Why

[1]

Definition

A *deep conditional family* is family of probabilistic generating pairs which are parameterized by a neural network (see Parameterized Distributions and Neural Networks). A *deep generative family* (or *deep latent variable model*, *DLVM*) is a neural network parameterized generative conditional family.

A *variational autoencoder* (or *VAE*) on *observations* X and *latents* Z is an ordered pair $(\{(p_z^{(\theta)}, p_{x|z}^{(\theta)})\}_{\theta \in \Theta}, \{q_{z|x}^{(\phi)}\}_{\phi \in \Phi})$ whose first coordinate is a deep generative family from Z to X and whose second coordinate is deep conditional family from X to Z.

We call

with *latent distribution* (density) $p_z : Z \to \mathsf{R}$ and *observation distribution* (density) $p_x : X \to \mathsf{R}$ is a ordered pair () discrete (continuous) latent set Z and discrete (continuous) observation set X is a tuple

[1]Future editions will include. Future editions may also change the name of this sheet. It is also likely that there will be added prerequisite sheets on variational inference.

Parameterizing distributions

Definition

where (a) ν is an autoencoder (which need not be regular, see Autoencoders), (b) $q_{z|x} : Z \times X \to \mathsf{R}$ is a conditional distribution (density) called the *recognition distribution* (*recognition density*), (c) $p_z : Z \to \mathsf{R}$ is a distribution (density) called the *latent prior model*, and (d) $p_{x|z} : X \times Z \to \mathsf{R}$ is a conditional distribution (density) called the *generating model*.

In other words, for (a) $q_{z|x}(\cdot, \xi) : Z \to \mathsf{R}$ is a distribution (density) for each $\xi \in X$ and for (d) $p_{x|z}(\cdot, \zeta) : X \to \mathsf{R}$ is a distribution (density) on X for each $\zeta \in Z$.

If the model has discrete latent set and discrete observation set (or continuous latent set and continuous observation set), the *joint distribution* (*joint density*) $p_{zx} : Z \times X \to \mathsf{R}$ is defined by $p_{zx} = p_z p_{x|z}$. The *observation distribution*

A *continuous-continuous variational autoencoder family* (*discrete-discrete*, *discrete-continuous*, *continuous-discrete*) is a tuple

$$(\nu, \{(q^{(\theta)}, p_z^{(\theta)}, p_{x|z}^{(\theta))})\}_{\theta \in \Theta}),$$

where:

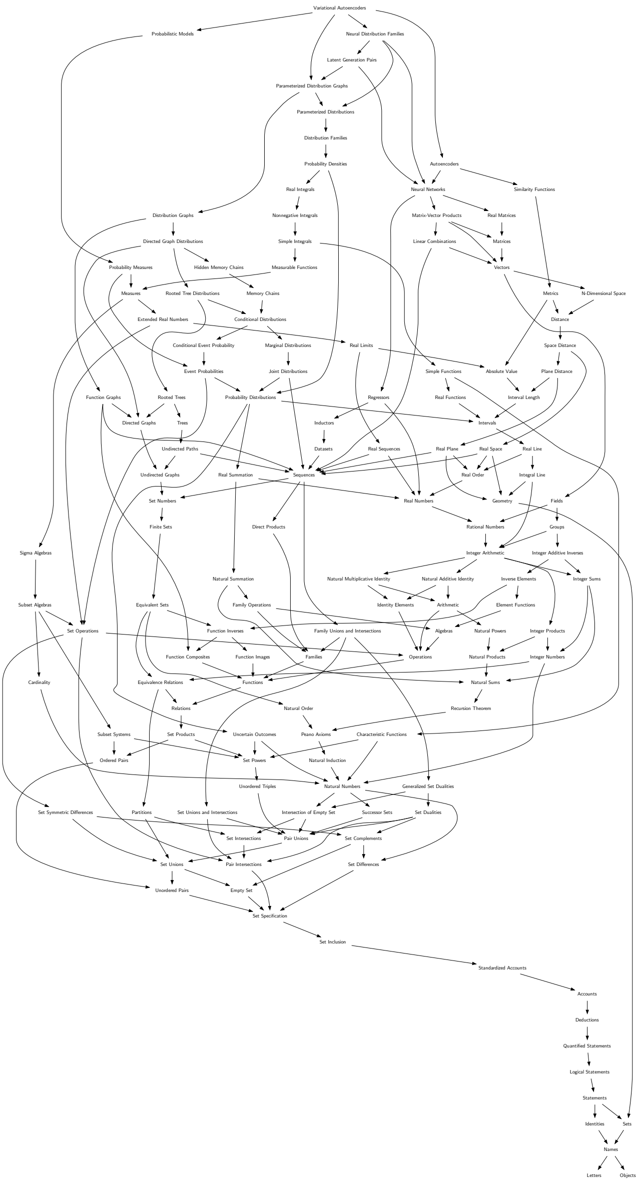
- ν is an autoencoder with encoder $f : \mathsf{R}^d \to \mathsf{R}^k$ and decoder $g : \mathsf{R}^\ell \to \mathsf{R}^m$. The autoencoder need not be regular, see Autoencoders).
- $\Theta \subset \mathsf{R}^p$. The *parameter set* (or *parameter space*).

