NumPyro Documentation

Uber AI Labs

Aug 12, 2020

Modeling:

CHAPTER 1

Pyro Primitives

1.1 param

param(*name*, *init_value=None*, ***kwargs*)

Annotate the given site as an optimizable parameter for use with `jax.experimental.optimizers`. For an example of how *param* statements can be used in inference algorithms, refer to `svi()`.

Parameters

- **name** (*str*) – name of site.
- **init_value** (*numpy.ndarray*) – initial value specified by the user. Note that the onus of using this to initialize the optimizer is on the user / inference algorithm, since there is no global parameter store in NumPyro.

Returns value for the parameter. Unless wrapped inside a handler like *substitute*, this will simply return the initial value.

1.2 sample

sample(*name*, *fn*, *obs=None*, *rng_key=None*, *sample_shape=()*, *infer=None*)

Returns a random sample from the stochastic function *fn*. This can have additional side effects when wrapped inside effect handlers like *substitute*.

Note: By design, *sample* primitive is meant to be used inside a NumPyro model. Then *seed* handler is used to inject a random state to *fn*. In those situations, *rng_key* keyword will take no effect.

Parameters

- **name** (*str*) – name of the sample site.
- **fn** – a stochastic function that returns a sample.

- **obs** (*numpy.ndarray*) – observed value
- **rng_key** (*jax.random.PRNGKey*) – an optional random key for *fn*.
- **sample_shape** – Shape of samples to be drawn.
- **infer** (*dict*) – an optional dictionary containing additional information for inference algorithms. For example, if *fn* is a discrete distribution, setting *infer={'enumerate': 'parallel'}* to tell MCMC marginalize this discrete latent site.

Returns sample from the stochastic *fn*.

1.3 plate

class plate(*name*, *size*, *subsample_size=None*, *dim=None*)

Construct for annotating conditionally independent variables. Within a *plate* context manager, *sample* sites will be automatically broadcasted to the size of the plate. Additionally, a scale factor might be applied by certain inference algorithms if *subsample_size* is specified.

Parameters

- **name** (*str*) – Name of the plate.
- **size** (*int*) – Size of the plate.
- **subsample_size** (*int*) – Optional argument denoting the size of the mini-batch. This can be used to apply a scaling factor by inference algorithms. e.g. when computing ELBO using a mini-batch.
- **dim** (*int*) – Optional argument to specify which dimension in the tensor is used as the plate dim. If *None* (default), the leftmost available dim is allocated.

1.4 plate_stack

plate_stack(*prefix*, *sizes*, *rightmost_dim=-1*)

Create a contiguous stack of *plate* s with dimensions:

```
rightmost_dim - len(sizes), ..., rightmost_dim
```

Parameters

- **prefix** (*str*) – Name prefix for plates.
- **sizes** (*iterable*) – An iterable of plate sizes.
- **rightmost_dim** (*int*) – The rightmost dim, counting from the right.

1.5 deterministic

deterministic(*name*, *value*)

Used to designate deterministic sites in the model. Note that most effect handlers will not operate on deterministic sites (except *trace()*), so deterministic sites should be side-effect free. The use case for deterministic nodes is to record any values in the model execution trace.

Parameters

- **name** (*str*) – name of the deterministic site.
- **value** (*numpy.ndarray*) – deterministic value to record in the trace.

1.6 factor

factor(*name*, *log_factor*)

Factor statement to add arbitrary log probability factor to a probabilistic model.

Parameters

- **name** (*str*) – Name of the trivial sample.
- **log_factor** (*numpy.ndarray*) – A possibly batched log probability factor.

1.7 module

module(*name*, *nn*, *input_shape=None*)

Declare a `stax` style neural network inside a model so that its parameters are registered for optimization via *param()* statements.

Parameters

- **name** (*str*) – name of the module to be registered.
- **nn** (*tuple*) – a tuple of *(init_fn, apply_fn)* obtained by a `stax` constructor function.
- **input_shape** (*tuple*) – shape of the input taken by the neural network.

Returns a *apply_fn* with bound parameters that takes an array as an input and returns the neural network transformed output array.

1.8 scan

scan(*f*, *init*, *xs*, *length=None*, *reverse=False*)

This primitive scans a function over the leading array axes of *xs* while carrying along state. See `jax.lax.scan()` for more information.

Usage:

```
>>> import numpy as np
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.contrib.control_flow import scan
>>>
>>> def gaussian_hmm(y=None, T=10):
...     def transition(x_prev, y_curr):
...         x_curr = numpyro.sample('x', dist.Normal(x_prev, 1))
...         y_curr = numpyro.sample('y', dist.Normal(x_curr, 1), obs=y_curr)
...         return x_curr, (x_curr, y_curr)
...
...     x0 = numpyro.sample('x_0', dist.Normal(0, 1))
...     _, (x, y) = scan(transition, x0, y, length=T)
...     return (x, y)
>>>
```

(continues on next page)

(continued from previous page)

```
>>> # here we do some quick tests
>>> with numpyro.handlers.seed(rng_seed=0):
...     x, y = gaussian_hmm(np.arange(10.))
>>> assert x.shape == (10,) and y.shape == (10,)
>>> assert np.all(y == np.arange(10))
>>>
>>> with numpyro.handlers.seed(rng_seed=0):  # generative
...     x, y = gaussian_hmm()
>>> assert x.shape == (10,) and y.shape == (10,)
```

Warning: This is an experimental utility function that allows users to use JAX control flow with NumPyro's effect handlers. Currently, *sample* and *deterministic* sites within the scan body *f* are supported. If you notice that any effect handlers or distributions are unsupported, please file an issue.

Note: It is ambiguous to align *scan* dimension inside a *plate* context. So the following pattern won't be supported

```
with numpyro.plate('N', 10):
    last, ys = scan(f, init, xs)
```

All *plate* statements should be put inside *f*. For example, the corresponding working code is

```
def g(*args, **kwargs):
    with numpyro.plate('N', 10):
        return f(*arg, **kwargs)

last, ys = scan(g, init, xs)
```

Note: Nested scan is currently not supported.

Note: We can scan over discrete latent variables in *f*. The joint density is evaluated using parallel-scan (reference [1]) over time dimension, which reduces parallel complexity to *O(log(length))*.

Currently, only the equivalence to `markov(history_size=1)` is supported. A `trace` of *scan* with discrete latent variables will contain the following sites:

- **init sites: those sites belong to the first trace of *f*. Each of** them will have name prefixed with *_init/*.
- **scanned sites: those sites collect the values of the remaining scan** loop over *f*. An addition time dimension *_time_foo* will be added to those sites, where *foo* is the name of the first site appeared in *f*.

Not all transition functions *f* are supported. All of the restrictions from Pyro's enumeration tutorial [2] still apply here. In addition, there should not have any site outside of *scan* depend on the first output of *scan* (the last carry value).

** References **

1. *Temporal Parallelization of Bayesian Smoothers*, Simo Sarkka, Angel F. Garcia-Fernandez (https://arxiv.org/abs/1905.13002)

2. *Inference with Discrete Latent Variables* (http://pyro.ai/examples/enumeration.html#Dependencies-among-plates)

Parameters

- **f** (*callable*) – a function to be scanned.
- **init** – the initial carrying state
- **xs** – the values over which we scan along the leading axis. This can be any JAX pytree (e.g. list/dict of arrays).
- **length** – optional value specifying the length of *xs* but can be used when *xs* is an empty pytree (e.g. None)
- **reverse** (*bool*) – optional boolean specifying whether to run the scan iteration forward (the default) or in reverse

Returns output of scan, quoted from `jax.lax.scan()` docs: "pair of type (c, [b]) where the first element represents the final loop carry value and the second element represents the stacked outputs of the second output of f when scanned over the leading axis of the inputs".

CHAPTER 2

Effect Handlers

This provides a small set of effect handlers in NumPyro that are modeled after Pyro's poutine module. For a tutorial on effect handlers more generally, readers are encouraged to read Poutine: A Guide to Programming with Effect Handlers in Pyro. These simple effect handlers can be composed together or new ones added to enable implementation of custom inference utilities and algorithms.

Example

As an example, we are using *seed*, *trace* and *substitute* handlers to define the *log_likelihood* function below. We first create a logistic regression model and sample from the posterior distribution over the regression parameters using `MCMC()`. The *log_likelihood* function uses effect handlers to run the model by substituting sample sites with values from the posterior distribution and computes the log density for a single data point. The *log_predictive_density* function computes the log likelihood for each draw from the joint posterior and aggregates the results for all the data points, but does so by using JAX's auto-vectorize transform called *vmap* so that we do not need to loop over all the data points.

```
>>> import jax.numpy as jnp
>>> from jax import random, vmap
>>> from jax.scipy.special import logsumexp
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro import handlers
>>> from numpyro.infer import MCMC, NUTS

>>> N, D = 3000, 3
>>> def logistic_regression(data, labels):
...     coefs = numpyro.sample('coefs', dist.Normal(jnp.zeros(D), jnp.ones(D)))
...     intercept = numpyro.sample('intercept', dist.Normal(0., 10.))
...     logits = jnp.sum(coefs * data + intercept, axis=-1)
...     return numpyro.sample('obs', dist.Bernoulli(logits=logits), obs=labels)

>>> data = random.normal(random.PRNGKey(0), (N, D))
>>> true_coefs = jnp.arange(1., D + 1.)
>>> logits = jnp.sum(true_coefs * data, axis=-1)
>>> labels = dist.Bernoulli(logits=logits).sample(random.PRNGKey(1))
```

(continues on next page)

(continued from previous page)

```
>>> num_warmup, num_samples = 1000, 1000
>>> mcmc = MCMC(NUTS(model=logistic_regression), num_warmup, num_samples)
>>> mcmc.run(random.PRNGKey(2), data, labels)
sample: 100%|| 1000/1000 [00:00<00:00, 1252.39it/s, 1 steps of size 5.83e-01. acc.
↪prob=0.85]
>>> mcmc.print_summary()


                mean       sd       5.5%      94.5%      n_eff       Rhat
    coefs[0]       0.96       0.07       0.85       1.07     455.35       1.01
    coefs[1]       2.05       0.09       1.91       2.20     332.00       1.01
    coefs[2]       3.18       0.13       2.96       3.37     320.27       1.00
   intercept      -0.03       0.02      -0.06       0.00     402.53       1.00

>>> def log_likelihood(rng_key, params, model, *args, **kwargs):
...     model = handlers.substitute(handlers.seed(model, rng_key), params)
...     model_trace = handlers.trace(model).get_trace(*args, **kwargs)
...     obs_node = model_trace['obs']
...     return obs_node['fn'].log_prob(obs_node['value'])

>>> def log_predictive_density(rng_key, params, model, *args, **kwargs):
...     n = list(params.values())[0].shape[0]
...     log_lk_fn = vmap(lambda rng_key, params: log_likelihood(rng_key, params,
↪model, *args, **kwargs))
...     log_lk_vals = log_lk_fn(random.split(rng_key, n), params)
...     return jnp.sum(logsumexp(log_lk_vals, 0) - jnp.log(n))

>>> print(log_predictive_density(random.PRNGKey(2), mcmc.get_samples(),
...       logistic_regression, data, labels))
-874.89813
```

2.1 block

class block(*fn=None*, *hide_fn=None*, *hide=None*)

Bases: numpyro.primitives.Messenger

Given a callable *fn*, return another callable that selectively hides primitive sites where *hide_fn* returns True from other effect handlers on the stack.

Parameters

- **fn** – Python callable with NumPyro primitives.
- **hide_fn** – function which when given a dictionary containing site-level metadata returns whether it should be blocked.

Example:

```
>>> from jax import random
>>> import numpyro
>>> from numpyro.handlers import block, seed, trace
>>> import numpyro.distributions as dist

>>> def model():
...     a = numpyro.sample('a', dist.Normal(0., 1.))
```

(continues on next page)

(continued from previous page)

```
...     return numpyro.sample('b', dist.Normal(a, 1.))

>>> model = seed(model, random.PRNGKey(0))
>>> block_all = block(model)
>>> block_a = block(model, lambda site: site['name'] == 'a')
>>> trace_block_all = trace(block_all).get_trace()
>>> assert not {'a', 'b'}.intersection(trace_block_all.keys())
>>> trace_block_a =  trace(block_a).get_trace()
>>> assert 'a' not in trace_block_a
>>> assert 'b' in trace_block_a
```

process_message(*msg*)

2.2 condition

class condition(*fn=None*, *data=None*, *condition_fn=None*)

Bases: numpyro.primitives.Messenger

Conditions unobserved sample sites to values from *data* or *condition_fn*. Similar to *substitute* except that it only affects *sample* sites and changes the *is_observed* property to *True*.

Parameters

- **fn** – Python callable with NumPyro primitives.
- **data** (*dict*) – dictionary of *numpy.ndarray* values keyed by site names.
- **condition_fn** – callable that takes in a site dict and returns a numpy array or *None* (in which case the handler has no side effect).

Example:

```
>>> from jax import random
>>> import numpyro
>>> from numpyro.handlers import condition, seed, substitute, trace
>>> import numpyro.distributions as dist

>>> def model():
...     numpyro.sample('a', dist.Normal(0., 1.))

>>> model = seed(model, random.PRNGKey(0))
>>> exec_trace = trace(condition(model, {'a': -1})).get_trace()
>>> assert exec_trace['a']['value'] == -1
>>> assert exec_trace['a']['is_observed']
```

process_message(*msg*)

2.3 do

class do(*fn=None*, *data=None*)

Bases: numpyro.primitives.Messenger

Given a stochastic function with some sample statements and a dictionary of values at names, set the return values of those sites equal to the values as if they were hard-coded to those values and introduce fresh sample sites with the same names whose values do not propagate. Composes freely with *condition()* to represent

counterfactual distributions over potential outcomes. See Single World Intervention Graphs [1] for additional details and theory.

This is equivalent to replacing *z = numpyro.sample("z", ...)* with *z = 1.* and introducing a fresh sample site *numpyro.sample("z", ...)* whose value is not used elsewhere.

References

[1] ***Single World Intervention Graphs: A Primer***, Thomas Richardson, James Robins

Parameters

- **fn** – a stochastic function (callable containing Pyro primitive calls)
- **data** – a `dict` mapping sample site names to interventions

Example:

```
>>> import jax.numpy as jnp
>>> import numpyro
>>> from numpyro.handlers import do, trace, seed
>>> import numpyro.distributions as dist
>>> def model(x):
...     s = numpyro.sample("s", dist.LogNormal())
...     z = numpyro.sample("z", dist.Normal(x, s))
...     return z ** 2
>>> intervened_model = handlers.do(model, data={"z": 1.})
>>> with trace() as exec_trace:
...     z_square = seed(intervened_model, 0)(1)
>>> assert exec_trace['z']['value'] != 1.
>>> assert not exec_trace['z']['is_observed']
>>> assert not exec_trace['z'].get('stop', None)
>>> assert z_square == 1
```

process_message(*msg*)

2.4 lift

class lift(*fn=None*, *prior=None*)

Bases: `numpyro.primitives.Messenger`

Given a stochastic function with `param` calls and a prior distribution, create a stochastic function where all param calls are replaced by sampling from prior. Prior should be a distribution or a dict of names to distributions.

Consider the following NumPyro program:

```
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.handlers import lift
>>>
>>> def model(x):
...     s = numpyro.param("s", 0.5)
...     z = numpyro.sample("z", dist.Normal(x, s))
...     return z ** 2
>>> lifted_model = lift(model, prior={"s": dist.Exponential(0.3)})
```

`lift` makes `param` statements behave like `sample` statements using the distributions in `prior`. In this example, site *s* will now behave as if it was replaced with `s = numpyro.sample("s", dist.Exponential(0.3))`.

Parameters

- **fn** – function whose parameters will be lifted to random values
- **prior** – prior function in the form of a Distribution or a dict of Distributions

process_message(*msg*)

2.5 mask

class mask(*fn=None*, *mask=True*)

Bases: `numpyro.primitives.Messenger`

This messenger masks out some of the sample statements elementwise.

Parameters **mask** – a boolean or a boolean-valued array for masking elementwise log probability of sample sites (*True* includes a site, *False* excludes a site).

process_message(*msg*)

2.6 reparam

class reparam(*fn=None*, *config=None*)

Bases: `numpyro.primitives.Messenger`

Reparametrizes each affected sample site into one or more auxiliary sample sites followed by a deterministic transformation [1].

To specify reparameterizers, pass a `config` dict or callable to the constructor. See the *numpyro.infer.reparam* module for available reparameterizers.

Note some reparameterizers can examine the `*args,**kwargs` inputs of functions they affect; these reparameterizers require using `handlers.reparam` as a decorator rather than as a context manager.

