



Why

1

Definition

A neural network ν commutes with a neural network μ if their associated predictors commute as functions.

An *autoencoder* (or *feedforward autoencoder*) is a pair of neural networks $((\phi_1, \dots, \phi_k), (\psi_1, \dots, \psi_\ell))$. If the networks commute and $\text{dom } \phi_1 = \text{dom } \psi_\ell$, we call the autoencoder *regular*. We call the predictor of the first network the *encoder* and the predictor the second network the *decoder*. We call the image of an input to the encoder an *embedding* (or *feature vector*, *representation*, *code*).

Compressive autoencoders

Let (ϕ, ψ) be regular and let $f : \mathbf{R}^d \rightarrow \mathbf{R}^k$ be the encoder and $g : \mathbf{R}^k \rightarrow \mathbf{R}^d$ be the decoder. If $k < d$, we call the autoencoder *compressive*. Otherwise, we call the autoencoder *noncompressive*. An autoencoder is *perfect* if $g \circ f$ is the identity function. Clearly, a compressive autoencoder can not be perfect.

Let us relax our notion of perfect by introducing a similarity function $\ell : \mathbf{R}^d \times \mathbf{R}^d \rightarrow \mathbf{R}$ (see **Similarity Functions**). An autoencoder is optimal with respect to ℓ if it minimizes

¹Future editions will include. Future editions may also change the name of this sheet.

$\int_{\mathbf{R}^d} \ell(g(f(z)), z) dz$. This integral may diverge. Even if it converges for some autoencoders, there may not be an optimal autoencoder, or a unique one.

If we parameterize a family of autoencoders $\{x_\theta\}_{\theta \in \Theta}$ by a compact set Θ , ... ²

It is natural to be interested in compressive autoencoders.

²Future editions will continue.