- $q^{(\theta)} : \mathsf{R}^h \to \mathsf{R}$ is a density (distribution, density, distribution), for each $\theta \in \Theta$. We call $\{q^{(\theta)}\}_{\theta\in\Theta}$ the *recognition model family.*

- $p_z^\theta : \mathsf{R}^h \to \mathsf{R}$ is a density (distribution, distribution, density), for each $\theta \in \Theta$. We call $\{p_z^{(\theta)}\}_\theta$ the *latent prior model family.*

- $p_{x|z}^{(\theta)} : \mathsf{R}^d \times \mathsf{R}^h \to \mathsf{R}$ is a conditional density (distribution, density, distribution). In other words, $p_{x|z}^{(\theta)}(\cdot, \zeta) : \mathsf{R}^d \to \mathsf{R}$ is a density (distribution, density, distribution) for every $\zeta \in \mathsf{R}^d$. We call $\{p_{x|z}^{(\theta)}\}_{\theta\in\Theta}$ the *observation model family.*

A *variational autoencoder* (or *VAE*) may refer to any of the above. The convention we have adopted is "latent type"-"observation type".


Variational Autoencoders
Probabilistic Models
Neural Distribution Families
Latent Generation Pairs
Parameterized Distribution Graphs
Parameterized Distributions
Distribution Families
Probability Densities
Autoencoders
Real Integrals
Neural Networks
Similarity Functions
Distribution Graphs
Nonnegative Integrals
Matrix-Vector Products
Real Matrices
Directed Graph Distributions
Simple Integrals
Linear Combinations
Matrices
Probability Measures
Hidden Memory Chains
Measurable Functions
Vectors
Measures
Rooted Tree Distributions
Memory Chains
Metrics
N-Dimensional Space
Extended Real Numbers
Conditional Distributions
Distance
Conditional Event Probability
Marginal Distributions
Real Limits
Space Distance
Event Probabilities
Joint Distributions
Simple Functions
Absolute Value
Plane Distance
Function Graphs
Rooted Trees
Probability Distributions
Regressors
Real Functions
Interval Length
Directed Graphs
Trees
Inductors
Intervals
Undirected Paths
Datasets
Real Sequences
Real Plane
Real Space
Real Line
Undirected Graphs
Real Summation
Sequences
Real Order
Integral Line
Set Numbers
Real Numbers
Geometry
Fields
Finite Sets
Direct Products
Rational Numbers
Groups
Sigma Algebras
Integer Arithmetic
Integer Additive Inverses
Natural Summation
Natural Multiplicative Identity
Natural Additive Identity
Inverse Elements
Integer Sums
Subset Algebras
Equivalent Sets
Family Operations
Identity Elements
Arithmetic
Element Functions
Set Operations
Function Inverses
Family Unions and Intersections
Algebras
Natural Powers
Integer Products
Function Composites
Function Images
Families
Operations
Natural Products
Integer Numbers
Cardinality
Equivalence Relations
Functions
Natural Sums
Relations
Natural Order
Recursion Theorem
Subset Systems
Set Products
Uncertain Outcomes
Peano Axioms
Characteristic Functions
Ordered Pairs
Set Powers
Natural Induction
Unordered Triples
Natural Numbers
Generalized Set Dualities
Set Symmetric Differences
Partitions
Set Unions and Intersections
Intersection of Empty Set
Successor Sets
Set Dualities
Set Intersections
Pair Unions
Set Complements
Set Unions
Pair Intersections
Set Differences
Unordered Pairs
Empty Set
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Names
Letters
Objects