NTNU
Kunnskap for en bedre verden

Department of Chemical Engineering

TKP4580 - Chemical Engineering, Specialization Project

Applications of constrained Bayesian optimization to process systems

Author: Geir Arne Vassnes

Supervisor: Sigurd Skogestad
Co-supervisor: Lucas Ferreira Bernadino

January 2023

Abstract

Bayesian Optimization was in this thesis used to optimize two different cases. One algebraic case with both equality and inequality constraints, and a modelled case where only inequality constraints were used. The first case used 100-120 seconds to complete 15 iterations of the method, while the second cased used 45-50 seconds on the same task. The method had trouble with modelling constraints, and was therefore too slow for on-line usage.

Table of Contents

1 Introduction

Optimization is the process of finding the best solution to a problem by adjusting different parameters. It is widely used in various industries, such as manufacturing, finance, and logistics, to make the most efficient use of resources and maximize profits. In the chemical industry, optimization is used to optimize the utilization of expensive reactants and achieve the desired amount of high-value products, especially when by-products are present in the process. These byproducts may be less valuable or even harmful, and their production should be minimized.

Traditional optimization methods often rely on evaluating gradients and finding local minima. The gradient is the rate of change of a function with respect to its inputs, and it is used to guide the optimization algorithm towards the optimal solution. However, this approach has some limitations. For example, in some cases, the gradient of an objective function is expensive to calculate, and therefore, the optimization algorithm may not have access to sufficient information to find the optimal solution. Additionally, some objectives have multiple local minima, which may give sub-optimal control, depending on the initial values.

Bayesian optimization (BO) is a gradient-free approach that uses statistical learning to find the global minimum within given input bounds, without the need for gradient calculations. BO uses a probabilistic model to predict the objective function's behavior, and it updates this model as new data is acquired. This allows the algorithm to explore the search space in an efficient and informed way, and it can also handle cases where the objective function is noisy or has multiple local minima. BO is also useful when the objective function is expensive to evaluate, such as in the case of simulating a complex chemical reaction.

This thesis will explore the application of BO in optimizing a Williams-Otto reactor and a general algebraic problem with constraints.

2 Theory

2.1 Gaussian processes

A Gaussian process (GP) is a type of statistical model that can be used for non-parametric regression and classification. It is a stochastic process, meaning that it is defined by a probability distribution over functions, rather than a single fixed function. [1]

The key property of a GP is that any finite collection of its points are jointly Gaussian distributed. This means that if you have a set of points from a GP, you can use the mean and covariance of those points to make predictions about the value of the process at other points.

One of the main advantages of GPs is that they provide a way to make predictions with uncertainty. Because a GP is a distribution over functions, you can use it to compute the probability that the true function lies within a certain region, or to compute the expected value of the function at a given point.

2.2 Bayesian Optimization

Bayesian optimization (BO) is a class of statistical-learning-based optimization methods focused on minimizing the value of a function. BO is designed as a black-box derivative-free global optimization using a GP as a basis for the results. [2]

In BO, the function to be optimized is treated as a black box, meaning that the algorithm does not require any knowledge of the function's gradient or structure. Instead, it relies on samples of the function's output (i.e., the function's value at different points) to construct a probabilistic model of the function. This model is a Gaussian process (GP). The GP provides a flexible and computationally efficient way to model the function's behavior, and it can handle a wide range of input domains and output ranges. [2]

The GP model is used to predict the function's behavior at new points, and the algorithm uses this prediction to guide the search for the global minimum. The algorithm starts by evaluating the function at a few initial points, and it then iteratively selects new points to evaluate based on the GP's prediction of the function's behavior. The algorithm updates the GP model with the new data and uses it to predict the function's behavior at new points. The algorithm will stop after a predefined number of iterations, when the algorithm reaches a satisfactory minimum or when the two last results are within a certain tolerance. [2]

3 Method

The bayesian optimization was done using the Trieste library in Python, the approach was used on two different cases, an algebraic approach with inequality and equality constraints, and a modelled approach with only inequality constraints.

3.1 Constrained Branin

For testing the efficency of constraints by using BO, the modified Branin function was used as an objective function with an equality and an inequality constraint. [3]

The objective function $f(x)$ is defined as shown in Equtaion 1.

$$f(x) = (15x_2 - \frac{5.1(15x_1 - 5)^2}{4\pi^2} + \frac{75x_1 - 25}{\pi} - 6)^2 + 10(1 - \frac{\cos(15x_1 - 4)}{8\pi} + 75x_1 - 25) \quad (1)$$

The inequality constraint $g(x)$ is defined as shown in Equation 2.