[1] Maria I. Gorinova, Dave Moore, Matthew D. Hoffman (2019) "Automatic Reparameterisation of Probabilistic Programs" https://arxiv.org/pdf/1906.03028.pdf

Parameters **config** (*dict or callable*) – Configuration, either a dict mapping site name to *Reparam*, or a function mapping site to *Reparam* or None.

process_message(*msg*)

2.7 replay

class replay(*fn=None*, *guide_trace=None*)

Bases: `numpyro.primitives.Messenger`

Given a callable *fn* and an execution trace *guide_trace*, return a callable which substitutes *sample* calls in *fn* with values from the corresponding site names in *guide_trace*.

Parameters

- **fn** – Python callable with NumPyro primitives.
- **guide_trace** – an OrderedDict containing execution metadata.

Example

```
>>> from jax import random
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.handlers import replay, seed, trace

>>> def model():
...     numpyro.sample('a', dist.Normal(0., 1.))

>>> exec_trace = trace(seed(model, random.PRNGKey(0))).get_trace()
>>> print(exec_trace['a']['value'])
-0.20584235
>>> replayed_trace = trace(replay(model, exec_trace)).get_trace()
>>> print(exec_trace['a']['value'])
-0.20584235
>>> assert replayed_trace['a']['value'] == exec_trace['a']['value']
```

process_message(*msg*)

2.8 scale

class scale(*fn=None*, *scale=1.0*)

Bases: numpyro.primitives.Messenger

This messenger rescales the log probability score.

This is typically used for data subsampling or for stratified sampling of data (e.g. in fraud detection where negatives vastly outnumber positives).

Parameters **scale** (*float*) – a positive scaling factor

process_message(*msg*)

2.9 scope

class scope(*fn=None*, *prefix=''*)

Bases: numpyro.primitives.Messenger

This handler prepend a prefix followed by a / to the name of sample sites.

Example:

```
.. doctest::
```

```
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.handlers import scope, seed, trace
>>>
>>> def model():
...     with scope(prefix="a"):
...         with scope(prefix="b"):
```

(continues on next page)

(continued from previous page)

```
...             return numpyro.sample("x", dist.Bernoulli(0.5))
...
>>> assert "a/b/x" in trace(seed(model, 0)).get_trace()
```

Parameters

- **fn** – Python callable with NumPyro primitives.
- **prefix** (*str*) – a string to prepend to sample names

process_message(*msg*)

2.10 seed

class seed(*fn=None*, *rng_seed=None*)

Bases: numpyro.primitives.Messenger

JAX uses a functional pseudo random number generator that requires passing in a seed `PRNGKey()` to every stochastic function. The *seed* handler allows us to initially seed a stochastic function with a `PRNGKey()`. Every call to the `sample()` primitive inside the function results in a splitting of this initial seed so that we use a fresh seed for each subsequent call without having to explicitly pass in a *PRNGKey* to each *sample* call.

Parameters

- **fn** – Python callable with NumPyro primitives.
- **rng_seed** (*int, jnp.ndarray scalar, or jax.random.PRNGKey*) – a random number generator seed.

Note: Unlike in Pyro, *numpyro.sample* primitive cannot be used without wrapping it in seed handler since there is no global random state. As such, users need to use *seed* as a contextmanager to generate samples from distributions or as a decorator for their model callable (See below).

Example:

```
>>> from jax import random
>>> import numpyro
>>> import numpyro.handlers
>>> import numpyro.distributions as dist

>>> # as context manager
>>> with handlers.seed(rng_seed=1):
...     x = numpyro.sample('x', dist.Normal(0., 1.))

>>> def model():
...     return numpyro.sample('y', dist.Normal(0., 1.))

>>> # as function decorator (/modifier)
>>> y = handlers.seed(model, rng_seed=1)()
>>> assert x == y
```

process_message(*msg*)

2.11 substitute

class substitute (*fn=None*, *data=None*, *substitute_fn=None*)

Bases: numpyro.primitives.Messenger

Given a callable *fn* and a dict *data* keyed by site names (alternatively, a callable *substitute_fn*), return a callable which substitutes all primitive calls in *fn* with values from *data* whose key matches the site name. If the site name is not present in *data*, there is no side effect.

If a *substitute_fn* is provided, then the value at the site is replaced by the value returned from the call to *substitute_fn* for the given site.

Parameters

- **fn** – Python callable with NumPyro primitives.
- **data** (*dict*) – dictionary of *numpy.ndarray* values keyed by site names.
- **substitute_fn** – callable that takes in a site dict and returns a numpy array or *None* (in which case the handler has no side effect).

Example:

```
>>> from jax import random
>>> import numpyro
>>> from numpyro.handlers import seed, substitute, trace
>>> import numpyro.distributions as dist

>>> def model():
...     numpyro.sample('a', dist.Normal(0., 1.))

>>> model = seed(model, random.PRNGKey(0))
>>> exec_trace = trace(substitute(model, {'a': -1})).get_trace()
>>> assert exec_trace['a']['value'] == -1
```

process_message (*msg*)

2.12 trace

class trace (*fn=None*)

Bases: numpyro.primitives.Messenger

Returns a handler that records the inputs and outputs at primitive calls inside *fn*.

Example

```
>>> from jax import random
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.handlers import seed, trace
>>> import pprint as pp

>>> def model():
...     numpyro.sample('a', dist.Normal(0., 1.))

>>> exec_trace = trace(seed(model, random.PRNGKey(0))).get_trace()
>>> pp.pprint(exec_trace)
OrderedDict([('a',
```

(continues on next page)

(continued from previous page)

```
             {'args': (),
              'fn': <numpyro.distributions.continuous.Normal object at␣
↪0x7f9e689b1eb8>,
              'is_observed': False,
              'kwargs': {'rng_key': DeviceArray([0, 0], dtype=uint32)},
              'name': 'a',
              'type': 'sample',
              'value': DeviceArray(-0.20584235, dtype=float32)})])
```

postprocess_message(*msg*)

get_trace(**args*, ***kwargs*)

Run the wrapped callable and return the recorded trace.

Parameters

- ***args** – arguments to the callable.
- ****kwargs** – keyword arguments to the callable.

Returns *OrderedDict* containing the execution trace.

CHAPTER 3

Base Distribution

3.1 Distribution

class Distribution(*batch_shape=()*, *event_shape=()*, *validate_args=None*)

Bases: `object`

Base class for probability distributions in NumPyro. The design largely follows from `torch.distributions`.

Parameters

- **batch_shape** – The batch shape for the distribution. This designates independent (possibly non-identical) dimensions of a sample from the distribution. This is fixed for a distribution instance and is inferred from the shape of the distribution parameters.
- **event_shape** – The event shape for the distribution. This designates the dependent dimensions of a sample from the distribution. These are collapsed when we evaluate the log probability density of a batch of samples using *.log_prob*.
- **validate_args** – Whether to enable validation of distribution parameters and arguments to *.log_prob* method.

As an example:

```
>>> import jax.numpy as jnp
>>> import numpyro.distributions as dist
>>> d = dist.Dirichlet(jnp.ones((2, 3, 4)))
>>> d.batch_shape
(2, 3)
>>> d.event_shape
(4,)
```

arg_constraints = {}

support = None

has_enumerate_support = False

is_discrete = False

reparametrized_params = []

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

static set_default_validate_args(*value*)

batch_shape
Returns the shape over which the distribution parameters are batched.

Returns batch shape of the distribution.

Return type tuple

event_shape
Returns the shape of a single sample from the distribution without batching.

Returns event shape of the distribution.

Return type tuple

event_dim

Returns Number of dimensions of individual events.

Return type int

shape(*sample_shape=()*)
The tensor shape of samples from this distribution.

Samples are of shape:

```
d.shape(sample_shape) == sample_shape + d.batch_shape + d.event_shape
```

Parameters **sample_shape** (*tuple*) – the size of the iid batch to be drawn from the distribution.

Returns shape of samples.

Return type tuple

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

sample_with_intermediates(*key*, *sample_shape=()*)
Same as `sample` except that any intermediate computations are returned (useful for *TransformedDistribution*).

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.

- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob(*value*)
Evaluates the log probability density for a batch of samples given by *value*.

Parameters **value** – A batch of samples from the distribution.

Returns an array with shape *value.shape[:-self.event_shape]*

Return type numpy.ndarray

mean
Mean of the distribution.

variance
Variance of the distribution.

to_event(*reinterpreted_batch_ndims=None*)
Interpret the rightmost *reinterpreted_batch_ndims* batch dimensions as dependent event dimensions.

Parameters **reinterpreted_batch_ndims** – Number of rightmost batch dims to interpret as event dims.

Returns An instance of *Independent* distribution.

Return type *Independent*

enumerate_support(*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

expand(*batch_shape*)
Returns a new *ExpandedDistribution* instance with batch dimensions expanded to *batch_shape*.

Parameters **batch_shape** (*tuple*) – batch shape to expand to.

Returns an instance of *ExpandedDistribution*.

Return type *ExpandedDistribution*

expand_by(*sample_shape*)
Expands a distribution by adding `sample_shape` to the left side of its *batch_shape*. To expand internal dims of `self.batch_shape` from 1 to something larger, use *expand()* instead.

Parameters **sample_shape** (*tuple*) – The size of the iid batch to be drawn from the distribution.

Returns An expanded version of this distribution.

Return type *ExpandedDistribution*

mask(*mask*)
Masks a distribution by a boolean or boolean-valued array that is broadcastable to the distributions *Distribution.batch_shape*.

Parameters **mask** (*bool or jnp.ndarray*) – A boolean or boolean valued array (*True* includes a site, *False* excludes a site).

Returns A masked copy of this distribution.

Return type *MaskedDistribution*

3.2 ExpandedDistribution

class ExpandedDistribution(*base_dist*, *batch_shape=()*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {}

expand(*batch_shape*)

Returns a new *ExpandedDistribution* instance with batch dimensions expanded to *batch_shape*.

Parameters **batch_shape** (*tuple*) – batch shape to expand to.

Returns an instance of *ExpandedDistribution*.

Return type *ExpandedDistribution*

has_enumerate_support

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

is_discrete

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

support

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(*value*)

Evaluates the log probability density for a batch of samples given by *value*.

Parameters **value** – A batch of samples from the distribution.

Returns an array with shape *value.shape[:-self.event_shape]*

Return type numpy.ndarray

enumerate_support(*expand=True*)

Returns an array with shape *len(support) x batch_shape* containing all values in the support.

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

3.3 ImproperUniform

class ImproperUniform(*support*, *batch_shape*, *event_shape*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

A helper distribution with zero *log_prob()* over the *support* domain.

Note: *sample* method is not implemented for this distribution. In autoguide and mcmc, initial parameters for improper sites are derived from *init_to_uniform* or *init_to_value* strategies.

Usage:

```
>>> from numpyro import sample
>>> from numpyro.distributions import ImproperUniform, Normal, constraints
>>>
>>> def model():
...     # ordered vector with length 10
...     x = sample('x', ImproperUniform(constraints.ordered_vector, (), event_
↪shape=(10,)))
...
...     # real matrix with shape (3, 4)
...     y = sample('y', ImproperUniform(constraints.real, (), event_shape=(3, 4)))
...
...     # a shape-(6, 8) batch of length-5 vectors greater than 3
...     z = sample('z', ImproperUniform(constraints.greater_than(3), (6, 8),␣
↪event_shape=(5,)))
```

If you want to set improper prior over all values greater than a, where a is another random variable, you might use

```
>>> def model():
...     a = sample('a', Normal(0, 1))
...     x = sample('x', ImproperUniform(constraints.greater_than(a), (), event_
↪shape=()))
```

or if you want to reparameterize it

```
>>> from numpyro.distributions import TransformedDistribution, transforms
>>> from numpyro.handlers import reparam
>>> from numpyro.infer.reparam import TransformReparam
>>>
>>> def model():
...     a = sample('a', Normal(0, 1))
...     with reparam(config={'x': TransformReparam()}):
...         x = sample('x',
...                    TransformedDistribution(ImproperUniform(constraints.
↪positive, (), ()),
...                                            transforms.AffineTransform(a, 1)))
```

Parameters

- **support** (*Constraint*) – the support of this distribution.
- **batch_shape** (*tuple*) – batch shape of this distribution. It is usually safe to set *batch_shape=()*.
- **event_shape** (*tuple*) – event shape of this distribution.

arg_constraints = {}

log_prob(**args*, ***kwargs*)

tree_flatten()

3.4 Independent

class Independent(*base_dist*, *reinterpreted_batch_ndims*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

Reinterprets batch dimensions of a distribution as event dims by shifting the batch-event dim boundary further to the left.

From a practical standpoint, this is useful when changing the result of *log_prob()*. For example, a univariate Normal distribution can be interpreted as a multivariate Normal with diagonal covariance:

```
>>> import numpyro.distributions as dist
>>> normal = dist.Normal(jnp.zeros(3), jnp.ones(3))
>>> [normal.batch_shape, normal.event_shape]
[(3,), ()]
>>> diag_normal = dist.Independent(normal, 1)
>>> [diag_normal.batch_shape, diag_normal.event_shape]
[(), (3,)]
```

Parameters

- **base_distribution** (*numpyro.distribution.Distribution*) – a distribution instance.
- **reinterpreted_batch_ndims** (*int*) – the number of batch dims to reinterpret as event dims.

arg_constraints = {}

support

has_enumerate_support

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

is_discrete

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

reparameterized_params

mean

Mean of the distribution.

variance

Variance of the distribution.

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*.

Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob(*value*)

Evaluates the log probability density for a batch of samples given by *value*.

Parameters **value** – A batch of samples from the distribution.

Returns an array with shape *value.shape[:-self.event_shape]*

Return type numpy.ndarray

expand(*batch_shape*)

Returns a new *ExpandedDistribution* instance with batch dimensions expanded to *batch_shape*.

Parameters **batch_shape** (*tuple*) – batch shape to expand to.

Returns an instance of *ExpandedDistribution*.

Return type *ExpandedDistribution*

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

3.5 MaskedDistribution

class MaskedDistribution(*base_dist*, *mask*)

Bases: *numpyro.distributions.distribution.Distribution*

Masks a distribution by a boolean array that is broadcastable to the distribution's *Distribution.batch_shape*. In the special case `mask is False`, computation of *log_prob()* , is skipped, and constant zero values are returned instead.

Parameters **mask** (*jnp.ndarray or bool*) – A boolean or boolean-valued array.

arg_constraints = {}

has_enumerate_support

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

is_discrete

bool(x) -> bool

Returns True when the argument x is true, False otherwise. The builtins True and False are the only two instances of the class bool. The class bool is a subclass of the class int, and cannot be subclassed.

support

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(*value*)
Evaluates the log probability density for a batch of samples given by *value*.

Parameters **value** – A batch of samples from the distribution.

Returns an array with shape *value.shape[:-self.event_shape]*

Return type numpy.ndarray

enumerate_support(*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

mean
Mean of the distribution.

variance
Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

3.6 TransformedDistribution

class TransformedDistribution(*base_distribution*, *transforms*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

Returns a distribution instance obtained as a result of applying a sequence of transforms to a base distribution. For an example, see `LogNormal` and `HalfNormal`.

Parameters

- **base_distribution** – the base distribution over which to apply transforms.
- **transforms** – a single transform or a list of transforms.
- **validate_args** – Whether to enable validation of distribution parameters and arguments to *.log_prob* method.

arg_constraints = {}

support

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

sample_with_intermediates(*key*, *sample_shape=()*)
Same as `sample` except that any intermediate computations are returned (useful for *TransformedDistribution*).

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

tree_flatten()

3.7 Unit

class Unit(*log_factor*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

Trivial nonnormalized distribution representing the unit type.

The unit type has a single value with no data, i.e. `value.size == 0`.

This is used for `numpyro.factor()` statements.

arg_constraints = {'log_factor': <numpyro.distributions.constraints._Real object>}

support = <numpyro.distributions.constraints._Real object>

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(*value*)
Evaluates the log probability density for a batch of samples given by *value*.

Parameters **value** – A batch of samples from the distribution.

Returns an array with shape *value.shape[:-self.event_shape]*

Return type numpy.ndarray

CHAPTER 4

Continuous Distributions

4.1 Beta

class Beta(*concentration1*, *concentration0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'concentration0': <numpyro.distributions.constraints._GreaterThan

support = <numpyro.distributions.constraints._Interval object>

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

4.2 Cauchy

class Cauchy(*loc=0.0*, *scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'scale']

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

4.3 Chi2

class Chi2(*df*, *validate_args=None*)

Bases: *numpyro.distributions.continuous.Gamma*

arg_constraints = {'df': <numpyro.distributions.constraints._GreaterThan object>}

4.4 Dirichlet

class Dirichlet(*concentration*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._Simplex object>

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.

- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.5 Exponential

class Exponential(*rate=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

reparametrized_params = ['rate']

arg_constraints = {'rate': <numpyro.distributions.constraints._GreaterThan object>}

support = <numpyro.distributions.constraints._GreaterThan object>

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape + batch_shape + event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.6 Gamma

class Gamma(*concentration*, *rate=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._GreaterThan object>

reparametrized_params = ['rate']

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.7 Gumbel

class Gumbel (*loc=0.0*, *scale=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'scale']

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.8 GaussianRandomWalk

class GaussianRandomWalk (*scale=1.0*, *num_steps=1*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'num_steps': <numpyro.distributions.constraints._IntegerGreaterTha

support = <numpyro.distributions.constraints._RealVector object>

reparametrized_params = ['scale']

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.9 HalfCauchy

class HalfCauchy(*scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

reparametrized_params = ['scale']

support = <numpyro.distributions.constraints._GreaterThan object>

arg_constraints = {'scale': <numpyro.distributions.constraints._GreaterThan object>}

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance
Variance of the distribution.

4.10 HalfNormal

class HalfNormal(*scale=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

reparametrized_params = ['scale']

support = <numpyro.distributions.constraints._GreaterThan object>

arg_constraints = {'scale': <numpyro.distributions.constraints._GreaterThan object>}

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.11 InverseGamma

class InverseGamma(*concentration*, *rate=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._GreaterThan object>

reparametrized_params = ['rate']

mean
Mean of the distribution.

variance
Variance of the distribution.

tree_flatten()

4.12 Laplace

class Laplace(*loc=0.0*, *scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'scale']

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

4.13 LKJ

class LKJ(*dimension*, *concentration=1.0*, *sample_method='onion'*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.TransformedDistribution*

LKJ distribution for correlation matrices. The distribution is controlled by `concentration` parameter η to make the probability of the correlation matrix M propotional to $\det(M)^{\eta-1}$. Because of that, when `concentration == 1`, we have a uniform distribution over correlation matrices.

When `concentration > 1`, the distribution favors samples with large large determinent. This is useful when we know a priori that the underlying variables are not correlated.

When `concentration < 1`, the distribution favors samples with small determinent. This is useful when we know a priori that some underlying variables are correlated.

Parameters

- **dimension** (*int*) – dimension of the matrices
- **concentration** (*ndarray*) – concentration/shape parameter of the distribution (often referred to as eta)
- **sample_method** (*str*) – Either "cvine" or "onion". Both methods are proposed in [1] and offer the same distribution over correlation matrices. But they are different in how to generate samples. Defaults to "onion".

References

[1] *Generating random correlation matrices based on vines and extended onion method*, Daniel Lewandowski, Dorota Kurowicka, Harry Joe

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._CorrMatrix object>

mean
Mean of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.14 LKJCholesky

class LKJCholesky(*dimension*, *concentration=1.0*, *sample_method='onion'*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

LKJ distribution for lower Cholesky factors of correlation matrices. The distribution is controlled by `concentration` parameter η to make the probability of the correlation matrix M generated from a Cholesky factor propotional to $\det(M)^{\eta-1}$. Because of that, when `concentration == 1`, we have a uniform distribution over Cholesky factors of correlation matrices.

When `concentration > 1`, the distribution favors samples with large diagonal entries (hence large determinent). This is useful when we know a priori that the underlying variables are not correlated.

When `concentration < 1`, the distribution favors samples with small diagonal entries (hence small determinent). This is useful when we know a priori that some underlying variables are correlated.

Parameters

- **dimension** (*int*) – dimension of the matrices
- **concentration** (*ndarray*) – concentration/shape parameter of the distribution (often referred to as eta)
- **sample_method** (*str*) – Either "cvine" or "onion". Both methods are proposed in [1] and offer the same distribution over correlation matrices. But they are different in how to generate samples. Defaults to "onion".

References

[1] *Generating random correlation matrices based on vines and extended onion method*, Daniel Lewandowski, Dorota Kurowicka, Harry Joe

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._CorrCholesky object>

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.15 LogNormal

class LogNormal(*loc=0.0*, *scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

reparametrized_params = ['loc', 'scale']

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

4.16 Logistic

class Logistic(*loc=0.0*, *scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'real']

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

4.17 MultivariateNormal

class MultivariateNormal(*loc=0.0*, *covariance_matrix=None*, *precision_matrix=None*, *scale_tril=None*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'covariance_matrix': <numpyro.distributions.constraints._PositiveD

support = <numpyro.distributions.constraints._RealVector object>

reparametrized_params = ['loc', 'covariance_matrix', 'precision_matrix', 'scale_tril']

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

covariance_matrix

precision_matrix

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.18 LowRankMultivariateNormal

class LowRankMultivariateNormal(*loc*, *cov_factor*, *cov_diag*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'cov_diag': <numpyro.distributions.constraints._GreaterThan object

support = <numpyro.distributions.constraints._RealVector object>

mean

Mean of the distribution.

variance

scale_tril

covariance_matrix

precision_matrix

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

entropy()

4.19 Normal

class Normal(*loc=0.0*, *scale=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'scale':

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'scale']

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

icdf(*q*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.20 Pareto

class Pareto(*scale*, *alpha*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'alpha': <numpyro.distributions.constraints._GreaterThan object>,

mean
Mean of the distribution.

variance
Variance of the distribution.

support

tree_flatten()

4.21 StudentT

class StudentT(*df*, *loc=0.0*, *scale=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'df': <numpyro.distributions.constraints._GreaterThan object>, 'lo

support = <numpyro.distributions.constraints._Real object>

reparametrized_params = ['loc', 'scale']

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

4.22 TruncatedCauchy

class TruncatedCauchy(*low=0.0*, *loc=0.0*, *scale=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'low': <

reparametrized_params = ['low', 'loc', 'scale']

support

mean
Mean of the distribution.

variance
Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.23 TruncatedNormal

class TruncatedNormal(*low=0.0*, *loc=0.0*, *scale=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'loc': <numpyro.distributions.constraints._Real object>, 'low': <

reparametrized_params = ['low', 'loc', 'scale']

support

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.24 TruncatedPolyaGamma

class TruncatedPolyaGamma(*batch_shape=()*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

truncation_point = 2.5

num_log_prob_terms = 7

num_gamma_variates = 8

arg_constraints = {}

support = <numpyro.distributions.constraints._Interval object>

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

4.25 Uniform

class Uniform(*low=0.0*, *high=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.TransformedDistribution*

arg_constraints = {'high': <numpyro.distributions.constraints._Dependent object>, 'lo

reparametrized_params = ['low', 'high']

support

mean

Mean of the distribution.

variance

Variance of the distribution.

tree_flatten()

classmethod tree_unflatten(*aux_data*, *params*)

CHAPTER 5

Discrete Distributions

5.1 Bernoulli

Bernoulli(*probs=None*, *logits=None*, *validate_args=None*)

5.2 BernoulliLogits

class BernoulliLogits(*logits=None*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'logits': <numpyro.distributions.constraints._Real object>}

support = <numpyro.distributions.constraints._Boolean object>

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

probs

mean
Mean of the distribution.

variance
Variance of the distribution.

enumerate_support(*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.3 BernoulliProbs

class BernoulliProbs(*probs*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'probs': <numpyro.distributions.constraints._Interval object>}

support = <numpyro.distributions.constraints._Boolean object>

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

enumerate_support(*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.4 BetaBinomial

class BetaBinomial(*concentration1*, *concentration0*, *total_count=1*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

Compound distribution comprising of a beta-binomial pair. The probability of success (`probs` for the `Binomial` distribution) is unknown and randomly drawn from a `Beta` distribution prior to a certain number of Bernoulli trials given by `total_count`.

Parameters

- **concentration1** (*numpy.ndarray*) – 1st concentration parameter (alpha) for the Beta distribution.
- **concentration0** (*numpy.ndarray*) – 2nd concentration parameter (beta) for the Beta distribution.
- **total_count** (*numpy.ndarray*) – number of Bernoulli trials.

arg_constraints = {'concentration0': <numpyro.distributions.constraints._GreaterThan

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

support

enumerate_support(*expand=True*)

Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.5 Binomial

Binomial(*total_count=1*, *probs=None*, *logits=None*, *validate_args=None*)

5.6 BinomialLogits

class BinomialLogits(*logits*, *total_count=1*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'logits': <numpyro.distributions.constraints._Real object>, 'total

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

probs

mean
Mean of the distribution.

variance
Variance of the distribution.

support

enumerate_support (*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.7 BinomialProbs

class BinomialProbs (*probs*, *total_count=1*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'probs': <numpyro.distributions.constraints._Interval object>, 'to

has_enumerate_support = True

is_discrete = True

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape + batch_shape + event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape + batch_shape + event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

support

enumerate_support (*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.8 Categorical

Categorical(*probs=None*, *logits=None*, *validate_args=None*)

5.9 CategoricalLogits

class CategoricalLogits(*logits*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'logits': <numpyro.distributions.constraints._RealVector object>}

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

probs

mean

Mean of the distribution.

variance

Variance of the distribution.

support

enumerate_support(*expand=True*)

Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.10 CategoricalProbs

class CategoricalProbs(*probs*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'probs': <numpyro.distributions.constraints._Simplex object>}

has_enumerate_support = True

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*.

Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

support

enumerate_support (*expand=True*)
Returns an array with shape *len(support) x batch_shape* containing all values in the support.

5.11 Delta

class Delta (*value=0.0*, *log_density=0.0*, *event_dim=0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'log_density': <numpyro.distributions.constraints._Real object>, '

support = <numpyro.distributions.constraints._Real object>

is_discrete = True

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

tree_flatten ()

classmethod tree_unflatten (*aux_data*, *params*)

5.12 GammaPoisson

class GammaPoisson(*concentration*, *rate=1.0*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

Compound distribution comprising of a gamma-poisson pair, also referred to as a gamma-poisson mixture. The `rate` parameter for the `Poisson` distribution is unknown and randomly drawn from a `Gamma` distribution.

Parameters

- **concentration** (*numpy.ndarray*) – shape parameter (alpha) of the Gamma distribution.
- **rate** (*numpy.ndarray*) – rate parameter (beta) for the Gamma distribution.

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._IntegerGreaterThan object>

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

Mean of the distribution.

variance

Variance of the distribution.

5.13 Geometric

Geometric(*probs=None*, *logits=None*, *validate_args=None*)

5.14 GeometricLogits

class GeometricLogits(*logits*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'logits': <numpyro.distributions.constraints._Real object>}

support = <numpyro.distributions.constraints._IntegerGreaterThan object>

is_discrete = True

probs

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

5.15 GeometricProbs

class GeometricProbs (*probs*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'probs': <numpyro.distributions.constraints._Interval object>}

support = <numpyro.distributions.constraints._IntegerGreaterThan object>

is_discrete = True

sample (*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob (**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

5.16 Multinomial

Multinomial (*total_count=1*, *probs=None*, *logits=None*, *validate_args=None*)

5.17 MultinomialLogits

class MultinomialLogits(*logits*, *total_count=1*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'logits': <numpyro.distributions.constraints._RealVector object>,

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

probs

mean

Mean of the distribution.

variance

Variance of the distribution.

support

5.18 MultinomialProbs

class MultinomialProbs(*probs*, *total_count=1*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'probs': <numpyro.distributions.constraints._Simplex object>, 'tota

is_discrete = True

sample(*key*, *sample_shape=()*)

Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

support

5.19 OrderedLogistic

class OrderedLogistic(*predictor*, *cutpoints*, *validate_args=None*)
Bases: *numpyro.distributions.discrete.CategoricalProbs*

A categorical distribution with ordered outcomes.

References:

1. *Stan Functions Reference, v2.20 section 12.6*, Stan Development Team

Parameters

- **predictor** (*numpy.ndarray*) – prediction in real domain; typically this is output of a linear model.
- **cutpoints** (*numpy.ndarray*) – positions in real domain to separate categories.

arg_constraints = {'cutpoints': <numpyro.distributions.constraints._OrderedVector obj

5.20 Poisson

class Poisson(*rate*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'rate': <numpyro.distributions.constraints._GreaterThan object>}

support = <numpyro.distributions.constraints._IntegerGreaterThan object>

is_discrete = True

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean
Mean of the distribution.

variance
Variance of the distribution.

5.21 PRNGIdentity

class PRNGIdentity
Bases: *numpyro.distributions.distribution.Distribution*

Distribution over `PRNGKey()`. This can be used to draw a batch of `PRNGKey()` using the *seed* handler. Only *sample* method is supported.

is_discrete = True

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

5.22 ZeroInflatedPoisson

class ZeroInflatedPoisson(*gate*, *rate=1.0*, *validate_args=None*)
Bases: *numpyro.distributions.distribution.Distribution*

A Zero Inflated Poisson distribution.

Parameters

- **gate** (*numpy.ndarray*) – probability of extra zeros.
- **rate** (*numpy.ndarray*) – rate of Poisson distribution.

arg_constraints = {'gate': <numpyro.distributions.constraints._Interval object>, 'rat

support = <numpyro.distributions.constraints._IntegerGreaterThan object>

is_discrete = True

sample(*key*, *sample_shape=()*)
Returns a sample from the distribution having shape given by *sample_shape* + *batch_shape* + *event_shape*. Note that when *sample_shape* is non-empty, leading dimensions (of size *sample_shape*) of the returned sample will be filled with iid draws from the distribution instance.

Parameters

- **key** (*jax.random.PRNGKey*) – the rng_key key to be used for the distribution.
- **sample_shape** (*tuple*) – the sample shape for the distribution.

Returns an array of shape *sample_shape* + *batch_shape* + *event_shape*

Return type numpy.ndarray

log_prob(**args*, ***kwargs*)

mean

variance

CHAPTER 6

Directional Distributions

6.1 VonMises

class VonMises(*loc*, *concentration*, *validate_args=None*)

Bases: *numpyro.distributions.distribution.Distribution*

arg_constraints = {'concentration': <numpyro.distributions.constraints._GreaterThan ob

support = <numpyro.distributions.constraints._Interval object>

sample(*key*, *sample_shape=()*)

Generate sample from von Mises distribution

Parameters

- **sample_shape** – shape of samples
- **key** – random number generator key

Returns samples from von Mises

log_prob(**args*, ***kwargs*)

mean

Computes circular mean of distribution. NOTE: same as location when mapped to support [-pi, pi]

variance

Computes circular variance of distribution

CHAPTER 7

Constraints

7.1 Constraint

class Constraint

Bases: object

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

check(*value*)

Returns a byte tensor of *sample_shape* + *batch_shape* indicating whether each event in value satisfies this constraint.

7.2 boolean

boolean = <numpyro.distributions.constraints._Boolean object>

7.3 corr_cholesky

corr_cholesky = <numpyro.distributions.constraints._CorrCholesky object>

7.4 corr_matrix

corr_matrix = <numpyro.distributions.constraints._CorrMatrix object>

7.5 dependent

dependent = <numpyro.distributions.constraints._Dependent object>

7.6 greater_than

greater_than(*lower_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.7 integer_interval

integer_interval(*lower_bound*, *upper_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.8 integer_greater_than

integer_greater_than(*lower_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.9 interval

interval(*lower_bound*, *upper_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.10 less_than

less_than(*upper_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.11 lower_cholesky

lower_cholesky = <numpyro.distributions.constraints._LowerCholesky object>

7.12 multinomial

multinomial (*upper_bound*)

Abstract base class for constraints.

A constraint object represents a region over which a variable is valid, e.g. within which a variable can be optimized.

