torch.utils.checkpoint

Note

Checkpointing is implemented by rerunning a forward-pass segment for each checkpointed segment during backward. This can cause persistent states like the RNG state to be advanced than they would without checkpointing. By default, checkpointing includes logic to juggle the RNG state such that checkpointed passes making use of RNG (through dropout for example) have deterministic output as compared to non-checkpointed passes. The logic to stash and restore RNG states can incur a moderate performance hit depending on the runtime of checkpointed operations. If deterministic output compared to non-checkpointed passes is not required, supply preserve_rng_state=False to checkpoint or checkpoint_sequential to omit stashing and restoring the RNG state during each checkpoint.

The stashing logic saves and restores the RNG state for the current device and the device of all cuda Tensor arguments to the run_fn . However, the logic has no way to anticipate if the user will move Tensors to a new device within the run_fn itself. Therefore, if you move Tensors to a new device ("new" meaning not belonging to the set of [current device + devices of Tensor arguments]) within run_fn , deterministic output compared to non-checkpointed passes is never guaranteed.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] checkpoint.rst, line 26)

Unknown directive type "currentmodule".

.. currentmodule:: torch.utils.checkpoint

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] checkpoint.rst, line 27)

Unknown directive type "autofunction".

.. autofunction:: checkpoint

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\pytorch-master\docs\source\[pytorch-master] [docs] [source] checkpoint.rst, line 28)

Unknown directive type "autofunction".

 \dots autofunction:: checkpoint_sequential