

Horovod - 贝叶斯优化源码分析

bayesian_optimization

概率代理函数:

GaussianProcessRegressor 类实现

1. Kernel: 平方指数核
2. PosteriorPrediction: 得到后验概率分布, 协助计算EI

采集函数

ExpectedImprovement函数

方法:

AddSample:

添加一组样本

NextSample

1. 更新概率密度函数, 根据EI计算得到下一个样本点。
2. 计算样本点时根据L-BFGS 得到极小值点, (多次计算, 每次从均匀分布中随机取值, 取所有次数中的最小值)

Parameter Manager

bayes_: bayesian_optimization.

1. TunableParameter

分为以下两类参数

BayesianParameter

以vector形式存放, 其中存放了可以调节的参数以及对应的调节区间。

目前以BayesianVariable 形式存放的参数只有

```
enum BayesianVariable { fusion_buffer_threshold_mb, cycle_time_ms };
```

TunableParameter

```
CategoricalParameter<bool> hierarchical_allreduce_;  
CategoricalParameter<bool> hierarchical_allgather_;  
CategoricalParameter<bool> cache_enabled_;
```

以成员的形式在Parameter Manager中。

2. 参数调节范围

在parameter manager初始化时被确定。

3. Test points (初始化采样) 在parameter manager初始化时被确定

a. 如果迭代次数在初始化次数之内，直接使用初始化采样点进行训练

4. 每进行5次训练，进行一次更新，更新时使用五次训练中值作为参数配置对应的score

5. parameter_chain_ 包含了 <joint_params_, hierarchical_allreduce_, hierarchical_allgather_, cache_enabled_>

```
param->Tune(best_score, &new_best_score) -->
```

```
ParameterManager::BayesianParameter::OnTune --->
```

```
bayes_->AddSample(value, score);
```

添加样本点

TensorFusionThresholdBytes

response_cache_

对于离散型参数，直接进行遍历搜索。

对于连续性，使用贝叶斯优化进行优化。

6. 只对协调器进行调整以确保一致性。

7. operation中进行parameter_manager的初始化，以及更新。