$$g(x) = (10 - 2x_1^2 + \frac{x_1^4}{3})x_1^2 + x_1x_2 + (x_2^2 - 4)x_2^2 + 4\sin(5\pi(1 - x_1)) + 4\sin(6\pi(1 - x_2)) - 6 \quad (2)$$

The equality constraint $h(x)$ is defined as shown in Equation 3.

$$h(x) = 20(x_1 - 0.7)^2 - 0.25 - x_2 = 0 \quad (3)$$

3.2 Williams-Otto reactor

For testing BO on a process system a reactor was modelled as described by Williams and Otto. [4]

The reactions of the reactor are given in Equation 4.

$$\begin{aligned} A + B &\xrightarrow{k_1} C \\ C + B &\xrightarrow{k_2} P + E \\ P + C &\xrightarrow{k_3} G \end{aligned} \quad (4)$$

The rates for the reactions are given in Equation 5.

$$\begin{aligned} k_1 &= k_{0,1}e^{-E_1/T_r} \\ k_2 &= k_{0,2}e^{-E_2/T_r} \\ k_3 &= k_{0,3}e^{-E_3/T_r} \end{aligned} \quad (5)$$

The dynamics of the reactor is given in Equation 6.

$$\begin{aligned} \frac{dx_A}{dt} &= \frac{F_A}{W} - \frac{(F_A + F_B)x_A}{W} - k_1x_Ax_B \\ \frac{dx_B}{dt} &= \frac{F_B}{W} - \frac{(F_A + F_B)x_B}{W} - k_1x_Ax_B - k_2x_Cx_B \\ \frac{dx_C}{dt} &= -\frac{(F_A + F_B)x_C}{W} + 2k_1x_Ax_B - 2k_2x_Cx_B - k_3x_Px_C \\ \frac{dx_P}{dt} &= -\frac{(F_A + F_B)x_P}{W} + k_2x_Cx_B - 0.5k_3x_Px_C \\ \frac{dx_E}{dt} &= -\frac{(F_A + F_B)x_E}{W} + 2k_2x_Cx_B \\ \frac{dx_G}{dt} &= -\frac{(F_A + F_B)x_G}{W} + 1.5k_3x_Px_C \end{aligned} \quad (6)$$

The model parameters for this case is given in Table 1.

The MVs for this reactor system are $u = [F_B, T_r]$ and the reactor has the objective function with constraints shown in Equation 7.

$$\begin{aligned} \min_u J^{ec} &= p_AF_A + p_BF_B - p_P(F_A + F_B)x_P - p_E(F_AF_B)x_E \\ &s.t. \\ g_1 &= F_B - 4.0 \leq 0 \\ g_2 &= T_r - 355.0 \leq 0 \\ g_3 &= x_G - 0.105 \leq 0 \end{aligned} \quad (7)$$

Parameter	Value
W	2105 kg
$k_{0,1}$	$1.6599 \cdot 10^{-6}$ kg/s
$k_{0,2}$	$7.2117 \cdot 10^{-8}$ kg/s
$k_{0,3}$	$2.6745 \cdot 10^{-12}$ kg/s
E_1	6666.7 K
E_2	8333.3 K
E_3	11111 K
p_A	79.23 \$/K
p_B	118.34 \$/K
p_P	1043.38 \$/K
p_E	20.92 \$/K
F_A	1.5 kg/s

Table 1: Model parameters for the WO-reactor

4 Results and discussion

For both methods, the reward of doing more iterations on the optimization is stepwise, this is characteristic for BO. Around 15 iterations seemed to be a safe estimate for both methods, as seen in Tables 2 and 4, and was used further for investigating variance by seeds.

For the constrained Branin, the real solution is x = [0.506, 0.501], with f(x) = 162.81, four out of six tested seeds, gets close to this solution, while the two others are far away as seen in Table 3. The optimization took 100-120 seconds.

The optimization seems to be struggling with equality constraints, as it currently needs to be modelled as two inequality constraints, with a set tolerance, for this case the equality constraint stayed below 1, which for some use-cases may be too high.

For the Williams-Otto reactor, the real solution was u = [3.43277, 355] with $J^{ec} = 87.0761$, all the tested seeds got close to this solution as seen in Table 5. The optimization still took 45-50 seconds, which is better than constrained Branin, but is still to slow for on-line optimization.

5 Conclusion

By using the Trieste package in python, constrained BO was done on the Branin function, and a WO-reactor. The optimization of both cases was too slow for on-line usage, and the method did not preform well on equality constraints.

6 Further work

For BO to be a viable method, the algorithm needs to be optimized. This is possible to do with making the algorithms by hand, instead of using an existing package, this also gives the opportunity to find a better solution for constraints, like creating smaller regions of focus, to avoid time being used on infeasible points.

