic – реализованная эвристическая оптимизация

Для того, чтобы условия сравнения были равными, начальное значение генератора псевдослучайных чисел выбрано равным для всех методов (кроме генетического алгоритма, его поставить в равные условия с прочими невозможно).

Значение параметра I0 (начальная инцидентность) положим 10000 для всех замеров.

Количество попыток задать начальные K, I0 (из равномерного псевдослучайного распределения) положим 25. Это, конечно, не имеет смысле в случае генетического алгоритма, там важен размер популяции и количество поколений. Положим их равными 500 и 50, соответственно.

Для разнообразия результатов были выбраны эпидемии города Москва в 1986—1989 годах. Поведение методов и скорость их сходимости сильно зависит от входных данных. Например, в 1986 все рассматриваемые методы показали хороший результат, одинаковый с точностью до 0.001, поэтому результаты не отражены в отчёте. Далее в таблицах следуют замеры скорости и значения R^2.

1987 год

|  |  |  |
| --- | --- | --- |
| **Method** | **R^2** | **Elapsed time (sec)** |
| SLSQP | 0.832389 | 225 |
| L-BFGS-B | 0.829263 | 650 |
| TNC | 0.832390 | 708 |
| Nelder-Mead | 0.909165 | 2426 |
| Genetic | 0.832380 | 8027 |

1988 год

|  |  |  |
| --- | --- | --- |
| **Method** | **R^2** | **Elapsed time (sec)** |
| SLSQP | 0.809949 | 145 |
| L-BFGS-B | 0.880573 | 865 |
| TNC | 0.905172 | 993 |
| Nelder-Mead | 0.905333 | 4176 |
| Genetic | 0.902974 | 8051 |

1988 год

|  |  |  |
| --- | --- | --- |
| **Method** | **R^2** | **Elapsed time (sec)** |
| SLSQP | 0. 906502 | 155 |
| L-BFGS-B | 0. 908231 | 3493 |
| TNC | 0. 908231 | 3474 |
| Nelder-Mead | 0. 908246 | 5028 |
| Genetic | 0. 908184 | 8039 |

По итогам работы можно селать некоторые выводы:

1. Генетический алгоритм уступает прочим даже при значительном увеличении времени вычисления. Решено не использовать его в дальнейших расчётах
2. Метод SLSQP выполняется гораздо быстрее других, но и проигрывает по целевому значению. Причины следует выяснить в дальнейшем
3. Метод Nelder-Mead лучше всех (незначительно) натягивает модель на данные, но примерно в два раза превосходит по времени вычисления. Решено не использовать в расчётах

Все технические детали сравнения и графики приложены к отчёту. Также изучена зависимость R^2 от величины популяции и количества поколений в генетическом алгоритме. К сожалению, результаты сильно уступают прочим методам.