Compressed Sensing with Variational Auto-Encoders

Francesco Tonolini, John Bradshaw, Alessandro Davide Ialongo

University of Cambridge

Objectives

In this project we aim to achieve the following:

- Reproduce the results of (Kingma and Welling, 2013).
- Apply the Auto-encoding Variational Bayes (AEVB) method in a compressed sensing setting to reconstruct images from low-dimensional noisy measurements.

Introduction

Variational Auto-encoders (VAEs: Kingma and Welling, 2013) provide an efficient, general method for inference in directed graphical models where we'd like to recover a latent representation z ("encoding") of our data-points x (the "decoded" observations). By training our VAEs on a large dataset, we hope to find a parsimonious encoding mechanism for our observations which we can use to either generate realistic new data or reconstruct an image from a noisy, compressed measurement.
This is the idea behind Compressed Sensing, where, by assuming sparsity in our signal (usually in the output or frequency domain) we can often reconstruct it beyond the limits imposed by the Shannon-Nyquist theorem, from just a few noisy measurements (usually linear projections of the original signal).

Variational Auto-Encoders

The key idea behind VAEs is to optimize a lower bound on the marginal likelihood of the data X:

$$\begin{aligned}\log p_\theta(X) &\geq \mathcal{L}(\theta,\phi;X) \\ &= \sum_i E_{q_\phi(z|x^{(i)})}\left[\log p_\theta(x^{(i)}|z)\right] \\ &\quad - D_{KL}\left[q_\phi(z|x^{(i)})||p_\theta(z)\right]\end{aligned} \quad (1)$$

By parameterizing the "encoding" $q_\phi(z|x^{(i)})$ and "decoding" $p_\theta(x^{(i)}|z)$ distributions via a flexible, differentiable function. In our case we are using a single-layer, non-linear perceptrons with free parameters ϕ, θ.

That is, we let the "decoder" $p_\theta(x^{(i)}|z)$ be a Bernoulli distribution with parameters y given by:

$$y = f_\sigma(W_2 \tanh(W_1 z + b_1) + b_2) \quad (2)$$

Where $\theta = \{W_1, w_2, b_1, b_2\}$ are free (variational) parameters of our model. Similarly, we let the "encoder" be a multivariate Gaussian distribution whose means and variances are given as the output of another perceptron.

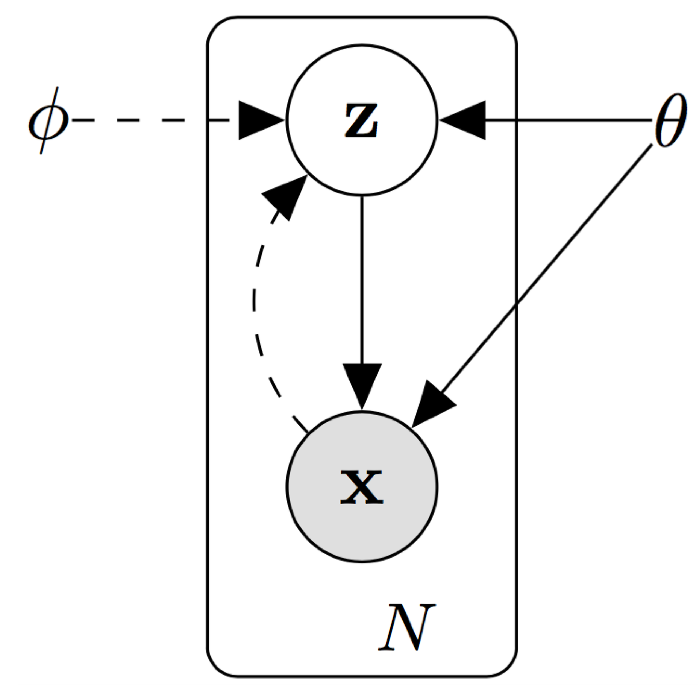

Figure 1: VAE Graphical Model

Compressed Sensing

In Compressed Sensing (CS) the aim is to recover an input x from a number of noisy linear projections (in our case 10% of the original dimensionality, i.e. 78 for MNIST):

$$x_e = \arg\max_x \quad -\frac{1}{2}||Ax - b||^2 + \lambda f(x) \quad (3)$$

Where $f(x)$ is usually a regularization term (e.g. l_1-norm) but in our case it is the VAE's optimized lower bound (pre-fit to relevant data):

$$f(x) = \mathcal{L}(\theta^*, \phi^*; x) \quad (4)$$

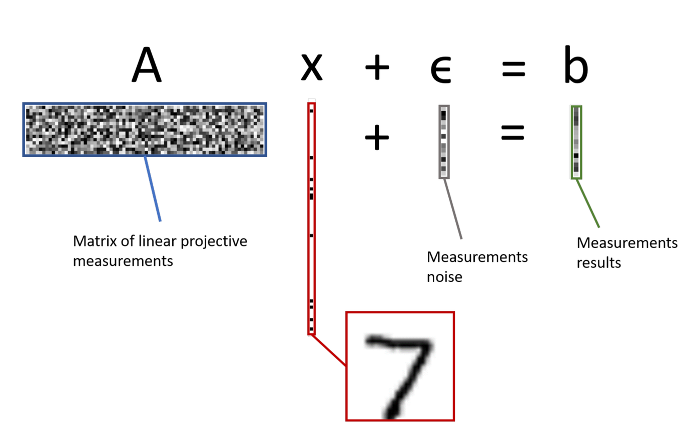

Figure 2: Compressed Sensing Schematic

Methods

All the algorithms and optimization objectives described so far were implemented in Tensorflow and ran on the MNIST digit dataset (LeCun et al, 1998) and a custom generated dataset of Alessandro Ialongo's face (which we thank for its patient cooperation). The perceptron parameters were optimized by backpropagation using Tensorflow's automatic differentiation. They were ran on an NVIDIA GTX 1080 GPU. The results that follow are the outcome of these simulations.

