

Гетероскедастичные гауссовские процессы и их применение для байесовской оптимизации

Колодяжный Максим

Научный руководитель:
Алексей Алексеевич Зайцев

25 апреля 2018 г.

- 1 Аппроксиматор на основе гауссовского процесса
 - Определение
 - Оценка параметров гауссовского процесса
- 2 Аппроксиматор на основе гетероскедастичного гауссовского процесса
 - Определение
 - Оценка параметров
- 3 Байесовская оптимизация
 - Определение
 - Пример Acquisition function - Expected Improvement
 - Пример работы оптимизации
- 4 Работа с библиотеками и экспериментальные результаты
 - Библиотеки GPy и GPyOpt
 - Анализ работы оптимизации на функциях с разным видом шума
 - Применение гетероскедастичной регрессии для байесовской оптимизации

Аппроксиматор на основе гауссовского процесса

$f(\cdot)$ - гауссовский процесс

x - точка пространства

$m(x) = \mathbb{E}[f(x)]$ - среднее

$k(x, x') = \mathbb{E}[(f(x) - m(x))(f(x') - m(x'))]$

- ковариационная функция

Апостериорное среднее:

$$\hat{f}(X_*) = K_* K^{-1} y, \quad (1)$$

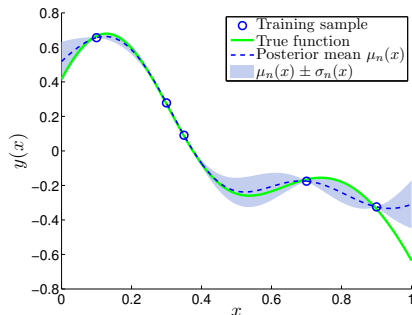
В большинстве случаев:

$$y(x) = f(x) + \varepsilon(x), \varepsilon \sim \mathcal{N}(0, \tilde{\sigma}^2) \quad (2)$$

Тогда:

$$\hat{f}(X_*) = K_* (K + \tilde{\sigma}^2 I_N)^{-1} y, \quad (3)$$

Параметры моделей оцениваются с использованием метода максимума правдоподобия



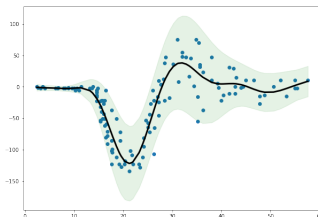
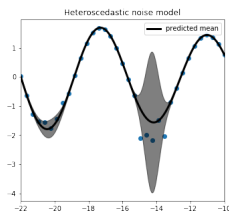
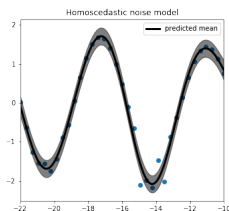
Аппроксиматор на основе гетероскедастичного гауссовского процесса

Определим гетероскедастичную модель следующим образом:

$$f(x) \sim \mathcal{GP}(0, k_f(x, x')), \quad \varepsilon_i \sim \mathcal{N}(0, r(x_i)), \quad r(x) = e^{g(x)}, \\ g(x) \sim \mathcal{GP}(\mu_0, k_g(x, x'))$$

Так как для такой модели, мы не можем получить правдоподобие в явном виде, будем максимизировать его нижнюю оценку Variational Lower Bound:

$$F(q(f), q(g)) = \log p(y) - \text{KL}(q(f)q(g) \parallel p(f, g \mid y)).$$



Метод глобальной оптимизации

- 1 Фиксируем модель данных.
 - 2 Оцениваем параметры модели, используя метод максимума правдоподобия или его аналоги.
 - 3 Выбираем точку с максимумом acquisition function для следующего вычисления целевой функции
 - 4 Делаем апдейт модели с использованием значения в новой точке
- итерируем между шагом 2 и 4 шагом до тех пор пока не будет исчерпан лимит на количество оценок.

