Simple Sparsification Improves Sparse Denoising Autoencoders in Denoising Highly Noisy Images

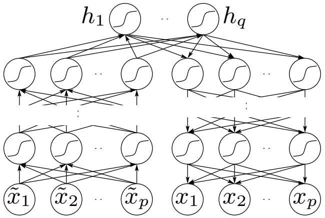
KyungHyun Cho
KyungHyun.Cho@aalto.fi

Aalto University School of Science,
Department of Information and Computer Science

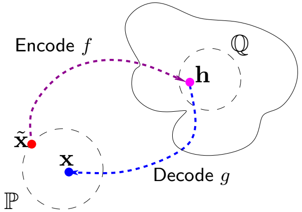
June 11, 2013

Sparse Denoising Autoencoder for *Denoising*

Sparse Denoising Autoencoder



Encoder/Decoder



Minimizes

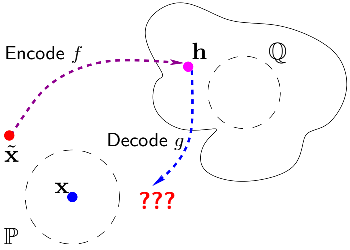
- $\sum_{n=1}^{N} \left\| \mathbf{x}^{(n)} - g\left(f(\kappa(\mathbf{x}^{(n)}))\right) \right\|_2^2 + \lambda\Omega(\mathbf{W}, \{\mathbf{x}^{(n)}\})$
- $\kappa(\mathbf{x})$ *explicitly corrupts* $\mathbf{x}$
- $\Omega(\mathbf{W}, \{\mathbf{x}^{(n)}\})$ *controls sparsity* of $\mathbf{h}$

Maps between

1. Data space $\mathbb{P} = \left\{ \mathbf{x} \in \mathbb{R}^p \,\middle|\, \exists \mathbf{x}^{(n)} \in D, \|\mathbf{x} - \mathbf{x}^{(n)}\|_2^2 \leq \epsilon \right\}$
2. Latent space $\mathbb{Q} \approx \left\{ \mathbf{h} = f(\mathbf{x}) \,\middle|\, \mathbf{x} \in \mathbb{P}, \|\mathbb{E}_{\mathbf{x}\in\mathbb{P}}\left[h_j\right] - \rho\|_2^2 = 0 \right\}$

What if *test* samples were *highly* noisy?

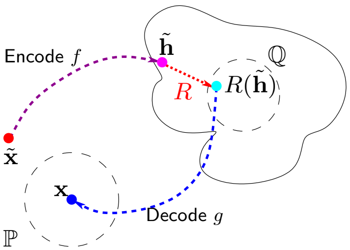
Problems



Out-of-domain Sample $\tilde{\mathbf{x}}$

1. f maps from $\mathbb{P}$, but $\tilde{\mathbf{x}} \notin \mathbb{P}$
2. g maps from $\mathbb{Q}$ to $\mathbb{P}$, but $\mathbf{h} \notin \mathbb{Q}$

Explicit Sparsification R



Sparsification R:

1. R brings $\tilde{\mathbf{h}} \notin \mathbb{Q}$ into $\mathbb{Q}$
2. Simple Sparsification (Shrinkage): $\max\left(\mathbf{h} - \max\left(\|\mathbf{h}\|_1 / q - \rho, 0\right), 0\right)$