7.13 nonnegative_integer

nonnegative_integer = <numpyro.distributions.constraints._IntegerGreaterThan object>

7.14 ordered_vector

ordered_vector = <numpyro.distributions.constraints._OrderedVector object>

7.15 positive

positive = <numpyro.distributions.constraints._GreaterThan object>

7.16 positive_definite

positive_definite = <numpyro.distributions.constraints._PositiveDefinite object>

7.17 positive_integer

positive_integer = <numpyro.distributions.constraints._IntegerGreaterThan object>

7.18 real

real = <numpyro.distributions.constraints._Real object>

7.19 real_vector

real_vector = <numpyro.distributions.constraints._RealVector object>

7.20 simplex

```
simplex = <numpyro.distributions.constraints._Simplex object>
```

7.21 unit_interval

```
unit_interval = <numpyro.distributions.constraints._Interval object>
```

CHAPTER 8

Transforms

8.1 biject_to

biject_to(*constraint*)

8.2 Transform

class Transform

Bases: object

domain = <numpyro.distributions.constraints._Real object>

codomain = <numpyro.distributions.constraints._Real object>

event_dim = 0

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

call_with_intermediates(*x*)

8.3 AbsTransform

class AbsTransform

Bases: *numpyro.distributions.transforms.Transform*

domain = <numpyro.distributions.constraints._Real object>

codomain = <numpyro.distributions.constraints._GreaterThan object>

inv(*y*)

8.4 AffineTransform

class AffineTransform(*loc*, *scale*, *domain=<numpyro.distributions.constraints._Real object>*)
Bases: *numpyro.distributions.transforms.Transform*

Note: When *scale* is a JAX tracer, we always assume that *scale > 0* when calculating *codomain*.

codomain

event_dim
int([x]) -> integer int(x, base=10) -> integer

Convert a number or string to an integer, or return 0 if no arguments are given. If x is a number, return x.__int__(). For floating point numbers, this truncates towards zero.

If x is not a number or if base is given, then x must be a string, bytes, or bytearray instance representing an integer literal in the given base. The literal can be preceded by '+' or '-' and be surrounded by whitespace. The base defaults to 10. Valid bases are 0 and 2-36. Base 0 means to interpret the base from the string as an integer literal. >>> int('0b100', base=0) 4

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.5 ComposeTransform

class ComposeTransform(*parts*)
Bases: *numpyro.distributions.transforms.Transform*

domain

codomain

event_dim
int([x]) -> integer int(x, base=10) -> integer

Convert a number or string to an integer, or return 0 if no arguments are given. If x is a number, return x.__int__(). For floating point numbers, this truncates towards zero.

If x is not a number or if base is given, then x must be a string, bytes, or bytearray instance representing an integer literal in the given base. The literal can be preceded by '+' or '-' and be surrounded by whitespace. The base defaults to 10. Valid bases are 0 and 2-36. Base 0 means to interpret the base from the string as an integer literal. >>> int('0b100', base=0) 4

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

call_with_intermediates(*x*)

8.6 CorrCholeskyTransform

class CorrCholeskyTransform
Bases: *numpyro.distributions.transforms.Transform*

Transforms a uncontrained real vector x with length $D * (D-1)/2$ into the Cholesky factor of a D-dimension correlation matrix. This Cholesky factor is a lower triangular matrix with positive diagonals and unit Euclidean norm for each row. The transform is processed as follows:

1. First we convert x into a lower triangular matrix with the following order:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ x_0 & 1 & 0 & 0 \\ x_1 & x_2 & 1 & 0 \\ x_3 & x_4 & x_5 & 1 \end{bmatrix}$$

2. For each row X_i of the lower triangular part, we apply a *signed* version of class *StickBreakingTransform* to transform X_i into a unit Euclidean length vector using the following steps:

   a. Scales into the interval $(-1, 1)$ domain: $r_i = \tanh(X_i)$.

   b. Transforms into an unsigned domain: $z_i = r_i^2$.

   c. Applies $s_i = StickBreakingTransform(z_i)$.

   d. Transforms back into signed domain: $y_i = (sign(r_i), 1) * \sqrt{s_i}$.

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._CorrCholesky object>

event_dim = 2

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.7 ExpTransform

class ExpTransform(*domain=<numpyro.distributions.constraints._Real object>*)

Bases: *numpyro.distributions.transforms.Transform*

codomain

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.8 IdentityTransform

class IdentityTransform(*event_dim=0*)

Bases: *numpyro.distributions.transforms.Transform*

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.9 InvCholeskyTransform

class InvCholeskyTransform(*domain=<numpyro.distributions.constraints._LowerCholesky object>*)

Bases: *numpyro.distributions.transforms.Transform*

Transform via the mapping $y = x@x.T$, where x is a lower triangular matrix with positive diagonal.

event_dim = 2

codomain

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.10 LowerCholeskyAffine

class LowerCholeskyAffine(*loc*, *scale_tril*)

Bases: *numpyro.distributions.transforms.Transform*

Transform via the mapping $y = loc + scale\_tril \ @ \ x$.

Parameters

- **loc** – a real vector.
- **scale_tril** – a lower triangular matrix with positive diagonal.

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._RealVector object>

event_dim = 1

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.11 LowerCholeskyTransform

class LowerCholeskyTransform

Bases: *numpyro.distributions.transforms.Transform*

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._LowerCholesky object>

event_dim = 2

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.12 OrderedTransform

class OrderedTransform

Bases: *numpyro.distributions.transforms.Transform*

Transform a real vector to an ordered vector.

References:

1. *Stan Reference Manual v2.20, section 10.6*, Stan Development Team

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._OrderedVector object>

event_dim = 1

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.13 PermuteTransform

class PermuteTransform(*permutation*)

Bases: *numpyro.distributions.transforms.Transform*

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._RealVector object>

event_dim = 1

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.14 PowerTransform

class PowerTransform(*exponent*)

Bases: *numpyro.distributions.transforms.Transform*

domain = <numpyro.distributions.constraints._GreaterThan object>

codomain = <numpyro.distributions.constraints._GreaterThan object>

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.15 SigmoidTransform

class SigmoidTransform

Bases: *numpyro.distributions.transforms.Transform*

codomain = <numpyro.distributions.constraints._Interval object>

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

8.16 StickBreakingTransform

class StickBreakingTransform

Bases: *numpyro.distributions.transforms.Transform*

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._Simplex object>

event_dim = 1

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

CHAPTER 9

Flows

9.1 InverseAutoregressiveTransform

class InverseAutoregressiveTransform(*autoregressive_nn*, *log_scale_min_clip=-5.0*, *log_scale_max_clip=3.0*)

Bases: *numpyro.distributions.transforms.Transform*

An implementation of Inverse Autoregressive Flow, using Eq (10) from Kingma et al., 2016,

$$\mathbf{y} = \mu_t + \sigma_t \odot \mathbf{x}$$

where $\mathbf{x}$ are the inputs, $\mathbf{y}$ are the outputs, μ_t, σ_t are calculated from an autoregressive network on $\mathbf{x}$, and $\sigma_t > 0$.

References

1. *Improving Variational Inference with Inverse Autoregressive Flow* [arXiv:1606.04934], Diederik P. Kingma, Tim Salimans, Rafal Jozefowicz, Xi Chen, Ilya Sutskever, Max Welling

domain = <numpyro.distributions.constraints._RealVector object>

codomain = <numpyro.distributions.constraints._RealVector object>

event_dim = 1

call_with_intermediates(*x*)

inv(*y*)

Parameters **y** (*numpy.ndarray*) – the output of the transform to be inverted

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

Calculates the elementwise determinant of the log jacobian.

Parameters

- **x** (*numpy.ndarray*) – the input to the transform
- **y** (*numpy.ndarray*) – the output of the transform

9.2 BlockNeuralAutoregressiveTransform

class BlockNeuralAutoregressiveTransform(*bn_arn*)

Bases: *numpyro.distributions.transforms.Transform*

An implementation of Block Neural Autoregressive flow.

References

1. *Block Neural Autoregressive Flow*, Nicola De Cao, Ivan Titov, Wilker Aziz

event_dim = 1

call_with_intermediates(*x*)

inv(*y*)

log_abs_det_jacobian(*x*, *y*, *intermediates=None*)

Calculates the elementwise determinant of the log jacobian.

Parameters

- **x** (*numpy.ndarray*) – the input to the transform
- **y** (*numpy.ndarray*) – the output of the transform

CHAPTER 10

Markov Chain Monte Carlo (MCMC)

class MCMC(*sampler*, *num_warmup*, *num_samples*, *num_chains=1*, *postprocess_fn=None*, *chain_method='parallel'*, *progress_bar=True*, *jit_model_args=False*)

Bases: `object`

Provides access to Markov Chain Monte Carlo inference algorithms in NumPyro.

Note: *chain_method* is an experimental arg, which might be removed in a future version.

Note: Setting *progress_bar=False* will improve the speed for many cases.

Parameters

- **sampler** (`MCMCKernel`) – an instance of *MCMCKernel* that determines the sampler for running MCMC. Currently, only `HMC` and `NUTS` are available.
- **num_warmup** (*int*) – Number of warmup steps.
- **num_samples** (*int*) – Number of samples to generate from the Markov chain.
- **num_chains** (*int*) – Number of Number of MCMC chains to run. By default, chains will be run in parallel using `jax.pmap()`, failing which, chains will be run in sequence.
- **postprocess_fn** – Post-processing callable - used to convert a collection of unconstrained sample values returned from the sampler to constrained values that lie within the support of the sample sites. Additionally, this is used to return values at deterministic sites in the model.
- **chain_method** (*str*) – One of 'parallel' (default), 'sequential', 'vectorized'. The method 'parallel' is used to execute the drawing process in parallel on XLA devices (CPUs/GPUs/TPUs), If there are not enough devices for 'parallel', we fall back to 'sequential' method to draw chains sequentially. 'vectorized' method is an experimental feature which vectorizes the drawing method, hence allowing us to collect samples in parallel on a single device.

- **progress_bar** (*bool*) – Whether to enable progress bar updates. Defaults to `True`.
- **jit_model_args** (*bool*) – If set to *True*, this will compile the potential energy computation as a function of model arguments. As such, calling *MCMC.run* again on a same sized but different dataset will not result in additional compilation cost.

warmup(*rng_key*, **args*, *extra_fields=()*, *collect_warmup=False*, *init_params=None*, ***kwargs*)
Run the MCMC warmup adaptation phase. After this call, the *run()* method will skip the warmup adaptation phase. To run *warmup* again for the new data, it is required to run *warmup()* again.

Parameters

- **rng_key** (*random.PRNGKey*) – Random number generator key to be used for the sampling.
- **args** – Arguments to be provided to the *numpyro.infer.mcmc.MCMCKernel.init()* method. These are typically the arguments needed by the *model*.
- **extra_fields** (*tuple or list*) – Extra fields (aside from `default_fields()`) from the state object (e.g. `numpyro.infer.mcmc.HMCState` for HMC) to collect during the MCMC run.
- **collect_warmup** (*bool*) – Whether to collect samples from the warmup phase. Defaults to *False*.
- **init_params** – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **kwargs** – Keyword arguments to be provided to the *numpyro.infer.mcmc.MCMCKernel.init()* method. These are typically the keyword arguments needed by the *model*.

run(*rng_key*, **args*, *extra_fields=()*, *init_params=None*, ***kwargs*)
Run the MCMC samplers and collect samples.

Parameters

- **rng_key** (*random.PRNGKey*) – Random number generator key to be used for the sampling. For multi-chains, a batch of *num_chains* keys can be supplied. If *rng_key* does not have batch_size, it will be split in to a batch of *num_chains* keys.
- **args** – Arguments to be provided to the *numpyro.infer.mcmc.MCMCKernel.init()* method. These are typically the arguments needed by the *model*.
- **extra_fields** (*tuple or list*) – Extra fields (aside from *z*, *diverging*) from `numpyro.infer.mcmc.HMCState` to collect during the MCMC run.
- **init_params** – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **kwargs** – Keyword arguments to be provided to the *numpyro.infer.mcmc.MCMCKernel.init()* method. These are typically the keyword arguments needed by the *model*.

Note: jax allows python code to continue even when the compiled code has not finished yet. This can cause troubles when trying to profile the code for speed. See https://jax.readthedocs.io/en/latest/async_dispatch.html and https://jax.readthedocs.io/en/latest/profiling.html for pointers on profiling jax programs.

---

get_samples(*group_by_chain=False*)
Get samples from the MCMC run.

Parameters **group_by_chain** (*bool*) – Whether to preserve the chain dimension. If True, all samples will have num_chains as the size of their leading dimension.

Returns Samples having the same data type as *init_params*. The data type is a *dict* keyed on site names if a model containing Pyro primitives is used, but can be any `jaxlib.pytree()`, more generally (e.g. when defining a *potential_fn* for HMC that takes *list* args).

get_extra_fields(*group_by_chain=False*)

Get extra fields from the MCMC run.

Parameters **group_by_chain** (*bool*) – Whether to preserve the chain dimension. If True, all samples will have num_chains as the size of their leading dimension.

Returns Extra fields keyed by field names which are specified in the *extra_fields* keyword of *run()*.

print_summary(*prob=0.9*, *exclude_deterministic=True*)

10.1 MCMC Kernels

class MCMCKernel

Bases: abc.ABC

Defines the interface for the Markov transition kernel that is used for MCMC inference.

Example:

```
>>> from collections import namedtuple
>>> from jax import random
>>> import jax.numpy as jnp
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.infer import MCMC

>>> MHState = namedtuple("MHState", ["z", "rng_key"])

>>> class MetropolisHastings(numpyro.infer.mcmc.MCMCKernel):
...     sample_field = "z"
...
...     def __init__(self, potential_fn, step_size=0.1):
...         self.potential_fn = potential_fn
...         self.step_size = step_size
...
...     def init(self, rng_key, num_warmup, init_params, model_args, model_
↪kwargs):
...         return MHState(init_params, rng_key)
...
...     def sample(self, state, model_args, model_kwargs):
...         z, rng_key = state
...         rng_key, key_proposal, key_accept = random.split(rng_key, 3)
...         z_proposal = dist.Normal(z, self.step_size).sample(key_proposal)
...         accept_prob = jnp.exp(self.potential_fn(z) - self.potential_fn(z_
↪proposal))
...         z_new = jnp.where(dist.Uniform().sample(key_accept) < accept_prob, z_
↪proposal, z)
...         return MHState(z_new, rng_key)

>>> def f(x):
```

(continues on next page)

(continued from previous page)

```
...     return ((x - 2) ** 2).sum()

>>> kernel = MetropolisHastings(f)
>>> mcmc = MCMC(kernel, num_warmup=1000, num_samples=1000)
>>> mcmc.run(random.PRNGKey(0), init_params=jnp.array([1., 2.]))
>>> samples = mcmc.get_samples()
```

postprocess_fn(*model_args*, *model_kwargs*)

Get a function that transforms unconstrained values at sample sites to values constrained to the site's support, in addition to returning deterministic sites in the model.

Parameters

- **model_args** – Arguments to the model.
- **model_kwargs** – Keyword arguments to the model.

init(*rng_key*, *num_warmup*, *init_params*, *model_args*, *model_kwargs*)

Initialize the *MCMCKernel* and return an initial state to begin sampling from.

Parameters

- **rng_key** (*random.PRNGKey*) – Random number generator key to initialize the kernel.
- **num_warmup** (*int*) – Number of warmup steps. This can be useful when doing adaptation during warmup.
- **init_params** (*tuple*) – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **model_args** – Arguments provided to the model.
- **model_kwargs** – Keyword arguments provided to the model.

Returns The initial state representing the state of the kernel. This can be any class that is registered as a pytree.

sample(*state*, *model_args*, *model_kwargs*)

Given the current *state*, return the next *state* using the given transition kernel.

Parameters

- **state** – A pytree class representing the state for the kernel. For HMC, this is given by *HMCState*. In general, this could be any class that supports *getattr*.
- **model_args** – Arguments provided to the model.
- **model_kwargs** – Keyword arguments provided to the model.

Returns Next *state*.

sample_field

The attribute of the *state* object passed to *sample()* that denotes the MCMC sample. This is used by *postprocess_fn()* and for reporting results in *MCMC.print_summary()*.

default_fields

The attributes of the *state* object to be collected by default during the MCMC run (when `MCMC.run()` is called).

get_diagnostics_str(*state*)

Given the current *state*, returns the diagnostics string to be added to progress bar for diagnostics purpose.

class HMC(*model=None*, *potential_fn=None*, *kinetic_fn=None*, *step_size=1.0*, *adapt_step_size=True*, *adapt_mass_matrix=True*, *dense_mass=False*, *target_accept_prob=0.8*, *trajectory_length=6.283185307179586*, *init_strategy=<function init_to_uniform>*, *find_heuristic_step_size=False*)

Bases: *numpyro.infer.mcmc.MCMCKernel*

Hamiltonian Monte Carlo inference, using fixed trajectory length, with provision for step size and mass matrix adaptation.