Пример Acquisition function - Expected Improvement

Обозначим $\Delta(x) = y_{\text{best}} - \mu_*(x)$,

тогда $\alpha_{\text{EI}} = \max(0, \Delta(x)) - \sigma_*(x) \varphi\left(\frac{\Delta(x)}{\sigma_*(x)}\right) + |\Delta(x)| \Phi\left(-\frac{|\Delta(x)|}{\sigma_*(x)}\right)$

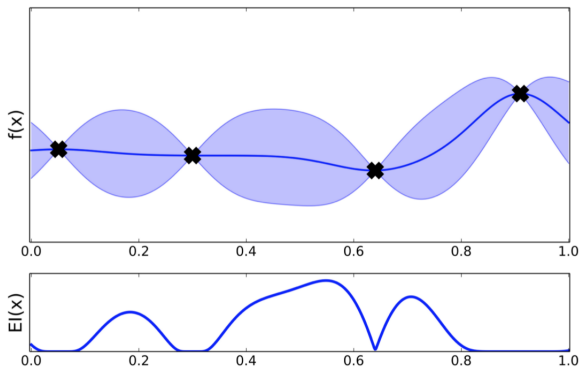
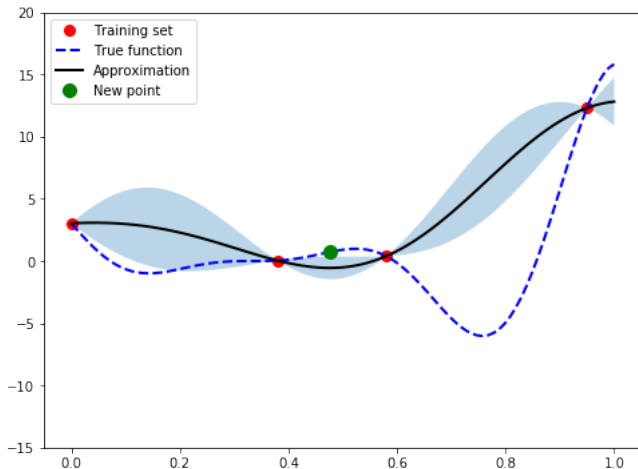
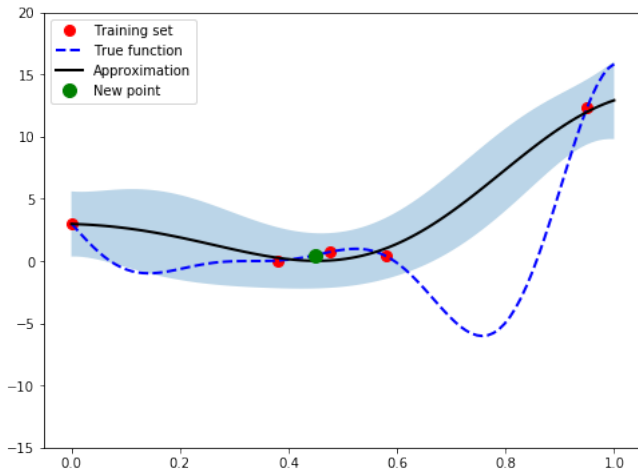


