

Entropy-Conserving Transformations Using Divergence-Free Vector Fields

Author: Varun Kapoor

Affiliation: Kapoorlabs, Paris, France; Universität Osnabrück, Osnabrück, Germany

Code: <https://github.com/Kapoorlabs-CAPED/entra>

Abstract

We present a method for transforming arbitrary probability distributions towards Gaussian form using divergence-free vector fields constructed from radial basis functions. Our approach exploits the maximum entropy property of Gaussian distributions: among all distributions with a given covariance matrix, the Gaussian has maximum entropy. By iteratively minimizing the covariance determinant using volume-preserving (divergence-free) transformations, we reshape distributions while conserving their fundamental entropy. The divergence-free basis functions are constructed using a differential operator applied to Gaussian radial basis functions, following the methodology of Lowitzsch [1].

Introduction

The Gaussian distribution holds a special place in probability theory and information theory. A fundamental theorem states that:

Among all probability distributions with a given mean and covariance matrix, the Gaussian distribution has the maximum differential entropy.

This maximum entropy principle has profound implications. Consider a distribution with entropy H_0 and covariance Σ . Since the Gaussian with the same covariance has entropy $H_{\text{Gaussian}}(\Sigma) \geq H_0$, with equality only when the distribution is Gaussian. If we apply a **volume-preserving transformation**, the entropy remains fixed at H_0 , but the covariance changes. By **minimizing the covariance determinant**, we reduce $H_{\text{Gaussian}}(\Sigma)$ until it equals H_0 , at which point the distribution must be Gaussian.

The key insight of our method is that **divergence-free vector fields define volume-preserving transformations**. By constructing transformations as linear combinations of divergence-free basis functions, we guarantee that:

1. The Jacobian determinant of the transformation equals 1 everywhere
2. Total probability volume is conserved
3. Differential entropy is conserved under the transformation

This allows us to iteratively minimize the covariance determinant while keeping entropy fixed, driving the distribution towards Gaussian form.

Theoretical Background

Maximum Entropy Property of Gaussians

For a D-dimensional random variable with mean μ and covariance matrix Σ , the differential entropy is bounded by:

$$H(X) \leq \frac{D}{2}(1 + \ln(2\pi)) + \frac{1}{2}\ln \det(\Sigma)$$

with equality if and only if X is Gaussian. The right-hand side is the entropy of a Gaussian with covariance Σ , denoted $H_{\text{Gaussian}}(\Sigma)$.

Corollary: If we have a distribution with fixed entropy H_0 , and we apply volume-preserving transformations that minimize $\det(\Sigma)$, the Gaussian entropy bound $H_{\text{Gaussian}}(\Sigma)$ decreases. When $H_{\text{Gaussian}}(\Sigma) = H_0$, the inequality becomes an equality, meaning the distribution has become Gaussian.

Volume-Preserving Transformations

A transformation $y = T(x)$ is volume-preserving if its Jacobian determinant satisfies:

$$\det\left(\frac{\partial T}{\partial x}\right) = 1$$

For infinitesimal transformations of the form $y = x + \varepsilon v(x)$, this condition becomes:

$$\nabla \cdot v = 0$$

That is, the velocity field v must be **divergence-free**. This is the incompressibility condition from fluid dynamics.

Entropy Conservation and Volume Preservation: The Complete Derivation

We now derive the fundamental connection between entropy conservation and volume preservation, showing that **volume preservation is both necessary and sufficient for entropy conservation**.

The Change of Variables Formula for Entropy

Consider a random variable X with probability density $p_X(x)$. Under a smooth invertible transformation $Y = T(X)$, the density of Y is given by:

$$p_Y(y) = \frac{p_X(T^{-1}(y))}{|\det(J_T(T^{-1}(y)))|}$$

where $J_T = \frac{\partial T}{\partial x}$ is the Jacobian matrix of the transformation.

Derivation of Entropy Change

The differential entropy of Y is:

$$H(Y) = - \int p_Y(y) \ln p_Y(y) dy$$

Substituting the density transformation formula:

$$H(Y) = - \int \frac{p_X(T^{-1}(y))}{|\det(J_T)|} \ln \left(\frac{p_X(T^{-1}(y))}{|\det(J_T)|} \right) dy$$

Applying the change of variables $y = T(x)$, so $dy = |\det(J_T)| dx$:

$$H(Y) = - \int p_X(x) \ln \left(\frac{p_X(x)}{|\det(J_T(x))|} \right) dx$$

Expanding the logarithm:

$$H(Y) = - \int p_X(x) \ln p_X(x) dx + \int p_X(x) \ln |\det(J_T(x))| dx$$

This gives us the **fundamental entropy transformation law**:

$$H(Y) = H(X) + E_X[\ln |\det(J_T(X))|]$$

The Entropy Conservation Condition

For entropy to be conserved ($H(Y) = H(X)$), we require:

$$E_X[\ln |\det(J_T(X))|] = 0$$

Volume Preservation Implies Entropy Conservation

If the transformation is volume-preserving, meaning $|\det(J_T(x))| = 1$ for all x , then:

$$\ln |\det(J_T(x))| = \ln(1) = 0$$

and therefore:

$$E_X[\ln |\det(J_T(X))|] = 0$$

Thus volume preservation is a **sufficient condition** for entropy conservation.