References:

1. *MCMC Using Hamiltonian Dynamics*, Radford M. Neal

Parameters

- **model** – Python callable containing Pyro *primitives*. If model is provided, *potential_fn* will be inferred using the model.
- **potential_fn** – Python callable that computes the potential energy given input parameters. The input parameters to *potential_fn* can be any python collection type, provided that *init_params* argument to *init_kernel* has the same type.
- **kinetic_fn** – Python callable that returns the kinetic energy given inverse mass matrix and momentum. If not provided, the default is euclidean kinetic energy.
- **step_size** (*float*) – Determines the size of a single step taken by the verlet integrator while computing the trajectory using Hamiltonian dynamics. If not specified, it will be set to 1.
- **adapt_step_size** (*bool*) – A flag to decide if we want to adapt step_size during warm-up phase using Dual Averaging scheme.
- **adapt_mass_matrix** (*bool*) – A flag to decide if we want to adapt mass matrix during warm-up phase using Welford scheme.
- **dense_mass** (*bool*) – A flag to decide if mass matrix is dense or diagonal (default when `dense_mass=False`)
- **target_accept_prob** (*float*) – Target acceptance probability for step size adaptation using Dual Averaging. Increasing this value will lead to a smaller step size, hence the sampling will be slower but more robust. Default to 0.8.
- **trajectory_length** (*float*) – Length of a MCMC trajectory for HMC. Default value is 2π.
- **init_strategy** (*callable*) – a per-site initialization function. See *Initialization Strategies* section for available functions.
- **find_heuristic_step_size** (*bool*) – whether to a heuristic function to adjust the step size at the beginning of each adaptation window. Defaults to False.

model

sample_field
The attribute of the *state* object passed to *sample()* that denotes the MCMC sample. This is used by *postprocess_fn()* and for reporting results in *MCMC.print_summary()*.

default_fields
The attributes of the *state* object to be collected by default during the MCMC run (when `MCMC.run()` is called).

get_diagnostics_str(*state*)
Given the current *state*, returns the diagnostics string to be added to progress bar for diagnostics purpose.

init(*rng_key*, *num_warmup*, *init_params=None*, *model_args=()*, *model_kwargs={}*)
Initialize the *MCMCKernel* and return an initial state to begin sampling from.

Parameters

- **rng_key** (*random.PRNGKey*) – Random number generator key to initialize the kernel.
- **num_warmup** (*int*) – Number of warmup steps. This can be useful when doing adaptation during warmup.
- **init_params** (*tuple*) – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **model_args** – Arguments provided to the model.
- **model_kwargs** – Keyword arguments provided to the model.

Returns

The initial state representing the state of the kernel. This can be any class that is registered as a pytree.

postprocess_fn(*args*, *kwargs*)
Get a function that transforms unconstrained values at sample sites to values constrained to the site's support, in addition to returning deterministic sites in the model.

Parameters

- **model_args** – Arguments to the model.
- **model_kwargs** – Keyword arguments to the model.

sample(*state*, *model_args*, *model_kwargs*)
Run HMC from the given *HMCState* and return the resulting *HMCState*.

Parameters

- **state** (*HMCState*) – Represents the current state.
- **model_args** – Arguments provided to the model.
- **model_kwargs** – Keyword arguments provided to the model.

Returns Next *state* after running HMC.

class NUTS(*model=None*, *potential_fn=None*, *kinetic_fn=None*, *step_size=1.0*, *adapt_step_size=True*, *adapt_mass_matrix=True*, *dense_mass=False*, *target_accept_prob=0.8*, *trajectory_length=None*, *max_tree_depth=10*, *init_strategy=<function init_to_uniform>*, *find_heuristic_step_size=False*)
Bases: *numpyro.infer.hmc.HMC*

Hamiltonian Monte Carlo inference, using the No U-Turn Sampler (NUTS) with adaptive path length and mass matrix adaptation.

References:

1. *MCMC Using Hamiltonian Dynamics*, Radford M. Neal
2. *The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo*, Matthew D. Hoffman, and Andrew Gelman.
3. *A Conceptual Introduction to Hamiltonian Monte Carlo'*, Michael Betancourt

Parameters

- **model** – Python callable containing Pyro *primitives*. If model is provided, *potential_fn* will be inferred using the model.

- **potential_fn** – Python callable that computes the potential energy given input parameters. The input parameters to *potential_fn* can be any python collection type, provided that *init_params* argument to *init_kernel* has the same type.
- **kinetic_fn** – Python callable that returns the kinetic energy given inverse mass matrix and momentum. If not provided, the default is euclidean kinetic energy.
- **step_size** (*float*) – Determines the size of a single step taken by the verlet integrator while computing the trajectory using Hamiltonian dynamics. If not specified, it will be set to 1.
- **adapt_step_size** (*bool*) – A flag to decide if we want to adapt step_size during warm-up phase using Dual Averaging scheme.
- **adapt_mass_matrix** (*bool*) – A flag to decide if we want to adapt mass matrix during warm-up phase using Welford scheme.
- **dense_mass** (*bool*) – A flag to decide if mass matrix is dense or diagonal (default when `dense_mass=False`)
- **target_accept_prob** (*float*) – Target acceptance probability for step size adaptation using Dual Averaging. Increasing this value will lead to a smaller step size, hence the sampling will be slower but more robust. Default to 0.8.
- **trajectory_length** (*float*) – Length of a MCMC trajectory for HMC. This arg has no effect in NUTS sampler.
- **max_tree_depth** (*int*) – Max depth of the binary tree created during the doubling scheme of NUTS sampler. Defaults to 10.
- **init_strategy** (*callable*) – a per-site initialization function. See *Initialization Strategies* section for available functions.
- **find_heuristic_step_size** (*bool*) – whether to a heuristic function to adjust the step size at the beginning of each adaptation window. Defaults to False.

class SA(*model=None*, *potential_fn=None*, *adapt_state_size=None*, *dense_mass=True*, *init_strategy=<function init_to_uniform>*)

Bases: *numpyro.infer.mcmc.MCMCKernel*

Sample Adaptive MCMC, a gradient-free sampler.

This is a very fast (in term of n_eff / s) sampler but requires many warmup (burn-in) steps. In each MCMC step, we only need to evaluate potential function at one point.

Note that unlike in reference [1], we return a randomly selected (i.e. thinned) subset of approximate posterior samples of size num_chains x num_samples instead of num_chains x num_samples x adapt_state_size.

Note: We recommend to use this kernel with *progress_bar=False* in `MCMC` to reduce JAX's dispatch overhead.

References:

1. *Sample Adaptive MCMC* (https://papers.nips.cc/paper/9107-sample-adaptive-mcmc), Michael Zhu

Parameters

- **model** – Python callable containing Pyro *primitives*. If model is provided, *potential_fn* will be inferred using the model.

- **potential_fn** – Python callable that computes the potential energy given input parameters. The input parameters to *potential_fn* can be any python collection type, provided that *init_params* argument to *init_kernel* has the same type.
- **adapt_state_size** (*int*) – The number of points to generate proposal distribution. Defaults to 2 times latent size.
- **dense_mass** (*bool*) – A flag to decide if mass matrix is dense or diagonal (default to `dense_mass=True`)
- **init_strategy** (*callable*) – a per-site initialization function. See *Initialization Strategies* section for available functions.

init(*rng_key*, *num_warmup*, *init_params=None*, *model_args=()*, *model_kwargs={}*)
Initialize the *MCMCKernel* and return an initial state to begin sampling from.

Parameters

- **rng_key** (*random.PRNGKey*) – Random number generator key to initialize the kernel.
- **num_warmup** (*int*) – Number of warmup steps. This can be useful when doing adaptation during warmup.
- **init_params** (*tuple*) – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **model_args** – Arguments provided to the model.
- **model_kwargs** – Keyword arguments provided to the model.

Returns

The initial state representing the state of the kernel. This can be any class that is registered as a pytree.

sample_field
The attribute of the *state* object passed to *sample()* that denotes the MCMC sample. This is used by *postprocess_fn()* and for reporting results in *MCMC.print_summary()*.

default_fields
The attributes of the *state* object to be collected by default during the MCMC run (when `MCMC.run()` is called).

get_diagnostics_str(*state*)
Given the current *state*, returns the diagnostics string to be added to progress bar for diagnostics purpose.

postprocess_fn(*args*, *kwargs*)
Get a function that transforms unconstrained values at sample sites to values constrained to the site's support, in addition to returning deterministic sites in the model.

Parameters

- **model_args** – Arguments to the model.
- **model_kwargs** – Keyword arguments to the model.

sample(*state*, *model_args*, *model_kwargs*)
Run SA from the given `SAState` and return the resulting `SAState`.

Parameters

- **state** (*SAState*) – Represents the current state.
- **model_args** – Arguments provided to the model.

- **model_kwargs** – Keyword arguments provided to the model.

Returns Next *state* after running SA.

hmc (*potential_fn=None*, *potential_fn_gen=None*, *kinetic_fn=None*, *algo='NUTS'*)

Hamiltonian Monte Carlo inference, using either fixed number of steps or the No U-Turn Sampler (NUTS) with adaptive path length.

References:

1. *MCMC Using Hamiltonian Dynamics*, Radford M. Neal
2. *The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo*, Matthew D. Hoffman, and Andrew Gelman.
3. *A Conceptual Introduction to Hamiltonian Monte Carlo'*, Michael Betancourt

Parameters

- **potential_fn** – Python callable that computes the potential energy given input parameters. The input parameters to *potential_fn* can be any python collection type, provided that *init_params* argument to *init_kernel* has the same type.
- **potential_fn_gen** – Python callable that when provided with model arguments / keyword arguments returns *potential_fn*. This may be provided to do inference on the same model with changing data. If the data shape remains the same, we can compile *sample_kernel* once, and use the same for multiple inference runs.
- **kinetic_fn** – Python callable that returns the kinetic energy given inverse mass matrix and momentum. If not provided, the default is euclidean kinetic energy.
- **algo** (*str*) – Whether to run `HMC` with fixed number of steps or `NUTS` with adaptive path length. Default is `NUTS`.

Returns a tuple of callables (*init_kernel*, *sample_kernel*), the first one to initialize the sampler, and the second one to generate samples given an existing one.

Warning: Instead of using this interface directly, we would highly recommend you to use the higher level `numpyro.infer.MCMC` API instead.

Example

```
>>> import jax
>>> from jax import random
>>> import jax.numpy as jnp
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.infer.hmc import hmc
>>> from numpyro.infer.util import initialize_model
>>> from numpyro.util import fori_collect

>>> true_coefs = jnp.array([1., 2., 3.])
>>> data = random.normal(random.PRNGKey(2), (2000, 3))
>>> dim = 3
>>> labels = dist.Bernoulli(logits=(true_coefs * data).sum(-1)).sample(random.
↪PRNGKey(3))
>>>
>>> def model(data, labels):
```

(continues on next page)

(continued from previous page)

```
...     coefs_mean = jnp.zeros(dim)
...     coefs = numpyro.sample('beta', dist.Normal(coefs_mean, jnp.ones(3)))
...     intercept = numpyro.sample('intercept', dist.Normal(0., 10.))
...     return numpyro.sample('y', dist.Bernoulli(logits=(coefs * data +␣
↪intercept).sum(-1)), obs=labels)
>>>
>>> model_info = initialize_model(random.PRNGKey(0), model, model_args=(data,␣
↪labels,))
>>> init_kernel, sample_kernel = hmc(model_info.potential_fn, algo='NUTS')
>>> hmc_state = init_kernel(model_info.param_info,
...                         trajectory_length=10,
...                         num_warmup=300)
>>> samples = fori_collect(0, 500, sample_kernel, hmc_state,
...                        transform=lambda state: model_info.postprocess_
↪fn(state.z))
>>> print(jnp.mean(samples['beta'], axis=0))
[0.9153987 2.0754058 2.9621222]
```

init_kernel(*init_params, num_warmup, step_size=1.0, inverse_mass_matrix=None, adapt_step_size=True, adapt_mass_matrix=True, dense_mass=False, target_accept_prob=0.8, trajectory_length=6.283185307179586, max_tree_depth=10, find_heuristic_step_size=False, model_args=(), model_kwargs=None, rng_key=DeviceArray([0, 0], dtype=uint32)*)

Initializes the HMC sampler.

Parameters

- **init_params** – Initial parameters to begin sampling. The type must be consistent with the input type to *potential_fn*.
- **num_warmup** (*int*) – Number of warmup steps; samples generated during warmup are discarded.
- **step_size** (*float*) – Determines the size of a single step taken by the verlet integrator while computing the trajectory using Hamiltonian dynamics. If not specified, it will be set to 1.
- **inverse_mass_matrix** (*numpy.ndarray*) – Initial value for inverse mass matrix. This may be adapted during warmup if adapt_mass_matrix = True. If no value is specified, then it is initialized to the identity matrix.
- **adapt_step_size** (*bool*) – A flag to decide if we want to adapt step_size during warm-up phase using Dual Averaging scheme.
- **adapt_mass_matrix** (*bool*) – A flag to decide if we want to adapt mass matrix during warm-up phase using Welford scheme.
- **dense_mass** (*bool*) – A flag to decide if mass matrix is dense or diagonal (default when `dense_mass=False`)
- **target_accept_prob** (*float*) – Target acceptance probability for step size adaptation using Dual Averaging. Increasing this value will lead to a smaller step size, hence the sampling will be slower but more robust. Default to 0.8.
- **trajectory_length** (*float*) – Length of a MCMC trajectory for HMC. Default value is 2π.
- **max_tree_depth** (*int*) – Max depth of the binary tree created during the doubling scheme of NUTS sampler. Defaults to 10.

- **find_heuristic_step_size** (*bool*) – whether to a heuristic function to adjust the step size at the beginning of each adaptation window. Defaults to False.
- **model_args** (*tuple*) – Model arguments if *potential_fn_gen* is specified.
- **model_kwargs** (*dict*) – Model keyword arguments if *potential_fn_gen* is specified.
- **rng_key** (*jax.random.PRNGKey*) – random key to be used as the source of randomness.

sample_kernel(*hmc_state*, *model_args=()*, *model_kwargs=None*)

Given an existing HMCState, run HMC with fixed (possibly adapted) step size and return a new HMCState.

Parameters

- **hmc_state** – Current sample (and associated state).
- **model_args** (*tuple*) – Model arguments if *potential_fn_gen* is specified.
- **model_kwargs** (*dict*) – Model keyword arguments if *potential_fn_gen* is specified.

Returns new proposed HMCState from simulating Hamiltonian dynamics given existing state.

HMCState = <class 'numpyro.infer.hmc.HMCState'>

A namedtuple() consisting of the following fields:

- **i** - iteration. This is reset to 0 after warmup.
- **z** - Python collection representing values (unconstrained samples from the posterior) at latent sites.
- **z_grad** - Gradient of potential energy w.r.t. latent sample sites.
- **potential_energy** - Potential energy computed at the given value of z.
- **energy** - Sum of potential energy and kinetic energy of the current state.
- **num_steps** - Number of steps in the Hamiltonian trajectory (for diagnostics).
- **accept_prob** - Acceptance probability of the proposal. Note that z does not correspond to the proposal if it is rejected.
- **mean_accept_prob** - Mean acceptance probability until current iteration during warmup adaptation or sampling (for diagnostics).
- **diverging** - A boolean value to indicate whether the current trajectory is diverging.
- **adapt_state** - A HMCAdaptState namedtuple which contains adaptation information during warmup:
  - **step_size** - Step size to be used by the integrator in the next iteration.
  - **inverse_mass_matrix** - The inverse mass matrix to be used for the next iteration.
  - **mass_matrix_sqrt** - The square root of mass matrix to be used for the next iteration. In case of dense mass, this is the Cholesky factorization of the mass matrix.
- **rng_key** - random number generator seed used for the iteration.

SAState = <class 'numpyro.infer.sa.SAState'>

A namedtuple() used in Sample Adaptive MCMC. This consists of the following fields:

- **i** - iteration. This is reset to 0 after warmup.
- **z** - Python collection representing values (unconstrained samples from the posterior) at latent sites.
- **potential_energy** - Potential energy computed at the given value of z.
- **accept_prob** - Acceptance probability of the proposal. Note that z does not correspond to the proposal if it is rejected.

- **mean_accept_prob** - Mean acceptance probability until current iteration during warmup or sampling (for diagnostics).
- **diverging** - A boolean value to indicate whether the new sample potential energy is diverging from the current one.
- **adapt_state** - A `SAAdaptState` namedtuple which contains adaptation information:
  - **zs** - Step size to be used by the integrator in the next iteration.
  - **pes** - Potential energies of *zs*.
  - **loc** - Mean of those *zs*.
  - **inv_mass_matrix_sqrt** - If using dense mass matrix, this is Cholesky of the covariance of *zs*. Otherwise, this is standard deviation of those *zs*.
- **rng_key** - random number generator seed used for the iteration.

10.2 MCMC Utilities

initialize_model(*rng_key*, *model*, *init_strategy=<function init_to_uniform>*, *dynamic_args=False*, *model_args=()*, *model_kwargs=None*)

(EXPERIMENTAL INTERFACE) Helper function that calls `get_potential_fn()` and *find_valid_initial_params()* under the hood to return a tuple of (*init_params_info*, *potential_fn*, *postprocess_fn*, *model_trace*).