Рис.: Значения функции EI на заданной функции $f(x)$

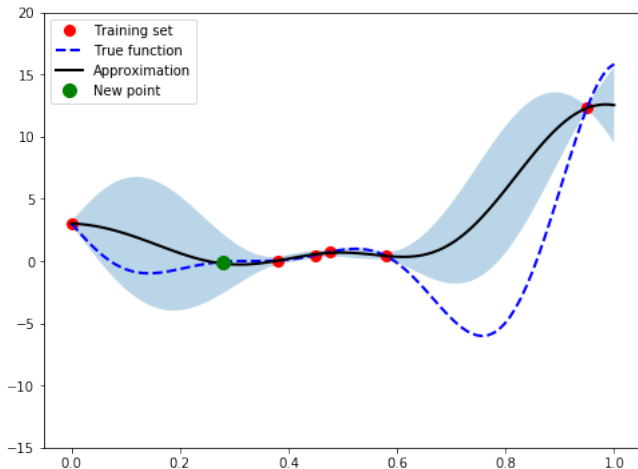
Пример работы оптимизации



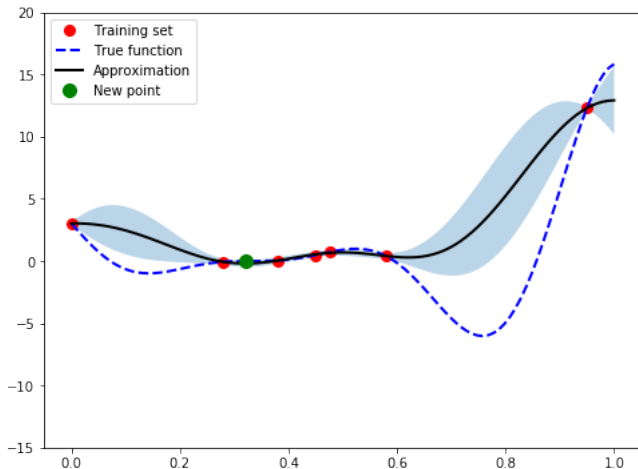
Пример работы оптимизации



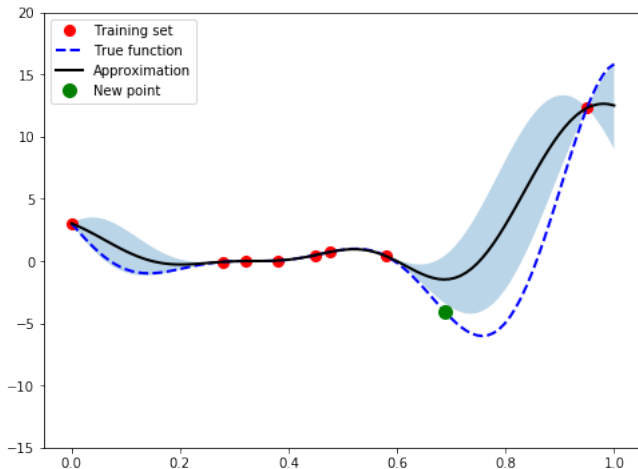
Пример работы оптимизации



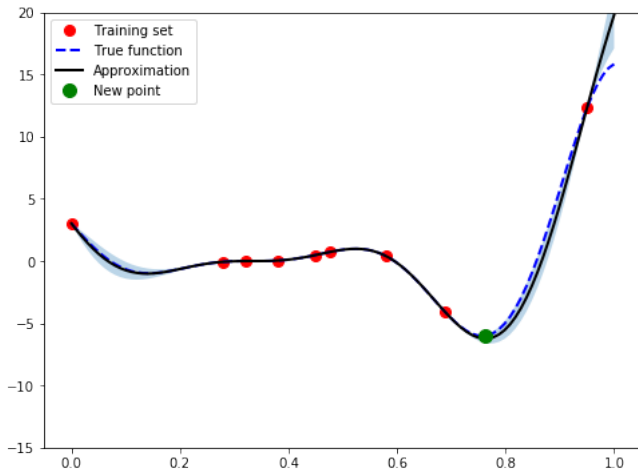
Пример работы оптимизации



Пример работы оптимизации



Пример работы оптимизации



- 1 Аппроксиматор на основе гауссовского процесса
 - Определение
 - Оценка параметров гауссовского процесса
- 2 Аппроксиматор на основе гетероскедастичного гауссовского процесса
 - Определение
 - Оценка параметров
- 3 Байесовская оптимизация
 - Определение
 - Пример Acquisition function - Expected Improvement
 - Пример работы оптимизации
- 4 Работа с библиотеками и экспериментальные результаты
 - Библиотеки GPy и GPyOpt
 - Анализ работы оптимизации на функциях с разным видом шума
 - Применение гетероскедастичной регрессии для байесовской оптимизации

Работа с библиотеками и экспериментальные результаты

SheffieldML / GPy

Watch 198 Star 732 Fork 285

Code Issues Pull requests Projects Wiki Insights

Gaussian processes framework in python

5,497 commits 20 branches 42 releases 49 contributors BSD-3-Clause

Branch: **dev** New pull request Create new file Upload files Find file Close or download

mergecode Merge pull request #113 from dmnp/MML_single_log			Latest commit d55cde1 on 23 Mar
GPy	Maintains consistency with numpy arrays	8 months ago	
benchmarks/regression	(chaching) changing all chacher limits to 3	2 years ago	
doc	fix: rtd	2 months ago	
approveyr_upload.bat	fix: plog: approveyr upload order using twine	7 months ago	
coveragerc	(travis) update for the coverage reports	2 years ago	
gitchangelog.rc	fix: plog: changelogrc and changelog update	2 years ago	
gIgnore	add offline plotting for plotly	11 months ago	
travis.yml	Use old deploy pip behavior	2 months ago	