Preliminary Results

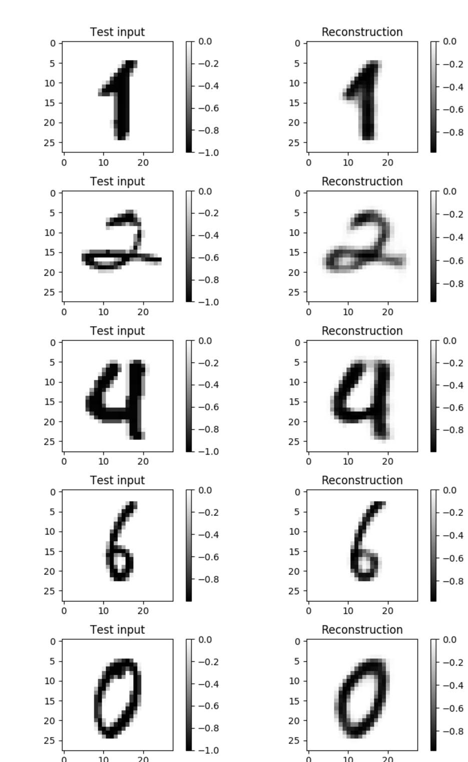

Figure 3: How well the VAE works as a reconstructor on the MNIST dataset

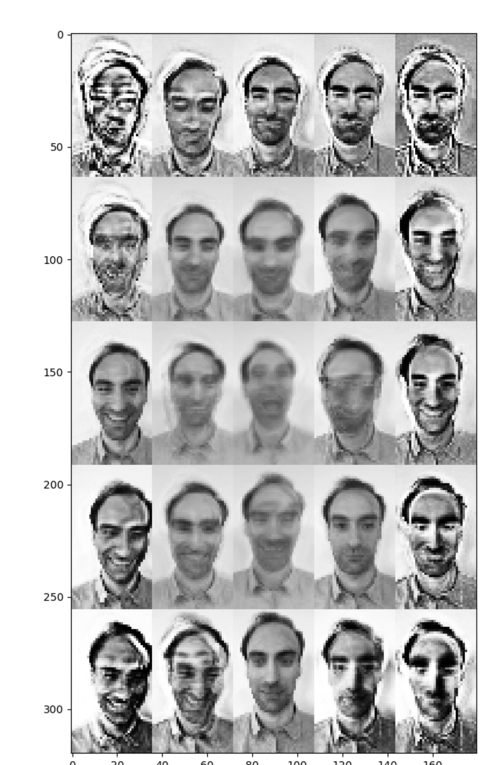

Figure 4: Visualization of 2D latent manifold for "Alessandro's Face" dataset

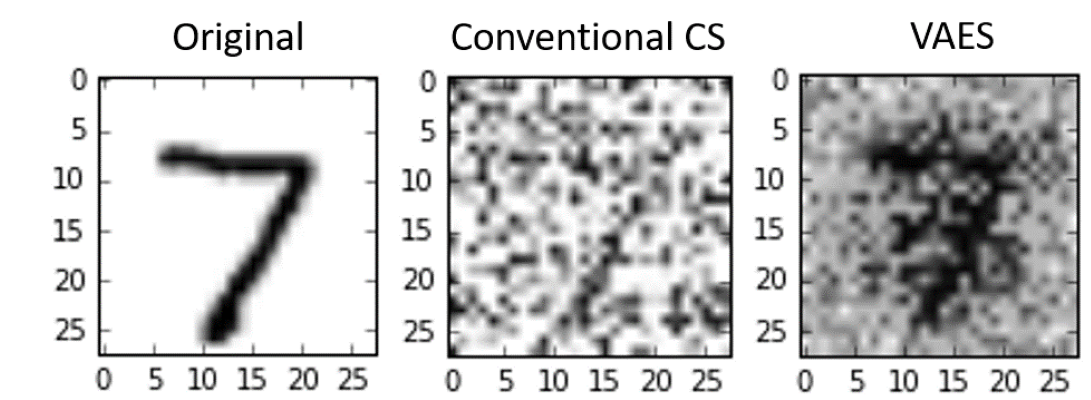

Figure 5: VAE + CS sparse reconstruction

Conclusion

While these are initial explorations, they seem to present encouraging evidence that the VAE is able to capture salient characteristics of hand-written digits and provide a useful additional criterion for the Compressed Sensing reconstruction objective, which can serve as a natural alternative to regularization.

References

- Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." arXiv preprint arXiv:1312.6114 (2013).
- Lustig, Michael, David L. Donoho, Juan M. Santos, and John M. Pauly. "Compressed sensing MRI." IEEE signal processing magazine 25, no. 2 (2008): 72-82.
- LeCun, Yann, Corinna Cortes, and Christopher JC Burges. "The MNIST database of handwritten digits." (1998).
- Jan Hendrik Metzen. Accessed March 10, 2017. https://jmetzen.github.io/.
- Danilo Jimenez Rezende, Shakir Mohamed, and Daan Wierstra. Stochastic backpropagation and variational inference in deep latent gaussian models. arXiv preprint arXiv:1401.4082, 2014.