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random number generator seed to sample from the prior. The returned *init_params* will have the batch shape `rng_key.shape[:-1]`.
- **model** – Python callable containing Pyro primitives.
- **init_strategy** (*callable*) – a per-site initialization function. See *Initialization Strategies* section for available functions.
- **dynamic_args** (*bool*) – if *True*, the *potential_fn* and *constraints_fn* are themselves dependent on model arguments. When provided a **model_args, **model_kwargs*, they return *potential_fn* and *constraints_fn* callables, respectively.
- **model_args** (*tuple*) – args provided to the model.
- **model_kwargs** (*dict*) – kwargs provided to the model.

Returns a namedtupe *ModelInfo* which contains the fields (*param_info*, *potential_fn*, *postprocess_fn*, *model_trace*), where *param_info* is a namedtuple *ParamInfo* containing values from the prior used to initiate MCMC, their corresponding potential energy, and their gradients; *postprocess_fn* is a callable that uses inverse transforms to convert unconstrained HMC samples to constrained values that lie within the site's support, in addition to returning values at *deterministic* sites in the model.

fori_collect(*lower*, *upper*, *body_fun*, *init_val*, *transform=<function identity>*, *progbar=True*, *return_last_val=False*, *collection_size=None*, ***progbar_opts*)

This looping construct works like `fori_loop()` but with the additional effect of collecting values from the loop body. In addition, this allows for post-processing of these samples via *transform*, and progress bar updates. Note that, *progbar=False* will be faster, especially when collecting a lot of samples. Refer to example usage in `hmc()`.

Parameters

- **lower** (*int*) – the index to start the collective work. In other words, we will skip collecting the first *lower* values.
- **upper** (*int*) – number of times to run the loop body.
- **body_fun** – a callable that takes a collection of *np.ndarray* and returns a collection with the same shape and *dtype*.
- **init_val** – initial value to pass as argument to *body_fun*. Can be any Python collection type containing *np.ndarray* objects.
- **transform** – a callable to post-process the values returned by *body_fn*.
- **progbar** – whether to post progress bar updates.
- **return_last_val** (*bool*) – If *True*, the last value is also returned. This has the same type as *init_val*.
- **collection_size** (*int*) – Size of the returned collection. If not specified, the size will be `upper - lower`. If the size is larger than `upper - lower`, only the top `upper - lower` entries will be non-zero.
- ****progbar_opts** – optional additional progress bar arguments. A *diagnostics_fn* can be supplied which when passed the current value from *body_fun* returns a string that is used to update the progress bar postfix. Also a *progbar_desc* keyword argument can be supplied which is used to label the progress bar.

Returns collection with the same type as *init_val* with values collected along the leading axis of *np.ndarray* objects.

consensus(*subposteriors*, *num_draws=None*, *diagonal=False*, *rng_key=None*)

Merges subposteriors following consensus Monte Carlo algorithm.

References:

1. *Bayes and big data: The consensus Monte Carlo algorithm*, Steven L. Scott, Alexander W. Blocker, Fernando V. Bonassi, Hugh A. Chipman, Edward I. George, Robert E. McCulloch

Parameters

- **subposteriors** (*list*) – a list in which each element is a collection of samples.
- **num_draws** (*int*) – number of draws from the merged posterior.
- **diagonal** (*bool*) – whether to compute weights using variance or covariance, defaults to *False* (using covariance).
- **rng_key** (*jax.random.PRNGKey*) – source of the randomness, defaults to *jax.random.PRNGKey(0)*.

Returns if *num_draws* is None, merges subposteriors without resampling; otherwise, returns a collection of *num_draws* samples with the same data structure as each subposterior.

parametric(*subposteriors*, *diagonal=False*)

Merges subposteriors following (embarrassingly parallel) parametric Monte Carlo algorithm.

References:

1. *Asymptotically Exact, Embarrassingly Parallel MCMC*, Willie Neiswanger, Chong Wang, Eric Xing

Parameters

- **subposteriors** (*list*) – a list in which each element is a collection of samples.

- **diagonal** (*bool*) – whether to compute weights using variance or covariance, defaults to *False* (using covariance).

Returns the estimated mean and variance/covariance parameters of the joined posterior

parametric_draws(*subposteriors*, *num_draws*, *diagonal=False*, *rng_key=None*)

Merges subposteriors following (embarrassingly parallel) parametric Monte Carlo algorithm.

References:

1. *Asymptotically Exact, Embarrassingly Parallel MCMC*, Willie Neiswanger, Chong Wang, Eric Xing

Parameters

- **subposteriors** (*list*) – a list in which each element is a collection of samples.
- **num_draws** (*int*) – number of draws from the merged posterior.
- **diagonal** (*bool*) – whether to compute weights using variance or covariance, defaults to *False* (using covariance).
- **rng_key** (*jax.random.PRNGKey*) – source of the randomness, defaults to *jax.random.PRNGKey(0)*.

Returns a collection of *num_draws* samples with the same data structure as each subposterior.

CHAPTER 11

Stochastic Variational Inference (SVI)

class SVI (*model*, *guide*, *optim*, *loss*, ***static_kwargs*)

Bases: object

Stochastic Variational Inference given an ELBO loss objective.

References

1. *SVI Part I: An Introduction to Stochastic Variational Inference in Pyro*, (http://pyro.ai/examples/svi_part_i.html)

Example:

```
>>> from jax import lax, random
>>> import jax.numpy as jnp
>>> import numpyro
>>> import numpyro.distributions as dist
>>> from numpyro.distributions import constraints
>>> from numpyro.infer import SVI, ELBO

>>> def model(data):
...     f = numpyro.sample("latent_fairness", dist.Beta(10, 10))
...     with numpyro.plate("N", data.shape[0]):
...         numpyro.sample("obs", dist.Bernoulli(f), obs=data)

>>> def guide(data):
...     alpha_q = numpyro.param("alpha_q", 15., constraint=constraints.positive)
...     beta_q = numpyro.param("beta_q", 15., constraint=constraints.positive)
...     numpyro.sample("latent_fairness", dist.Beta(alpha_q, beta_q))

>>> data = jnp.concatenate([jnp.ones(6), jnp.zeros(4)])
>>> optimizer = numpyro.optim.Adam(step_size=0.0005)
>>> svi = SVI(model, guide, optimizer, loss=ELBO())
>>> init_state = svi.init(random.PRNGKey(0), data)
>>> state = lax.fori_loop(0, 2000, lambda i, state: svi.update(state, data)[0],␣
↪init_state)
>>> # or to collect losses during the loop
```

(continues on next page)

(continued from previous page)

```
>>> # state, losses = lax.scan(lambda state, i: svi.update(state, data), init_
↪state, jnp.arange(2000))
>>> params = svi.get_params(state)
>>> inferred_mean = params["alpha_q"] / (params["alpha_q"] + params["beta_q"])
```

Parameters

- **model** – Python callable with Pyro primitives for the model.
- **guide** – Python callable with Pyro primitives for the guide (recognition network).
- **optim** – an instance of _NumpyroOptim.
- **loss** – ELBO loss, i.e. negative Evidence Lower Bound, to minimize.
- **static_kwargs** – static arguments for the model / guide, i.e. arguments that remain constant during fitting.

Returns tuple of *(init_fn, update_fn, evaluate)*.

init(*rng_key*, **args*, ***kwargs*)

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random number generator seed.
- **args** – arguments to the model / guide (these can possibly vary during the course of fitting).
- **kwargs** – keyword arguments to the model / guide (these can possibly vary during the course of fitting).

Returns tuple containing initial SVIState, and *get_params*, a callable that transforms unconstrained parameter values from the optimizer to the specified constrained domain

get_params(*svi_state*)

Gets values at *param* sites of the *model* and *guide*.

Parameters **svi_state** – current state of the optimizer.

update(*svi_state*, **args*, ***kwargs*)

Take a single step of SVI (possibly on a batch / minibatch of data), using the optimizer.

Parameters

- **svi_state** – current state of SVI.
- **args** – arguments to the model / guide (these can possibly vary during the course of fitting).
- **kwargs** – keyword arguments to the model / guide (these can possibly vary during the course of fitting).

Returns tuple of *(svi_state, loss)*.

evaluate(*svi_state*, **args*, ***kwargs*)

Take a single step of SVI (possibly on a batch / minibatch of data).

Parameters

- **svi_state** – current state of SVI.
- **args** – arguments to the model / guide (these can possibly vary during the course of fitting).

- **kwargs** – keyword arguments to the model / guide.

Returns evaluate ELBO loss given the current parameter values (held within *svi_state.optim_state*).

11.1 ELBO

class ELBO(*num_particles=1*)
Bases: object

A trace implementation of ELBO-based SVI. The estimator is constructed along the lines of references [1] and [2]. There are no restrictions on the dependency structure of the model or the guide.

This is the most basic implementation of the Evidence Lower Bound, which is the fundamental objective in Variational Inference. This implementation has various limitations (for example it only supports random variables with reparameterized samplers) but can be used as a template to build more sophisticated loss objectives.

For more details, refer to http://pyro.ai/examples/svi_part_i.html.

References:

1. *Automated Variational Inference in Probabilistic Programming*, David Wingate, Theo Weber
2. *Black Box Variational Inference*, Rajesh Ranganath, Sean Gerrish, David M. Blei

Parameters num_particles – The number of particles/samples used to form the ELBO (gradient) estimators.

loss(*rng_key*, *param_map*, *model*, *guide*, **args*, ***kwargs*)
Evaluates the ELBO with an estimator that uses num_particles many samples/particles.

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random number generator seed.
- **param_map** (*dict*) – dictionary of current parameter values keyed by site name.
- **model** – Python callable with NumPyro primitives for the model.
- **guide** – Python callable with NumPyro primitives for the guide.
- **args** – arguments to the model / guide (these can possibly vary during the course of fitting).
- **kwargs** – keyword arguments to the model / guide (these can possibly vary during the course of fitting).

Returns negative of the Evidence Lower Bound (ELBO) to be minimized.

11.2 RenyiELBO

class RenyiELBO(*alpha=0*, *num_particles=2*)
Bases: *numpyro.infer.elbo.ELBO*

An implementation of Renyi's α-divergence variational inference following reference [1]. In order for the objective to be a strict lower bound, we require $\alpha \geq 0$. Note, however, that according to reference [1], depending on the dataset $\alpha < 0$ might give better results. In the special case $\alpha = 0$, the objective function is that of the important weighted autoencoder derived in reference [2].

Note: Setting $\alpha < 1$ gives a better bound than the usual ELBO.

Parameters

- **alpha** (*float*) – The order of α-divergence. Here $\alpha \neq 1$. Default is 0.
- **num_particles** – The number of particles/samples used to form the objective (gradient) estimator. Default is 2.

References:

1. *Renyi Divergence Variational Inference*, Yingzhen Li, Richard E. Turner
2. *Importance Weighted Autoencoders*, Yuri Burda, Roger Grosse, Ruslan Salakhutdinov

loss(*rng_key*, *param_map*, *model*, *guide*, **args*, ***kwargs*)

Evaluates the Renyi ELBO with an estimator that uses num_particles many samples/particles.

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random number generator seed.
- **param_map** (*dict*) – dictionary of current parameter values keyed by site name.
- **model** – Python callable with NumPyro primitives for the model.
- **guide** – Python callable with NumPyro primitives for the guide.
- **args** – arguments to the model / guide (these can possibly vary during the course of fitting).
- **kwargs** – keyword arguments to the model / guide (these can possibly vary during the course of fitting).

Returns negative of the Renyi Evidence Lower Bound (ELBO) to be minimized.

CHAPTER 12

Automatic Guide Generation

12.1 AutoContinuous

class AutoContinuous(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*)

Bases: numpyro.infer.autoguide.AutoGuide

Base class for implementations of continuous-valued Automatic Differentiation Variational Inference [1].

Each derived class implements its own `_get_posterior()` method.

Assumes model structure and latent dimension are fixed, and all latent variables are continuous.

Reference:

1. *Automatic Differentiation Variational Inference*, Alp Kucukelbir, Dustin Tran, Rajesh Ranganath, Andrew Gelman, David M. Blei

Parameters

- **model** (*callable*) – A NumPyro model.
- **prefix** (*str*) – a prefix that will be prefixed to all param internal sites.
- **init_strategy** (*callable*) – A per-site initialization function. See *Initialization Strategies* section for available functions.

get_base_dist()

Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

get_transform(*params*)

Returns the transformation learned by the guide to generate samples from the unconstrained (approximate) posterior.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

Returns the transform of posterior distribution

Return type *Transform*

get_posterior(*params*)
Returns the posterior distribution.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

sample_posterior(*rng_key*, *params*, *sample_shape=()*)
Get samples from the learned posterior.

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random key to be used draw samples.
- **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.
- **sample_shape** (*tuple*) – batch shape of each latent sample, defaults to ().

Returns a dict containing samples drawn the this guide.

Return type dict

median(*params*)
Returns the posterior median value of each latent variable.

Parameters **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.

Returns A dict mapping sample site name to median tensor.

Return type dict

quantiles(*params*, *quantiles*)
Returns posterior quantiles each latent variable. Example:

```
print(guide.quantiles(opt_state, [0.05, 0.5, 0.95]))
```

Parameters

- **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.
- **quantiles** (*list*) – A list of requested quantiles between 0 and 1.

Returns A dict mapping sample site name to a list of quantile values.

Return type dict

12.2 AutoBNAFNormal

class AutoBNAFNormal(*model, prefix='auto', init_strategy=<function init_to_uniform>, num_flows=1, hidden_factors=[8, 8]*)
Bases: *numpyro.infer.autoguide.AutoContinuous*

This implementation of *AutoContinuous* uses a Diagonal Normal distribution transformed via a *BlockNeuralAutoregressiveTransform* to construct a guide over the entire latent space. The guide does not depend on the model's `*args, **kwargs`.

Usage:

```
guide = AutoBNAFNormal(model, num_flows=1, hidden_factors=[50, 50], ...)
svi = SVI(model, guide, ...)
```

References

1. *Block Neural Autoregressive Flow*, Nicola De Cao, Ivan Titov, Wilker Aziz

Parameters

- **model** (*callable*) – a generative model.
- **prefix** (*str*) – a prefix that will be prefixed to all param internal sites.
- **init_strategy** (*callable*) – A per-site initialization function.
- **num_flows** (*int*) – the number of flows to be used, defaults to 3.
- **hidden_factors** (*list*) – Hidden layer i has `hidden_factors[i]` hidden units per input dimension. This corresponds to both a and b in reference [1]. The elements of hidden_factors must be integers.

get_base_dist()
Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

12.3 AutoDiagonalNormal

class AutoDiagonalNormal(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*, *init_scale=0.1*)
Bases: *numpyro.infer.autoguide.AutoContinuous*

This implementation of *AutoContinuous* uses a Normal distribution with a diagonal covariance matrix to construct a guide over the entire latent space. The guide does not depend on the model's `*args, **kwargs`.

Usage:

```
guide = AutoDiagonalNormal(model, ...)
svi = SVI(model, guide, ...)
```

get_base_dist()
Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

get_transform(*params*)
Returns the transformation learned by the guide to generate samples from the unconstrained (approximate) posterior.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

Returns the transform of posterior distribution

Return type *Transform*

get_posterior(*params*)
Returns a diagonal Normal posterior distribution.

median(*params*)
Returns the posterior median value of each latent variable.

Parameters **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.

Returns A dict mapping sample site name to median tensor.

Return type dict

quantiles(*params*, *quantiles*)
Returns posterior quantiles each latent variable. Example:

```
print(guide.quantiles(opt_state, [0.05, 0.5, 0.95]))
```

Parameters

- **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.
- **quantiles** (*list*) – A list of requested quantiles between 0 and 1.

Returns A dict mapping sample site name to a list of quantile values.

Return type dict

12.4 AutoMultivariateNormal

class AutoMultivariateNormal(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*, *init_scale=0.1*)
Bases: *numpyro.infer.autoguide.AutoContinuous*

This implementation of *AutoContinuous* uses a MultivariateNormal distribution to construct a guide over the entire latent space. The guide does not depend on the model's `*args, **kwargs`.

Usage:

```
guide = AutoMultivariateNormal(model, ...)
svi = SVI(model, guide, ...)
```

get_base_dist()
Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

get_transform(*params*)
Returns the transformation learned by the guide to generate samples from the unconstrained (approximate) posterior.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

Returns the transform of posterior distribution

Return type *Transform*

get_posterior(*params*)
Returns a multivariate Normal posterior distribution.

median(*params*)
Returns the posterior median value of each latent variable.