SheffieldML / GPyOpt

Watch 46 Star 274 Fork 90

Code Issues Pull requests Projects Wiki Insights

Gaussian Process Optimization using GPy

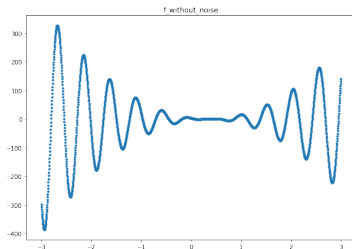
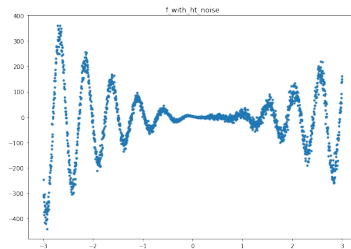
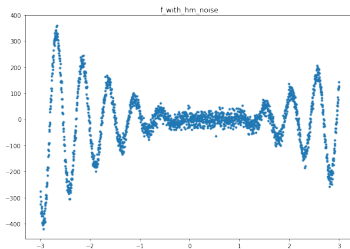
458 commits 6 branches 0 releases 25 contributors BSD-3-Clause

Branch: **master** New pull request Create new file Upload files Find file Close or download

Bugfix: Exploration steps too early with single init run			Latest commit 8551641 19 days ago
GPyOpt	Bugfix: Exploration steps too early with single init run	19 days ago	
docs	Updated docs	28 days ago	
examples	update six-hump camel example	2 years ago	
manual	Updates to the library	4 months ago	
gIgnore	Updated docs	28 days ago	
travis.yml	Update travis.yml	7 months ago	
AUTHORS.txt	Released version 1.2.0	6 months ago	
LICENSE.txt	collection of notebooks added	3 years ago	

- 1 Аппроксиматор на основе гауссовского процесса
 - Определение
 - Оценка параметров гауссовского процесса
- 2 Аппроксиматор на основе гетероскедастичного гауссовского процесса
 - Определение
 - Оценка параметров
- 3 Байесовская оптимизация
 - Определение
 - Пример Acquisition function - Expected Improvement
 - Пример работы оптимизации
- 4 Работа с библиотеками и экспериментальные результаты
 - Библиотеки GPy и GPyOpt
 - Анализ работы оптимизации на функциях с разным видом шума
 - Применение гетероскедастичной регрессии для байесовской оптимизации

Работа с библиотеками и экспериментальные результаты



Работа с библиотеками и экспериментальные результаты

https://github.com/col14m/Heteroscedastic_noise/blob/master/BayesianOptimizationReport.ipynb

- ссылка на полный отчет.

Метрика, по которой оценивалась точность работы оптимизации - это количество попаданий в 5% окрестность точки минимума.

Функция (№)	гомоскедастичный шум	гетероскедастичный шум
1	54%	44%
2	14%	18%
3	14%	10%
4	72%	58%
5	68%	56%

Таблица: Результаты экспериментов

- 1 Аппроксиматор на основе гауссовского процесса
 - Определение
 - Оценка параметров гауссовского процесса
- 2 Аппроксиматор на основе гетероскедастичного гауссовского процесса
 - Определение
 - Оценка параметров
- 3 Байесовская оптимизация
 - Определение
 - Пример Acquisition function - Expected Improvement
 - Пример работы оптимизации
- 4 Работа с библиотеками и экспериментальные результаты
 - Библиотеки GPy и GPyOpt
 - Анализ работы оптимизации на функциях с разным видом шума
 - Применение гетероскедастичной регрессии для байесовской оптимизации

Работа с библиотеками и экспериментальные результаты

SheffieldML / GPy

Watch 198 Star 732 Fork 285

Code Issues Pull requests Projects Wiki Insights

Gaussian processes framework in python

5,497 commits 20 branches 42 releases 49 contributors BSD-3-Clause

Branch: **default** New pull request Create new file Upload files Find file Close or download

mergecode Merge pull request #113 from danielpm/ml_single_log

Latest commit d55cde1 on 23 Mar

GPy	Maintains consistency with numpy arrays	a month ago
benchmarks/regression	(chaching) changing all chacher limits to 3	2 years ago
doc	fix: rtd	2 months ago
approveyr_twine_upload.bat	fix: pip; approveyr/upload order using twine	7 months ago
coveragerc	[travis] update for the coverage reports	2 years ago
gitchangelog.rc	fix: pip; changelogrc and changelog update	2 years ago
gIgnore	add offline plotting for plotly	11 months ago
travis.yml	Use old deploy pip behavior	2 months ago

