Mathematical Foundations and Applications of Distribution Metrics

1 Overview

This document explains the mathematical foundations and practical applications of the functions implemented in `distribution_metrics.py`. The module is designed to compare discrete probability distributions arising in classical machine learning, statistical physics, and quantum computing (e.g., measurement outcomes from quantum circuits).

2 Kernel Methods

2.1 Radial Basis Function (RBF) Kernel

Function: `rbf_kernel(x, y,` γ`)`

Mathematics

The RBF (Gaussian) kernel is defined as:

$$k(x, y) = \exp\left(-\gamma(x - y)^2\right)$$

where:

- x, y are scalar inputs
- $\gamma > 0$ controls smoothness

Interpretation

Measures similarity between two points. Nearby points have kernel value close to 1, distant points decay exponentially.

Applications

- Kernel methods (SVM, kernel PCA)
- Maximum Mean Discrepancy (MMD)
- Distribution comparison in quantum generative models

2.2 Multi-Kernel Matrix

Function: `compute_kernel_matrix(space, gammas)`

Mathematics

For a discrete space $\{x_i\}$:

$$K_{ij} = \frac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} \exp\left(-\gamma (x_i - x_j)^2\right)$$

Interpretation

Averaging multiple bandwidths stabilizes sensitivity across scales.

Applications

- Robust MMD estimation
- Comparing probability distributions over bitstrings

3 Maximum Mean Discrepancy (MMD)

3.1 Kernel Expectation

Function: `kernel_expectation(p, q, K)`

Mathematics

$$\mathbb{E}_{p,q}[k] = p^\top K q$$

Meaning

Computes expected kernel similarity between two distributions.

3.2 MMD Loss

Function: `mmd_loss(p, q, K)`

Mathematics

$$\mathtt{MMD}^2(p, q) = (p - q)^\top K (p - q)$$

Interpretation

Zero if and only if distributions are identical (for characteristic kernels).

Applications

- Training quantum circuit Born machines
- Distribution alignment without likelihoods
- Two-sample testing

4 Information-Theoretic Metrics

4.1 Kullback–Leibler Divergence

Function: `kl_divergence(p, q)`

Mathematics

$$D_{\mathrm{KL}}(p\|q) = \sum_i p_i \log \frac{p_i}{q_i}$$

Properties

- Asymmetric
- Non-negative
- Infinite if $q_i = 0$ and $p_i > 0$

Applications

- Variational inference
- Model fitting
- Quantum state approximation

4.2 Safe Logarithm

Function: `safe_log(x)`

Prevents numerical instability by enforcing:

$$\log(\max(x, \varepsilon))$$

5 Probability Utilities

5.1 Normalization

Function: `normalize_probs`

Ensures:

$$\sum_i p_i = 1$$

Critical for valid probability distributions.

5.2 Bitstring Conversions

Functions:

- `int_to_bitstring`
- `bitstring_to_int`
- `probs_to_bitstrings`

Applications

- Quantum measurement post-processing
- Mapping computational basis states
- Support estimation

5.3 Chi Metric

Function: compute_chi

Definition

$$\chi = \mathbb{E}[\mathbf{1}_{s\in\mathcal{V}}]$$

Measures fraction of samples lying in valid support.

6 Optimal Transport and Sinkhorn Distance

6.1 Cost Matrix

Function: build_cost_matrix

Mathematics

$$C_{ij} = (x_i - x_j)^2$$

Represents squared Euclidean transport cost.

6.2 Sinkhorn Kernel

Function: sinkhorn_kernel

$$K = \exp(-C/\varepsilon)$$

Introduces entropy regularization.

6.3 Sinkhorn Loss

Function: sinkhorn_loss

Optimization Problem

$$\min_{T\in\Pi(p,q)} \langle T, C\rangle - \varepsilon H(T)$$

where:

$$H(T) = -\sum_{ij} T_{ij}\log T_{ij}$$

Algorithm

Uses Sinkhorn-Knopp scaling:

$$T = \texttt{diag}(u)K\texttt{diag}(v)$$

Interpretation

Smooth approximation of Wasserstein distance.

Applications

- Comparing distributions with geometry
- Quantum state transport metrics
- Training generative models with support mismatch

6.4 Sinkhorn Report

Function: sinkhorn_report
Wrapper for diagnostic output.

7 Relevance to Quantum Computing

These metrics are directly applicable to:

- Quantum Circuit Born Machines (QCBM)
- Variational Quantum Algorithms
- Measurement distribution comparison
- Hybrid quantum-classical training loops

8 Conclusion

The module provides a mathematically grounded, numerically stable toolkit for high-fidelity distribution comparison across classical and quantum domains.