Parameters **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.

Returns A dict mapping sample site name to median tensor.

Return type dict

quantiles(*params*, *quantiles*)

Returns posterior quantiles each latent variable. Example:

```
print(guide.quantiles(opt_state, [0.05, 0.5, 0.95]))
```

Parameters

- **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.
- **quantiles** (*list*) – A list of requested quantiles between 0 and 1.

Returns A dict mapping sample site name to a list of quantile values.

Return type dict

12.5 AutoIAFNormal

class AutoIAFNormal(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*, *num_flows=3*, *hidden_dims=None*, *skip_connections=False*, *nonlinearity=(<function elementwise.<locals>.<lambda>>, <function elementwise.<locals>.<lambda>>)*)

Bases: *numpyro.infer.autoguide.AutoContinuous*

This implementation of *AutoContinuous* uses a Diagonal Normal distribution transformed via a *InverseAutoregressiveTransform* to construct a guide over the entire latent space. The guide does not depend on the model's `*args, **kwargs`.

Usage:

```
guide = AutoIAFNormal(model, hidden_dims=[20], skip_connections=True, ...)
svi = SVI(model, guide, ...)
```

Parameters

- **model** (*callable*) – a generative model.
- **prefix** (*str*) – a prefix that will be prefixed to all param internal sites.
- **init_strategy** (*callable*) – A per-site initialization function.
- **num_flows** (*int*) – the number of flows to be used, defaults to 3.
- **hidden_dims** (*list*) – the dimensionality of the hidden units per layer. Defaults to `[latent_dim, latent_dim]`.
- **skip_connections** (*bool*) – whether to add skip connections from the input to the output of each flow. Defaults to False.
- **nonlinearity** (*callable*) – the nonlinearity to use in the feedforward network. Defaults to `jax.experimental.stax.Elu()`.

get_base_dist()

Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

12.6 AutoLaplaceApproximation

class AutoLaplaceApproximation(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*)

Bases: *numpyro.infer.autoguide.AutoContinuous*

Laplace approximation (quadratic approximation) approximates the posterior $\log p(z|x)$ by a multivariate normal distribution in the unconstrained space. Under the hood, it uses Delta distributions to construct a MAP guide over the entire (unconstrained) latent space. Its covariance is given by the inverse of the hessian of $-\log p(x, z)$ at the MAP point of z.

Usage:

```
guide = AutoLaplaceApproximation(model, ...)
svi = SVI(model, guide, ...)
```

get_base_dist()

Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

get_transform(*params*)

Returns the transformation learned by the guide to generate samples from the unconstrained (approximate) posterior.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

Returns the transform of posterior distribution

Return type *Transform*

get_posterior(*params*)

Returns a multivariate Normal posterior distribution.

sample_posterior(*rng_key*, *params*, *sample_shape=()*)

Get samples from the learned posterior.

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random key to be used draw samples.
- **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.
- **sample_shape** (*tuple*) – batch shape of each latent sample, defaults to ().

Returns a dict containing samples drawn the this guide.

Return type dict

median(*params*)

Returns the posterior median value of each latent variable.

Parameters **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.

Returns A dict mapping sample site name to median tensor.

Return type dict

quantiles(*params*, *quantiles*)

Returns posterior quantiles each latent variable. Example:

```
print(guide.quantiles(opt_state, [0.05, 0.5, 0.95]))
```

Parameters

- **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.
- **quantiles** (*list*) – A list of requested quantiles between 0 and 1.

Returns A dict mapping sample site name to a list of quantile values.

Return type dict

12.7 AutoLowRankMultivariateNormal

class AutoLowRankMultivariateNormal(*model*, *prefix='auto'*, *init_strategy=<function init_to_uniform>*, *init_scale=0.1*, *rank=None*)

Bases: *numpyro.infer.autoguide.AutoContinuous*

This implementation of *AutoContinuous* uses a LowRankMultivariateNormal distribution to construct a guide over the entire latent space. The guide does not depend on the model's `*args, **kwargs`.

Usage:

```
guide = AutoLowRankMultivariateNormal(model, rank=2, ...)
svi = SVI(model, guide, ...)
```

get_base_dist()

Returns the base distribution of the posterior when reparameterized as a *TransformedDistribution*. This should not depend on the model's **args, **kwargs*.

get_transform(*params*)

Returns the transformation learned by the guide to generate samples from the unconstrained (approximate) posterior.

Parameters **params** (*dict*) – Current parameters of model and autoguide. The parameters can be obtained using *get_params()* method from *SVI*.

Returns the transform of posterior distribution

Return type *Transform*

get_posterior(*params*)

Returns a lowrank multivariate Normal posterior distribution.

median(*params*)

Returns the posterior median value of each latent variable.

Parameters **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.

Returns A dict mapping sample site name to median tensor.

Return type dict

quantiles(*params*, *quantiles*)

Returns posterior quantiles each latent variable. Example:

```
print(guide.quantiles(opt_state, [0.05, 0.5, 0.95]))
```

Parameters

- **params** (*dict*) – A dict containing parameter values. The parameters can be obtained using *get_params()* method from *SVI*.
- **quantiles** (*list*) – A list of requested quantiles between 0 and 1.

Returns A dict mapping sample site name to a list of quantile values.

Return type dict

CHAPTER 13

Reparameterizers

The *numpyro.infer.reparam* module contains reparameterization strategies for the *numpyro.handlers.reparam* effect. These are useful for altering geometry of a poorly-conditioned parameter space to make the posterior better shaped. These can be used with a variety of inference algorithms, e.g. `Auto*Normal` guides and MCMC.

class Reparam
Bases: `abc.ABC`

Base class for reparameterizers.

13.1 Loc-Scale Decentering

class LocScaleReparam(*centered=None*, *shape_params=()*)
Bases: *numpyro.infer.reparam.Reparam*

Generic decentering reparameterizer [1] for latent variables parameterized by `loc` and `scale` (and possibly additional `shape_params`).

This reparameterization works only for latent variables, not likelihoods.

References:

1. *Automatic Reparameterisation of Probabilistic Programs*, Maria I. Gorinova, Dave Moore, Matthew D. Hoffman (2019)

Parameters

- **centered** (*float*) – optional centered parameter. If None (default) learn a per-site per-element centering parameter in `[0,1]`. If 0, fully decenter the distribution; if 1, preserve the centered distribution unchanged.
- **shape_params** (*tuple or list*) – list of additional parameter names to copy unchanged from the centered to decentered distribution.

__call__(*name*, *fn*, *obs*)

Parameters

- **name** (*str*) – A sample site name.
- **fn** (*Distribution*) – A distribution.
- **obs** (*numpy.ndarray*) – Observed value or None.

Returns A pair (`new_fn`, `value`).

13.2 Neural Transport

class NeuTraReparam(*guide*, *params*)

Bases: *numpyro.infer.reparam.Reparam*

Neural Transport reparameterizer [1] of multiple latent variables.

This uses a trained `AutoContinuous` guide to alter the geometry of a model, typically for use e.g. in MCMC. Example usage:

```
# Step 1. Train a guide
guide = AutoIAFNormal(model)
svi = SVI(model, guide, ...)
# ...train the guide...

# Step 2. Use trained guide in NeuTra MCMC
neutra = NeuTraReparam(guide)
model = netra.reparam(model)
nuts = NUTS(model)
# ...now use the model in HMC or NUTS...
```

This reparameterization works only for latent variables, not likelihoods. Note that all sites must share a single common *NeuTraReparam* instance, and that the model must have static structure.

[1] Hoffman, M. et al. (2019) "NeuTra-lizing Bad Geometry in Hamiltonian Monte Carlo Using Neural Transport" https://arxiv.org/abs/1903.03704

Parameters

- **guide** (`AutoContinuous`) – A guide.
- **params** – trained parameters of the guide.

reparam(*fn=None*)

__call__(*name*, *fn*, *obs*)

Parameters

- **name** (*str*) – A sample site name.
- **fn** (*Distribution*) – A distribution.
- **obs** (*numpy.ndarray*) – Observed value or None.

Returns A pair (`new_fn`, `value`).

transform_sample(*latent*)

Given latent samples from the warped posterior (with possible batch dimensions), return a *dict* of samples from the latent sites in the model.

Parameters **latent** – sample from the warped posterior (possibly batched).

Returns a *dict* of samples keyed by latent sites in the model.

Return type dict

13.3 Transformed Distributions

class TransformReparam

Bases: *numpyro.infer.reparam.Reparam*

Reparameterizer for `TransformedDistribution` latent variables.

This is useful for transformed distributions with complex, geometry-changing transforms, where the posterior has simple shape in the space of `base_dist`.

This reparameterization works only for latent variables, not likelihoods.

__call__(*name*, *fn*, *obs*)

Parameters

- **name** (*str*) – A sample site name.
- **fn** (*Distribution*) – A distribution.
- **obs** (*numpy.ndarray*) – Observed value or None.

Returns A pair (`new_fn`, `value`).

CHAPTER 14

Funsor-based NumPyro

14.1 Effect handlers

class enum(*fn=None*, *first_available_dim=None*)

Bases: numpyro.contrib.funsor.enum_messenger.BaseEnumMessenger

Enumerates in parallel over discrete sample sites marked `infer={"enumerate": "parallel"}`.

Parameters

- **fn** (*callable*) – Python callable with NumPyro primitives.
- **first_available_dim** (*int*) – The first tensor dimension (counting from the right) that is available for parallel enumeration. This dimension and all dimensions left may be used internally by Pyro. This should be a negative integer or None.

process_message(*msg*)

class infer_config(*fn*, *config_fn*)

Bases: numpyro.primitives.Messenger

Given a callable *fn* that contains Pyro primitive calls and a callable *config_fn* taking a trace site and returning a dictionary, updates the value of the infer kwarg at a sample site to config_fn(site).

Parameters

- **fn** – a stochastic function (callable containing Pyro primitive calls)
- **config_fn** – a callable taking a site and returning an infer dict

process_message(*msg*)

class markov

Markov dependency declaration.

This is a statistical equivalent of a memory management arena.

Parameters

- **fn** (*callable*) – Python callable with NumPyro primitives.

- **history** (*int*) – The number of previous contexts visible from the current context. Defaults to 1. If zero, this is similar to *numpyro.primitives.plate*.
- **keep** (*bool*) – If true, frames are replayable. This is important when branching: if `keep=True`, neighboring branches at the same level can depend on each other; if `keep=False`, neighboring branches are independent (conditioned on their shared ancestors).

class plate(*name*, *size*, *subsample_size=None*, *dim=None*)

Bases: `numpyro.contrib.funsor.enum_messenger.GlobalNamedMessenger`

An alternative implementation of *numpyro.primitives.plate* primitive. Note that only this version is compatible with enumeration.

There is also a context manager *plate_to_enum_plate()* which converts *numpyro.plate* statements to this version.

Parameters

- **name** (*str*) – Name of the plate.
- **size** (*int*) – Size of the plate.
- **subsample_size** (*int*) – Optional argument denoting the size of the mini-batch. This can be used to apply a scaling factor by inference algorithms. e.g. when computing ELBO using a mini-batch.
- **dim** (*int*) – Optional argument to specify which dimension in the tensor is used as the plate dim. If *None* (default), the leftmost available dim is allocated.

process_message(*msg*)

class to_data

A primitive to extract a python object from a `Funsor`.

Parameters

- **x** (*Funsor*) – A funsor object
- **name_to_dim** (*OrderedDict*) – An optional inputs hint which maps dimension names from *x* to dimension positions of the returned value.
- **dim_type** (*int*) – Either 0, 1, or 2. This optional argument indicates a dimension should be treated as 'local', 'global', or 'visible', which can be used to interact with the global `DimStack`.

Returns A non-funsor equivalent to *x*.

class to_funsor

A primitive to convert a Python object to a `Funsor`.

Parameters

- **x** – An object.
- **output** (*funsor.domains.Domain*) – An optional output hint to uniquely convert a data to a Funsor (e.g. when *x* is a string).
- **dim_to_name** (*OrderedDict*) – An optional mapping from negative batch dimensions to name strings.
- **dim_type** (*int*) – Either 0, 1, or 2. This optional argument indicates a dimension should be treated as 'local', 'global', or 'visible', which can be used to interact with the global `DimStack`.

Returns A Funsor equivalent to *x*.

Return type funsor.terms.Funsor

class trace(*fn=None*)

Bases: *numpyro.handlers.trace*

This version of *trace* handler records information necessary to do packing after execution.

Each sample site is annotated with a "dim_to_name" dictionary, which can be passed directly to *to_funsor()*.

postprocess_message(*msg*)

14.2 Inference Utilities

config_enumerate(*fn*, *default='parallel'*)

Configures enumeration for all relevant sites in a NumPyro model.

When configuring for exhaustive enumeration of discrete variables, this configures all sample sites whose distribution satisfies `.has_enumerate_support == True`.

This can be used as either a function:

```
model = config_enumerate(model)
```

or as a decorator:

```
@config_enumerate
def model(*args, **kwargs):
    ...
```

Note: Currently, only `default='parallel'` is supported.

Parameters

- **fn** (*callable*) – Python callable with NumPyro primitives.
- **default** (*str*) – Which enumerate strategy to use, one of "sequential", "parallel", or None. Defaults to "parallel".

log_density(*model*, *model_args*, *model_kwargs*, *params*)

Similar to *numpyro.infer.util.log_density()* but works for models with discrete latent variables. Internally, this uses `funsor` to marginalize discrete latent sites and evaluate the joint log probability.

Parameters

- **model** – Python callable containing NumPyro primitives. Typically, the model has been enumerated by using *enum* handler:

```
def model(*args, **kwargs):
    ...

log_joint = log_density(enum(config_enumerate(model)), args,␣
↪kwargs, params)
```

- **model_args** (*tuple*) – args provided to the model.

- **model_kwargs** (*dict*) – kwargs provided to the model.
- **params** (*dict*) – dictionary of current parameter values keyed by site name.

Returns log of joint density and a corresponding model trace

plate_to_enum_plate()

A context manager to replace *numpyro.plate* statement by a funsor-based *plate*.

This is useful when doing inference for the usual NumPyro programs with *numpyro.plate* statements. For example, to get trace of a *model* whose discrete latent sites are enumerated, we can use:

```
enum_model = numpyro.contrib.funsor.enum(model)
with plate_to_enum_plate():
    model_trace = numpyro.contrib.funsor.trace(enum_model).get_trace(
        *model_args, **model_kwargs)
```

CHAPTER 15

Optimizers

Optimizer classes defined here are light wrappers over the corresponding optimizers sourced from `jax.experimental.optimizers` with an interface that is better suited for working with NumPyro inference algorithms.

15.1 Adam

class Adam(**args*, ***kwargs*)

Wrapper class for the JAX optimizer: `adam()`

get_params(*state: Tuple[int, _OptState]*) → _Params

Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]

Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]

Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.2 Adagrad

class Adagrad(**args*, ***kwargs*)
Wrapper class for the JAX optimizer: `adagrad()`

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.3 ClippedAdam

class ClippedAdam(**args*, *clip_norm=10.0*, ***kwargs*)
Adam optimizer with gradient clipping.

Parameters **clip_norm** (*float*) – All gradient values will be clipped between *[-clip_norm, clip_norm]*.

Reference:

A Method for Stochastic Optimization, Diederik P. Kingma, Jimmy Ba https://arxiv.org/abs/1412.6980

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g*, *state*)
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.4 Momentum

class Momentum(**args*, ***kwargs*)
Wrapper class for the JAX optimizer: `momentum()`

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.5 RMSProp

class RMSProp(**args*, ***kwargs*)
Wrapper class for the JAX optimizer: `rmsprop()`

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.6 RMSPropMomentum

class **RMSPropMomentum**(**args*, ***kwargs*)
Wrapper class for the JAX optimizer: `rmsprop_momentum()`

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.7 SGD

class **SGD**(**args*, ***kwargs*)
Wrapper class for the JAX optimizer: `sgd()`

get_params(*state: Tuple[int, _OptState]*) → _Params
Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]
Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]
Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

15.8 SM3

class SM3(**args*, ***kwargs*)

Wrapper class for the JAX optimizer: `sm3()`

get_params(*state: Tuple[int, _OptState]*) → _Params

Get current parameter values.

Parameters **state** – current optimizer state.

Returns collection with current value for parameters.

init(*params: _Params*) → Tuple[int, _OptState]

Initialize the optimizer with parameters designated to be optimized.

Parameters **params** – a collection of numpy arrays.

Returns initial optimizer state.

update(*g: _Params, state: Tuple[int, _OptState]*) → Tuple[int, _OptState]

Gradient update for the optimizer.