SheffieldML / GPyOpt

Watch 46 Star 274 Fork 90

Code Issues Pull requests Projects Wiki Insights

458 commits 6 branches 0 releases 25 contributors BSD-3-Clause

Branch: **master** New pull request Create new file Upload files Find file Close or download

harvard and spallone Bugfix: Exploration steps too early with single init run

Latest commit 8551641 19 days ago

GPyOpt	Bugfix: Exploration steps too early with single init run	19 days ago
docs	Updated docs	28 days ago
examples	update six-hump camel example	2 years ago
manual	Updates to the library	4 months ago
gIgnore	Updated docs	28 days ago
travis.yml	Update travis.yml	7 months ago
AUTHORS.txt	Released version 1.2.0	6 months ago
LICENSE.txt	collection of notebooks added	3 years ago

Работа с библиотеками и экспериментальные результаты

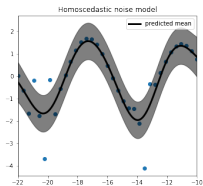
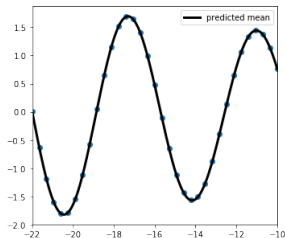


Рис.: Функция №1

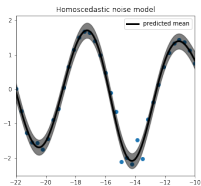
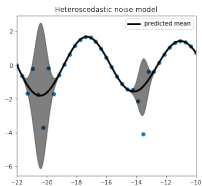
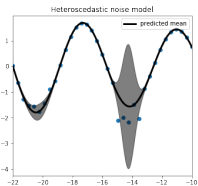


Рис.: Функция №2



Работа с библиотеками и экспериментальные результаты

https://github.com/col14m/Heteroscedastic_noise/blob/master/BayesOptHetModel.ipynb

- ссылка на полный отчет.

Метрика, по которой оценивалась точность работы оптимизации - это количество попаданий в 5% окрестность точки минимума.

Функция (№)	гомоскедастичная модель	гетероскедастичная модель
1	57%	66%
2	26%	46%

Таблица: Результаты экспериментов

Выводы

- экспериментально было получено, что в среднем на большинстве функций на гетероскедастичной модели оптимизация работает хуже примерно на 10%, чем на гомоскедастичной
- использование гетероскедастичной модели позволило на определенных видах функций повысить точность оптимизации от 10% до 20%
- гетероскедастичная модель из статьи не стабильная
- в библиотеках GPy и GPyOpt нет нормальной гетероскедастичной модели

Дальнейшее развитие работы

- введение регуляризации для параметров гетероскедастичной модели
- добавление данной гетероскедастичной модели в GPy и GPyOpt
- написание и публикация статьи

- Е.В. Бурнаев, М.Е. Панов, А.А. Зайцев, Регрессия на основе нестационарных гауссовских процессов с байесовской регуляризацией, 2015
- Miguel Lázaro-Gredilla, Michalis K. Titsias, Variational Heteroscedastic Gaussian Process Regression, 2011
- А.А. Зайцев, Variable Fidelity Regression Using Low Fidelity Function Blackbox and Sparsification, 2016
- Е.В. Бурнаев, Gaussian Processes and Bayesian Optimization, 2017
- https://github.com/col14m/Heteroscedastic_noise/blob/master/BayesOptHetModel.ipynb
- https://github.com/col14m/Heteroscedastic_noise/blob/master/BayesOptHetModel.ipynb

Спасибо за внимание!

Логарифм правдоподобия:

$$\log p(y|X, a, \tilde{\sigma}) = -\frac{1}{2} \left(y^T (K + \tilde{\sigma}^2 I_N)^{-1} y + \log |K + \tilde{\sigma}^2 I_N| + N \log 2\pi \right),$$

$$\log p(y|X, a, \tilde{\sigma}) \rightarrow \max_{a, \tilde{\sigma}}.$$

Аппроксиматор на основе гетероскедастичного гауссовского процесса

Вычислим распределение $q^*(f)$, которое максимизирует $F(q(f), q(g))$:

$$q^*(f) = \operatorname{argmax}_F = \frac{p(f)}{Z(q(g))} e^{\int q(g) \log p(y|f,g) dy}$$

$$\text{где } Z(q(g)) = \int e^{\int q(g) \log p(y|f,g) dy} p(f) df$$

Аппроксиматор на основе гетероскедастичного гауссовского процесса

В результате мы получим функцию:

$$F(q(g)) = \log Z(q(g)) - \text{KL}(q(f) \parallel p(g))$$

Данная функция зависит только от $q(g)$,
где $q(g) = \mathcal{N}(g \mid \mu, \Sigma)$