Bibliography

[1] Novi Quadrianto, Kristian Kersting, and Zhao Xu. *Gaussian Process*, pages 428–439. Springer US, Boston, MA, 2010. ISBN 978-0-387-30164-8. doi: 10.1007/978-0-387-30164-8_324.

[2] Peter I. Frazier. A tutorial on bayesian optimization. https://arxiv.org/abs/1807.02811, 2018. Grabbed 02.12.2022

[3] Congwen Lu and Joel A. Paulson. No-regret bayesian optimization with unknown equality and inequality constraints using exact penalty functions. *IFAC-PapersOnLine*, 55(7):895–902, 2022. ISSN 2405-8963. doi: https://doi.org/10.1016/j.ifacol.2022.07.558.

[4] Theodore J. Williams and Robert E. Otto. A generalized chemical processing model for the investigation of computer control. *Transactions of the American Institute of Electrical Engineers, Part I: Communication and Electronics*, 79(5):458–473, 1960. doi: 10.1109/TCE.1960.6367296.

Appendix

A Results Constrained Branin

Iterations	x_1	x_2	f(x)	g(x)	h(x)	Time [s]
5	0.464	0.869	201.27	1.575	-1.62e-3	42.012
6	0.464	0.869	201.27	1.575	-1.62e-3	54.308
7	0.464	0.869	201.27	1.575	-1.62e-3	60.361
8	0.499	0.562	166.07	3.469	-6.02e-3	63.032
9	0.499	0.562	166.07	3.469	-6.02e-3	66.236
10	0.499	0.562	166.07	3.469	-6.02e-3	72.762
11	0.499	0.562	166.07	3.469	-6.02e-3	78.158
12	0.499	0.562	166.07	3.469	-6.02e-3	89.841
13	0.499	0.562	166.07	3.469	-6.02e-3	103.867
14	0.499	0.562	166.07	3.469	-6.02e-3	103.891
15	0.499	0.562	166.07	3.469	-6.02e-3	99.261
16	0.499	0.562	166.07	3.469	-6.02e-3	108.184
17	0.499	0.562	166.07	3.469	-6.02e-3	116.824
18	0.499	0.562	166.07	3.469	-6.02e-3	119.684
19	0.499	0.562	166.07	3.469	-6.02e-3	126.883
20	0.499	0.562	166.07	3.469	-6.02e-3	130.294

Table 2: Results from optimization of Constrained Branin at different iterations

Seed	x_1	x_2	f(x)	g(x)	h(x)	Time [s]
1404	0.499	0.562	166.07	3.469	-6.02e-3	99.740
1715	0.537	0.278	166.72	3.373	-9.04e-5	109.241
1999	0.828	0.534	426.30	3.757	-4.56e-1	120.565
5394	0.594	0.526	164.52	1.782	-7.26e-3	113.121
6293	0.833	0.658	457.78	2.315	-5.54e-1	110.105
9123	0.495	0.590	167.11	3.671	7.47e-5	108.522

Table 3: Results from optimization of Constrained Branin at different seeds

B Results WO-reactor

Iterations	F_B	T_r	J^{ec}	g_3	Time [s]
5	2.6593	350.78	-73.721	1.44e-2	23.290
6	2.6593	350.78	-73.721	1.44e-2	20.874
7	2.6593	350.78	-73.721	1.44e-2	23.685
8	2.6593	350.78	-73.721	1.44e-2	27.011
9	2.6593	350.78	-73.721	1.44e-2	31.496
10	2.6593	350.78	-73.721	1.44e-2	33.988
11	3.4205	354.97	-87.066	3.98e-4	36.594
12	3.4205	354.97	-87.066	3.98e-4	39.296
13	3.4205	354.97	-87.066	3.98e-4	42.766
14	3.4205	354.97	-87.066	3.98e-4	45.339
15	3.4205	354.97	-87.066	3.98e-4	48.324
16	3.4205	354.97	-87.066	3.98e-4	51.463
17	3.4205	354.97	-87.066	3.98e-4	54.384
18	3.4205	354.97	-87.066	3.98e-4	56.116
19	3.4205	354.97	-87.066	3.98e-4	59.758
20	3.4205	354.97	-87.066	3.98e-4	61.959

Table 4: Results from optimization of WO-reactor at different iterations

Seed	F_B	T_r	J^{ec}	g_3	Time [s]
1404	3.4205	354.97	-87.066	3.98e-4	50.015
1715	3.4326	355.00	-87.076	8.67e-6	46.224
1999	3.3583	354.78	-86.887	2.21e-3	47.239
5394	3.4054	354.86	-87.004	5.02e-4	45.853
6293	3.3091	353.99	-86.358	4.40e-4	49.854
9123	3.4275	355.00	-87.079	2.32e-4	49.946

Table 5: Results from optimization of WO-reactor at different seeds