Divergence-Free Fields Generate Volume-Preserving Flows

Consider a velocity field $v(x)$ and the flow it generates:

$$\frac{dx}{dt} = v(x)$$

Let $\Phi_t(x_0)$ denote the position at time t of a particle starting at x_0 . The Jacobian determinant of this flow map evolves according to **Liouville's equation**:

$$\frac{d}{dt} \det(J_{\Phi_t}) = \det(J_{\Phi_t}) \cdot (\nabla \cdot v)$$

Taking the logarithm:

$$\frac{d}{dt} \ln|\det(J_{\Phi_t})| = \nabla \cdot v$$

The Divergence-Free Condition

If the vector field is divergence-free ($\nabla \cdot v = 0$), then:

$$\frac{d}{dt} \ln|\det(J_{\Phi_t})| = 0$$

This means $\ln|\det(J_{\Phi_t})|$ is constant along the flow. Since at $t = 0$ we have $\Phi_0(x) = x$ (identity map), the Jacobian is $J_{\Phi_0} = I$ and $\det(J_{\Phi_0}) = 1$.

Therefore, for all t :

$$|\det(J_{\Phi_t})| = 1$$

The Complete Chain of Implications

We have established:

$$\nabla \cdot v = 0 \Rightarrow |\det(J)| = 1 \Rightarrow H(Y) = H(X)$$

In words:

1. **Divergence-free vector field** \Rightarrow **Volume-preserving transformation**
2. **Volume-preserving transformation** \Rightarrow **Entropy conservation**

This is why we construct our transformations from divergence-free basis functions: it guarantees entropy conservation by construction, allowing us to minimize the covariance determinant while keeping entropy fixed.

Divergence-Free Basis Functions

Construction via Differential Operator

Following Lowitzsch ¹, we construct divergence-free vector fields by applying the differential operator:

$$\hat{O} = -\nabla^2 + \nabla \nabla^T$$

to scalar radial basis functions. For a Gaussian RBF centered at c_l :

$$\phi_l(x) = \exp\left(-\frac{\|x - c_l\|^2}{2\sigma^2}\right)$$

the operator produces a $D \times D$ matrix-valued function:

$$\Phi_l(x) = \hat{O}\phi_l(x)$$

Each column of this matrix is a divergence-free vector field.

Explicit Form for Gaussian RBFs

For a Gaussian RBF, the components of the tensor basis are:

$$\Phi_{ij}(x) = \left[-\delta_{ij} \nabla^2 + \frac{\partial^2}{\partial x_i \partial x_j} \right] \phi(x)$$

where:

$$\nabla^2 \phi = \frac{1}{\sigma^4} (\|x - c\|^2 - D\sigma^2) \phi$$

$$\frac{\partial^2 \phi}{\partial x_i \partial x_j} = \frac{1}{\sigma^4} [(x_i - c_i)(x_j - c_j) - \delta_{ij} \sigma^2] \phi$$

Algorithm

The algorithm proceeds in two stages:

Stage 1: Initial Optimization

1. Sample J points from the initial distribution on a regular grid
2. Place L centers for the divergence-free basis functions
3. Construct the tensor basis Φ of shape (J, L, D, D)
4. Define transformation: $y' = y + \sum_l \Phi_l(y) c_l$ with coefficients c of shape (L, D)
5. Minimize $\det(\text{Cov}(y'))$ using Levenberg-Marquardt optimization
6. Compute the effective basis: $V_l = \Phi_l \cdot c_l$

Stage 2: Iterative Refinement (Outer Loop)

After Stage 1, the tensor basis (J, L, D, D) with coefficients (L, D) collapses to an effective basis (J, L, D) representing L vector fields.

For each outer iteration:

1. Create transformation: $y' = y + \sum_l c_l V_l$ with L scalar coefficients
2. Minimize $\det(\text{Cov}(y'))$ over the L scalar coefficients
3. Update the basis: $V_l \leftarrow c_l V_l$
4. Transform points: $y \leftarrow y'$
5. Repeat until convergence

Entropy Verification

To verify entropy conservation, we employ multiple entropy estimators:

Parametric Estimators

Uniform distribution (initial):

$$H_{\text{uniform}} = \ln(V)$$

where V is the volume of the support.

Gaussian distribution (after transformation):

$$H_{\text{Gaussian}} = \frac{D}{2} (1 + \ln(2\pi)) + \frac{1}{2} \ln \det(\Sigma)$$

Non-parametric Estimator

The Kozachenko-Leonenko k-NN estimator provides a distribution-free entropy estimate:

$$\hat{H} = \frac{D}{N} \sum_{i=1}^N \ln(2\rho_{k,i}) + \ln(V_D) + \psi(N) - \psi(k)$$

where $\rho_{k,i}$ is the distance to the k-th nearest neighbor of point i, V_D is the volume of the unit ball in D dimensions, and ψ is the digamma function.

For a truly volume-preserving transformation, the k-NN entropy estimate should remain constant before and after transformation.

Results

The algorithm successfully transforms uniform distributions towards Gaussian form:

- The covariance determinant decreases monotonically
- 2D projections of 3D distributions show transformation from rectangular (uniform) to elliptical (Gaussian) shapes
- The k-NN entropy estimate remains approximately constant, confirming volume preservation

Conclusion

We have demonstrated a principled method for transforming arbitrary distributions towards Gaussian form using divergence-free vector fields. The key theoretical insights are:

1. Volume-preserving (divergence-free) transformations conserve entropy
2. For any distribution, the Gaussian with the same covariance has higher or equal entropy
3. Minimizing the covariance determinant while preserving entropy reduces the gap between the distribution's entropy and the Gaussian bound, until they become equal

When the distribution's entropy equals the Gaussian entropy bound for its covariance, the distribution must be Gaussian. The divergence-free basis functions constructed from Gaussian RBFs provide a flexible and computationally tractable framework for implementing these entropy-conserving, covariance-minimizing transformations.

References