

Deciding which acquisition function is best based on results from tests

AUTHOR

Emma Bowen

Timing test

The results from this test vary, but only by a few milliseconds. This is likely due to using a different runner (where the code is running) each time on Github as there can be background processes, etc. It could also be due to slightly different loading times for the R packages each time. Any random sampling has been fixed by setting a seed. On average, after five runs:

```
# Timing test output

case_study8: 2.1716514636

looped_ver_case_study8_KG: 3.2781787488

looped_ver_case_study8_multi_point_AEI: 2.21322831978

looped_ver_case_study8_multi_point_EI: 2.13455526
```

We can see that `looped_ver_case_study8_multi_point_EI` is the quickest to run but is very close to `case_study8.R`. `looped_ver_case_study8_multi_point_AEI` is also very close. However, `looped_ver_case_study8_KG` takes much longer to run. We need to bear in mind that these are all with the pre-evaluated objective function. This means we do not have to evaluate the objective function, we only have to look up the pre-evaluated value, which vastly reduces the running time.

Total evaluations test

These results do not vary, and we get the output:

```
# Total Evaluations Summary

case_study8 = 54

looped_ver_case_study8_KG = 50

looped_ver_case_study8_multi_point_AEI = 56

looped_ver_case_study8_multi_point_EI = 56
```

This clearly shows that the Knowledge Gradient acquisition function requires the least amount of evaluations of the objective function. When we do not have pre-evaluated points, this will make it the least expensive to run.

Number of Rounds test

The results from this test don't change.

```
#Rounds Completed Summary

case_study8 = 7 rounds

looped_ver_case_study8_KG = 7 rounds

looped_ver_case_study8_multi_point_AEI = 7 rounds

looped_ver_case_study8_multi_point_EI = 7 rounds
```

As we evaluate the objective function in parallel for however many points we pick each round (normally eight, but it may not be in the last round), this means they will all be even in terms of timing.

Convergence Test

The results show that all the files converge to the correct answer in 100/100 runs. There may be slight variations when fitting the Gaussian Processes which could lead to slightly different values for the acquisition functions. This could lead to a different ordering of the candidate points and so different selected next points. However, after 100 runs we can be relatively confident that this does not lead to the process converging to the wrong point.

Conclusion

Despite currently taking more time on average than the other methods, the Knowledge Gradient acquisition function method will be the best to use here because it requires less evaluations and so is less expensive to run. However, it does have the same number of rounds as the others, which is important because we evaluate the new points each round in parallel, and so it doesn't have an overwhelming lead in this particular example. The results from the convergence test are essentially irrelevant for this comparison as all methods got the same result.