Parameters

- **g** – gradient information for parameters.
- **state** – current optimizer state.

Returns new optimizer state after the update.

CHAPTER 16

Diagnostics

This provides a small set of utilities in NumPyro that are used to diagnose posterior samples.

16.1 Autocorrelation

autocorrelation (*x*, *axis=0*)

Computes the autocorrelation of samples at dimension `axis`.

Parameters

- **x** (*numpy.ndarray*) – the input array.
- **axis** (*int*) – the dimension to calculate autocorrelation.

Returns autocorrelation of `x`.

Return type numpy.ndarray

16.2 Autocovariance

autocovariance (*x*, *axis=0*)

Computes the autocovariance of samples at dimension `axis`.

Parameters

- **x** (*numpy.ndarray*) – the input array.
- **axis** (*int*) – the dimension to calculate autocovariance.

Returns autocovariance of `x`.

Return type numpy.ndarray

16.3 Effective Sample Size

effective_sample_size(*x*)

Computes effective sample size of input `x`, where the first dimension of `x` is chain dimension and the second dimension of `x` is draw dimension.

References:

1. *Introduction to Markov Chain Monte Carlo*, Charles J. Geyer
2. *Stan Reference Manual version 2.18*, Stan Development Team

Parameters **x** (*numpy.ndarray*) – the input array.

Returns effective sample size of `x`.

Return type numpy.ndarray

16.4 Gelman Rubin

gelman_rubin(*x*)

Computes R-hat over chains of samples `x`, where the first dimension of `x` is chain dimension and the second dimension of `x` is draw dimension. It is required that `x.shape[0] >= 2` and `x.shape[1] >= 2`.

Parameters **x** (*numpy.ndarray*) – the input array.

Returns R-hat of `x`.

Return type numpy.ndarray

16.5 Split Gelman Rubin

split_gelman_rubin(*x*)

Computes split R-hat over chains of samples `x`, where the first dimension of `x` is chain dimension and the second dimension of `x` is draw dimension. It is required that `x.shape[1] >= 4`.

Parameters **x** (*numpy.ndarray*) – the input array.

Returns split R-hat of `x`.

Return type numpy.ndarray

16.6 HPDI

hpdi(*x*, *prob=0.9*, *axis=0*)

Computes "highest posterior density interval" (HPDI) which is the narrowest interval with probability mass `prob`.

Parameters

- **x** (*numpy.ndarray*) – the input array.
- **prob** (*float*) – the probability mass of samples within the interval.
- **axis** (*int*) – the dimension to calculate hpdi.

Returns quantiles of `x` at `(1 - prob) / 2` and `(1 + prob) / 2`.

Return type numpy.ndarray

16.7 Summary

summary(*samples*, *prob=0.9*, *group_by_chain=True*)

Returns a summary table displaying diagnostics of `samples` from the posterior. The diagnostics displayed are mean, standard deviation, median, the 90% Credibility Interval *hpdi()*, *effective_sample_size()*, and *split_gelman_rubin()*.

Parameters

- **samples** (*dict or numpy.ndarray*) – a collection of input samples with left most dimension is chain dimension and second to left most dimension is draw dimension.
- **prob** (*float*) – the probability mass of samples within the HPDI interval.
- **group_by_chain** (*bool*) – If True, each variable in *samples* will be treated as having shape *num_chains x num_samples x sample_shape*. Otherwise, the corresponding shape will be *num_samples x sample_shape* (i.e. without chain dimension).

print_summary(*samples*, *prob=0.9*, *group_by_chain=True*)

Prints a summary table displaying diagnostics of `samples` from the posterior. The diagnostics displayed are mean, standard deviation, median, the 90% Credibility Interval *hpdi()*, *effective_sample_size()*, and *split_gelman_rubin()*.

Parameters

- **samples** (*dict or numpy.ndarray*) – a collection of input samples with left most dimension is chain dimension and second to left most dimension is draw dimension.
- **prob** (*float*) – the probability mass of samples within the HPDI interval.
- **group_by_chain** (*bool*) – If True, each variable in *samples* will be treated as having shape *num_chains x num_samples x sample_shape*. Otherwise, the corresponding shape will be *num_samples x sample_shape* (i.e. without chain dimension).

CHAPTER 17

Runtime Utilities

17.1 enable_validation

enable_validation(*is_validate=True*)

Enable or disable validation checks in NumPyro. Validation checks provide useful warnings and errors, e.g. NaN checks, validating distribution arguments and support values, etc. which is useful for debugging.

Note: This utility does not take effect under JAX's JIT compilation or vectorized transformation `jax.vmap()`.

Parameters **is_validate** (*bool*) – whether to enable validation checks.

17.2 validation_enabled

validation_enabled(*is_validate=True*)

Context manager that is useful when temporarily enabling/disabling validation checks.

Parameters **is_validate** (*bool*) – whether to enable validation checks.

17.3 enable_x64

enable_x64(*use_x64=True*)

Changes the default array type to use 64 bit precision as in NumPy.

Parameters **use_x64** (*bool*) – when *True*, JAX arrays will use 64 bits by default; else 32 bits.

17.4 set_platform

set_platform(*platform=None*)
Changes platform to CPU, GPU, or TPU. This utility only takes effect at the beginning of your program.

Parameters **platform** (*str*) – either 'cpu', 'gpu', or 'tpu'.

17.5 set_host_device_count

set_host_device_count(*n*)
By default, XLA considers all CPU cores as one device. This utility tells XLA that there are *n* host (CPU) devices available to use. As a consequence, this allows parallel mapping in JAX `jax.pmap()` to work in CPU platform.

Note: This utility only takes effect at the beginning of your program. Under the hood, this sets the environment variable *XLA_FLAGS=–xla_force_host_platform_device_count=[num_devices]*, where *[num_device]* is the desired number of CPU devices *n*.

Warning: Our understanding of the side effects of using the *xla_force_host_platform_device_count* flag in XLA is incomplete. If you observe some strange phenomenon when using this utility, please let us know through our issue or forum page. More information is available in this JAX issue.

Parameters **n** (*int*) – number of CPU devices to use.

CHAPTER 18

Inference Utilities

18.1 Predictive

class Predictive(*model*, *posterior_samples=None*, *guide=None*, *params=None*, *num_samples=None*, *return_sites=None*, *parallel=False*, *batch_ndims=1*)

Bases: object

This class is used to construct predictive distribution. The predictive distribution is obtained by running model conditioned on latent samples from *posterior_samples*.

Warning: The interface for the *Predictive* class is experimental, and might change in the future.

Parameters

- **model** – Python callable containing Pyro primitives.
- **posterior_samples** (*dict*) – dictionary of samples from the posterior.
- **guide** (*callable*) – optional guide to get posterior samples of sites not present in *posterior_samples*.
- **params** (*dict*) – dictionary of values for param sites of model/guide.
- **num_samples** (*int*) – number of samples
- **return_sites** (*list*) – sites to return; by default only sample sites not present in *posterior_samples* are returned.
- **parallel** (*bool*) – whether to predict in parallel using JAX vectorized map jax.vmap(). Defaults to False.
- **batch_ndims** – the number of batch dimensions in posterior samples. Some usages:
  - set *batch_ndims=0* to get prediction for 1 single sample

- set *batch_ndims=1* to get prediction for *posterior_samples* with shapes *(num_samples x ...)*
- set *batch_ndims=2* to get prediction for *posterior_samples* with shapes *(num_chains x N x ...)*. Note that if *num_samples* argument is not None, its value should be equal to *num_chains x N*.

Returns dict of samples from the predictive distribution.

18.2 log_density

log_density(*model*, *model_args*, *model_kwargs*, *params*)

(EXPERIMENTAL INTERFACE) Computes log of joint density for the model given latent values `params`.

Parameters

- **model** – Python callable containing NumPyro primitives.
- **model_args** (*tuple*) – args provided to the model.
- **model_kwargs** (*dict*) – kwargs provided to the model.
- **params** (*dict*) – dictionary of current parameter values keyed by site name.

Returns log of joint density and a corresponding model trace

18.3 transform_fn

transform_fn(*transforms*, *params*, *invert=False*)

(EXPERIMENTAL INTERFACE) Callable that applies a transformation from the *transforms* dict to values in the *params* dict and returns the transformed values keyed on the same names.

Parameters

- **transforms** – Dictionary of transforms keyed by names. Names in *transforms* and *params* should align.
- **params** – Dictionary of arrays keyed by names.
- **invert** – Whether to apply the inverse of the transforms.

Returns *dict* of transformed params.

18.4 constrain_fn

constrain_fn(*model*, *model_args*, *model_kwargs*, *params*, *return_deterministic=False*)

(EXPERIMENTAL INTERFACE) Gets value at each latent site in *model* given unconstrained parameters *params*. The *transforms* is used to transform these unconstrained parameters to base values of the corresponding priors in *model*. If a prior is a transformed distribution, the corresponding base value lies in the support of base distribution. Otherwise, the base value lies in the support of the distribution.

Parameters

- **model** – a callable containing NumPyro primitives.
- **model_args** (*tuple*) – args provided to the model.

- **model_kwargs** (*dict*) – kwargs provided to the model.
- **params** (*dict*) – dictionary of unconstrained values keyed by site names.
- **return_deterministic** (*bool*) – whether to return the value of *deterministic* sites from the model. Defaults to *False*.

Returns *dict* of transformed params.

18.5 potential_energy

potential_energy(*model*, *model_args*, *model_kwargs*, *params*, *enum=False*)

(EXPERIMENTAL INTERFACE) Computes potential energy of a model given unconstrained params. The *inv_transforms* is used to transform these unconstrained parameters to base values of the corresponding priors in *model*. If a prior is a transformed distribution, the corresponding base value lies in the support of base distribution. Otherwise, the base value lies in the support of the distribution.

Parameters

- **model** – a callable containing NumPyro primitives.
- **model_args** (*tuple*) – args provided to the model.
- **model_kwargs** (*dict*) – kwargs provided to the model.
- **params** (*dict*) – unconstrained parameters of *model*.
- **enum** (*bool*) – whether to enumerate over discrete latent sites.

Returns potential energy given unconstrained parameters.

18.6 log_likelihood

log_likelihood(*model*, *posterior_samples*, **args*, *parallel=False*, *batch_ndims=1*, ***kwargs*)

(EXPERIMENTAL INTERFACE) Returns log likelihood at observation nodes of model, given samples of all latent variables.

Parameters

- **model** – Python callable containing Pyro primitives.
- **posterior_samples** (*dict*) – dictionary of samples from the posterior.
- **args** – model arguments.
- **batch_ndims** – the number of batch dimensions in posterior samples. Some usages:
  - set *batch_ndims=0* to get prediction for 1 single sample
  - set *batch_ndims=1* to get prediction for *posterior_samples* with shapes *(num_samples x ...)*
  - set *batch_ndims=2* to get prediction for *posterior_samples* with shapes *(num_chains x N x ...)*
- **kwargs** – model kwargs.

Returns dict of log likelihoods at observation sites.

18.7 find_valid_initial_params

find_valid_initial_params(*rng_key*, *model*, *init_strategy=<function init_to_uniform>*, *enum=False*, *model_args=()*, *model_kwargs=None*, *prototype_params=None*)

(EXPERIMENTAL INTERFACE) Given a model with Pyro primitives, returns an initial valid unconstrained value for all the parameters. This function also returns the corresponding potential energy, the gradients, and an *is_valid* flag to say whether the initial parameters are valid. Parameter values are considered valid if the values and the gradients for the log density have finite values.

Parameters

- **rng_key** (*jax.random.PRNGKey*) – random number generator seed to sample from the prior. The returned *init_params* will have the batch shape `rng_key.shape[:-1]`.
- **model** – Python callable containing Pyro primitives.
- **init_strategy** (*callable*) – a per-site initialization function.
- **enum** (*bool*) – whether to enumerate over discrete latent sites.
- **model_args** (*tuple*) – args provided to the model.
- **model_kwargs** (*dict*) – kwargs provided to the model.
- **prototype_params** (*dict*) – an optional prototype parameters, which is used to define the shape for initial parameters.

Returns tuple of *init_params_info* and *is_valid*, where *init_params_info* is the tuple containing the initial params, their potential energy, and their gradients.

18.8 Initialization Strategies

18.8.1 init_to_feasible

init_to_feasible(*site=None*)

Initialize to an arbitrary feasible point, ignoring distribution parameters.

18.8.2 init_to_median

init_to_median(*site=None*, *num_samples=15*)

Initialize to the prior median. For priors with no *.sample* method implemented, we defer to the *init_to_uniform()* strategy.

Parameters **num_samples** (*int*) – number of prior points to calculate median.

18.8.3 init_to_sample

init_to_sample(*site=None*)

Initialize to a prior sample. For priors with no *.sample* method implemented, we defer to the *init_to_uniform()* strategy.

18.8.4 init_to_uniform

init_to_uniform(*site=None*, *radius=2*)
Initialize to a random point in the area *(-radius, radius)* of unconstrained domain.

Parameters **radius** (*float*) – specifies the range to draw an initial point in the unconstrained domain.

18.8.5 init_to_value

init_to_value(*site=None*, *values={}*)
Initialize to the value specified in *values*. We defer to *init_to_uniform()* strategy for sites which do not appear in *values*.

Parameters **values** (*dict*) – dictionary of initial values keyed by site name.

18.9 Tensor Indexing

vindex(*tensor*, *args*)
Vectorized advanced indexing with broadcasting semantics.

See also the convenience wrapper *Vindex*.

This is useful for writing indexing code that is compatible with batching and enumeration, especially for selecting mixture components with discrete random variables.

For example suppose `x` is a parameter with `len(x.shape) == 3` and we wish to generalize the expression `x[i, :, j]` from integer `i, j` to tensors `i, j` with batch dims and enum dims (but no event dims). Then we can write the generalize version using *Vindex*

```
xij = Vindex(x)[i, :, j]

batch_shape = broadcast_shape(i.shape, j.shape)
event_shape = (x.size(1),)
assert xij.shape == batch_shape + event_shape
```

To handle the case when `x` may also contain batch dimensions (e.g. if `x` was sampled in a plated context as when using vectorized particles), *vindex()* uses the special convention that `Ellipsis` denotes batch dimensions (hence `...` can appear only on the left, never in the middle or in the right). Suppose `x` has event dim 3. Then we can write:

```
old_batch_shape = x.shape[:-3]
old_event_shape = x.shape[-3:]

xij = Vindex(x)[..., i, :, j]   # The ... denotes unknown batch shape.

new_batch_shape = broadcast_shape(old_batch_shape, i.shape, j.shape)
new_event_shape = (x.size(1),)
assert xij.shape = new_batch_shape + new_event_shape
```

Note that this special handling of `Ellipsis` differs from the NEP [1].

Formally, this function assumes:

1. Each arg is either `Ellipsis`, `slice(None)`, an integer, or a batched integer tensor (i.e. with empty event shape). This function does not support Nontrivial slices or boolean tensor masks. `Ellipsis` can only appear on the left as `args[0]`.

2. If `args[0] is not Ellipsis` then `tensor` is not batched, and its event dim is equal to `len(args)`.
3. If `args[0] is Ellipsis` then `tensor` is batched and its event dim is equal to `len(args[1:])`. Dims of `tensor` to the left of the event dims are considered batch dims and will be broadcasted with dims of tensor args.

Note that if none of the args is a tensor with `len(shape) > 0`, then this function behaves like standard indexing:

```
if not any(isinstance(a, jnp.ndarray) and len(a.shape) > 0 for a in args):
    assert Vindex(x)[args] == x[args]
```

References

[1] https://www.numpy.org/neps/nep-0021-advanced-indexing.html introduces `vindex` as a helper for vectorized indexing. This implementation is similar to the proposed notation `x.vindex[]` except for slightly different handling of `Ellipsis`.

Parameters

- **tensor** (*jnp.ndarray*) – A tensor to be indexed.
- **args** (*tuple*) – An index, as args to `__getitem__`.

Returns A nonstandard interpetation of `tensor[args]`.

Return type jnp.ndarray

class Vindex(*tensor*)

Bases: `object`

Convenience wrapper around *vindex()*.

The following are equivalent:

```
Vindex(x)[..., i, j, :]
vindex(x, (Ellipsis, i, j, slice(None)))
```

Parameters **tensor** (*jnp.ndarray*) – A tensor to be indexed.

Returns An object with a special `__getitem__()` method.

CHAPTER 19

Indices and tables

- genindex
- search

Python Module Index

n

Index

Symbols

A

B

C

D

E

F

G

H

I

L

M

N

O

P

Q

R

S

T

U

V

W

Z