Item 311
git_comments:

1. **comment:** @roywei: Removing fixed seed as flakiness in this test is fixed tracked at
   https://github.com/apache/incubator-mxnet/issues/11686
   label: test

git_commits:

1. **summary:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot (#11978)
   message: [MXNET-771] Fix Flaky Test test_executor.py:test_dot (#11978) * use assert_almost_equal,
   increase rtol, reduce matrix size * remove seed in test_bind * add seed 0 to test_bind, it is still flaky * add
   comments for tracking
   label: test

github_issues:

1. **title:** test_executor.test_dot has fixed seed that can mask flakiness
   body: The unit test in title have been using fixed seed to mask flakiness. Suggested action: 1. Evaluate
   whether the test is flaky without fixed seed. If not, remove seed. Else move to 2 2. If test is flaky,
   determine whether it's an actual uncaught edge case. If so, fix the operator. Else move to 3 3. If numerical
   instability is inevitable, adjust tolerance level appropriately.
   label: test
2. **title:** test_executor.test_dot has fixed seed that can mask flakiness
   body: The unit test in title have been using fixed seed to mask flakiness. Suggested action: 1. Evaluate
   whether the test is flaky without fixed seed. If not, remove seed. Else move to 2 2. If test is flaky,
   determine whether it's an actual uncaught edge case. If so, fix the operator. Else move to 3 3. If numerical
   instability is inevitable, adjust tolerance level appropriately.
3. **title:** test_executor.test_dot has fixed seed that can mask flakiness
   body: The unit test in title have been using fixed seed to mask flakiness. Suggested action: 1. Evaluate
   whether the test is flaky without fixed seed. If not, remove seed. Else move to 2 2. If test is flaky,
   determine whether it's an actual uncaught edge case. If so, fix the operator. Else move to 3 3. If numerical
   instability is inevitable, adjust tolerance level appropriately.
4. **title:** test_executor.test_dot has fixed seed that can mask flakiness
   body: The unit test in title have been using fixed seed to mask flakiness. Suggested action: 1. Evaluate
   whether the test is flaky without fixed seed. If not, remove seed. Else move to 2 2. If test is flaky,
   determine whether it's an actual uncaught edge case. If so, fix the operator. Else move to 3 3. If numerical
   instability is inevitable, adjust tolerance level appropriately.
   label: test
5. **title:** test_executor.test_dot has fixed seed that can mask flakiness
   body: The unit test in title have been using fixed seed to mask flakiness. Suggested action: 1. Evaluate
   whether the test is flaky without fixed seed. If not, remove seed. Else move to 2 2. If test is flaky,
   determine whether it's an actual uncaught edge case. If so, fix the operator. Else move to 3 3. If numerical
   instability is inevitable, adjust tolerance level appropriately.
   label: test

github_issues_comments:

1. **body:** Thanks for filing this issue. We will investigate this Flaky test
   label: test
2. Reproducible using seed: `MXNET_TEST_SEED=1509417289` Will investigate and try to increase
   tolerance. ``` Ran 1 test in 34.693s FAILED (failures=1) ubuntu@ip-172-31-81-100:~/incubator-mxnet$
   MXNET_TEST_SEED=1509417289 nosetests --logging-level=DEBUG --verbose -s
   tests/python/unittest/test_e xecutor.py:test_dot /usr/local/lib/python3.5/dist-packages/nose/util.py:453:
   DeprecationWarning: inspect.getargspec() is deprecated, use inspect.signature() instead
   inspect.getargspec(func) [INFO] Setting module np/mx/python random seeds, use
   MXNET_MODULE_SEED=633029204 to reproduce. /home/ubuntu/incubator-
   mxnet/tests/python/unittest/common.py:244: DeprecationWarning: The 'warn' method is deprecated, use
   'warning' instead logger.warn('*** test-level seed set: all "@with_seed()" tests run deterministically ***')

```
[WARNING] *** test-level seed set: all "@with_seed()" tests run deterministically ***
test_executor.test_dot ... [INFO] Setting test np/mx/python random seeds, use
MXNET_TEST_SEED=1509417289 to reproduce. FAIL
====================================================================== FAIL:
test_executor.test_dot ---------------------------------------------------------------------- Traceback (most recent
call last): File "/usr/local/lib/python3.5/dist-packages/nose/case.py", line 198, in runTest
self.test(*self.arg) File "/home/ubuntu/incubator-mxnet/tests/python/unittest/common.py", line 172, in
test_new orig_test(*args, **kwargs) File "/home/ubuntu/incubator-
mxnet/tests/python/unittest/test_executor.py", line 132, in test_dot sf = mx.symbol.dot) File
"/home/ubuntu/incubator-mxnet/tests/python/unittest/test_executor.py", line 78, in
check_bind_with_uniform assert reldiff(rhs_grad.asnumpy(), rhs_grad2) < 1e-6 AssertionError: -----------
--------- >> begin captured logging << -------------------- common: INFO: Setting module np/mx/python
random seeds, use MXNET_MODULE_SEED=633029204 to reproduce. common: WARNING: *** test-
level seed set: all "@with_seed()" tests run deterministically *** common: INFO: Setting test
np/mx/python random seeds, use MXNET_TEST_SEED=1509417289 to reproduce. --------------------
>> end captured logging << ---------------------- ----------------------------------------------------------------------
Ran 1 test in 0.153s FAILED (failures=1) ```
```
3. #11978 should fix it

github_pulls:

1. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot
   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test
   is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors.
   The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to
   make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4.
   removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in
   `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ###
   Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with
   [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny
   changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test
   coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) -
   Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests
   will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is
   well-documented: - For user-facing API changes, API doc string has been updated. - For new C++
   functions in header files, their functionalities and arguments are documented. - For new examples,
   README.md is added to explain the what the example does, the source of the dataset, expected
   performance on test set and reference to the original paper if applicable - Check the API doc at
   http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To
   the my best knowledge, examples are either not affected by this change, or have been fixed to be
   compatible with this change ### Changes ### See description ## Comments ## - If this change is a
   backward incompatible change, why must this change be made. - Interesting edge cases to note here
   label: test
2. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot
   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test
   is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors.
   The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to
   make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4.
   removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in
   `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ###
   Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with
   [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny
   changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test
   coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) -
   Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests
   will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is
   well-documented: - For user-facing API changes, API doc string has been updated. - For new C++
   functions in header files, their functionalities and arguments are documented. - For new examples,
   README.md is added to explain the what the example does, the source of the dataset, expected

performance on test set and reference to the original paper if applicable - Check the API doc at http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To the my best knowledge, examples are either not affected by this change, or have been fixed to be compatible with this change ### Changes ### See description ## Comments ## - If this change is a backward incompatible change, why must this change be made. - Interesting edge cases to note here

3. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot

   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors. The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4. removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ### Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) - Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is well-documented: - For user-facing API changes, API doc string has been updated. - For new C++ functions in header files, their functionalities and arguments are documented. - For new examples, README.md is added to explain the what the example does, the source of the dataset, expected performance on test set and reference to the original paper if applicable - Check the API doc at http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To the my best knowledge, examples are either not affected by this change, or have been fixed to be compatible with this change ### Changes ### See description ## Comments ## - If this change is a backward incompatible change, why must this change be made. - Interesting edge cases to note here

4. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot

   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors. The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4. removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ### Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) - Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is well-documented: - For user-facing API changes, API doc string has been updated. - For new C++ functions in header files, their functionalities and arguments are documented. - For new examples, README.md is added to explain the what the example does, the source of the dataset, expected performance on test set and reference to the original paper if applicable - Check the API doc at http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To the my best knowledge, examples are either not affected by this change, or have been fixed to be compatible with this change ### Changes ### See description ## Comments ## - If this change is a backward incompatible change, why must this change be made. - Interesting edge cases to note here

   label: code-design

5. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot

   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors. The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4. removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ### Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) -

Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is well-documented: - For user-facing API changes, API doc string has been updated. - For new C++ functions in header files, their functionalities and arguments are documented. - For new examples, README.md is added to explain the what the example does, the source of the dataset, expected performance on test set and reference to the original paper if applicable - Check the API doc at http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To the my best knowledge, examples are either not affected by this change, or have been fixed to be compatible with this change ### Changes ### See description ## Comments ## - If this change is a backward incompatible change, why must this change be made. - Interesting edge cases to note here
label: test

6. **title:** [MXNET-771] Fix Flaky Test test_executor.py:test_dot
   body: ## Description ## Fix #11686, removed fixed seed in test_executor.py:test_dot 1. Verified this test is flaky and the values for comparison are slightly off 2. Flakiness comes from sym.dot of 2 large tensors. The input tensor size was original set to be a random size in range of (1,500). Reduced 500 to 200 to make it more stable. 3. The test still fails with 1e-6 tolerance, increased it to 1e-5 (reasonable for fp32) 4. removed helper function `reldiff` (calculating relative difference) with `assert_almost_equal` in `mxnet.test_utils` for value comparison 5. removed fixed seed 6. Passed 10,000 runs ## Checklist ## ### Essentials ### Please feel free to remove inapplicable items for your PR. - [x] The PR title starts with [MXNET-771](https://issues.apache.org/jira/browse/MXNET-771) created (except PRs with tiny changes) - [x] Changes are complete (i.e. I finished coding on this PR) - [x] All changes have test coverage: - Unit tests are added for small changes to verify correctness (e.g. adding a new operator) - Nightly tests are added for complicated/long-running ones (e.g. changing distributed kvstore) - Build tests will be added for build configuration changes (e.g. adding a new build option with NCCL) - [x] Code is well-documented: - For user-facing API changes, API doc string has been updated. - For new C++ functions in header files, their functionalities and arguments are documented. - For new examples, README.md is added to explain the what the example does, the source of the dataset, expected performance on test set and reference to the original paper if applicable - Check the API doc at http://mxnet-ci-doc.s3-accelerate.dualstack.amazonaws.com/PR-$PR_ID/$BUILD_ID/index.html - [x] To the my best knowledge, examples are either not affected by this change, or have been fixed to be compatible with this change ### Changes ### See description ## Comments ## - If this change is a backward incompatible change, why must this change be made. - Interesting edge cases to note here
   label: test

github_pulls_comments:

1. **body:** @kalyc since `reldiff` is also used in #11685 , with this change, the test_bind() does not fail deterministically when you remove fixed seed. However, it is still flaky and will fail with `MXNET_TEST_SEED=1875642038` please take a look
   label: test
2. How many trials have you tried with your new rtol and atol values?
3. @haojin2 10,000 runs, updated in description

github_pulls_reviews:

1. **body:** one nit: please also add something similar to this: https://github.com/apache/incubator-mxnet/pull/11888/files#diff-cb652780258e73a9cd08568f38929aa2R2509 to keep track of this flaky test.
   label: test

jira_issues:

1. **summary:** Flaky test in text_executor:test_dot
   description: [https://github.com/apache/incubator-mxnet/issues/11686]
2. **summary:** Flaky test in text_executor:test_dot
   description: [https://github.com/apache/incubator-mxnet/issues/11686]
3. **summary:** Flaky test in text_executor:test_dot
   description: [https://github.com/apache/incubator-mxnet/issues/11686]
   label: test

jira_issues_comments:

1. fixed with https://github.com/apache/incubator-mxnet/pull/11978