Optimal Rank-Constrained Mappings for Linear ED Architectures

Alexander DeLise[1], Kyle Loh[2], Krish Patel[3], Meredith Teague[4], Dr. Andrea Arnold[5], Dr. Matthias Chung[3]

[1]Department of Mathematics, Department of Scientific Computing, Florida State University, Tallahassee, FL [2]Department of Mathematics and Statistics, Florida Atlantic University, Boca Raton, FL [3]Department of Mathematics, Emory University, Atlanta, GA, [4]Department of Quantitative Theory and Methods, Department of Computer Science, Emory University, Atlanta, GA, [5]Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA

Motivation

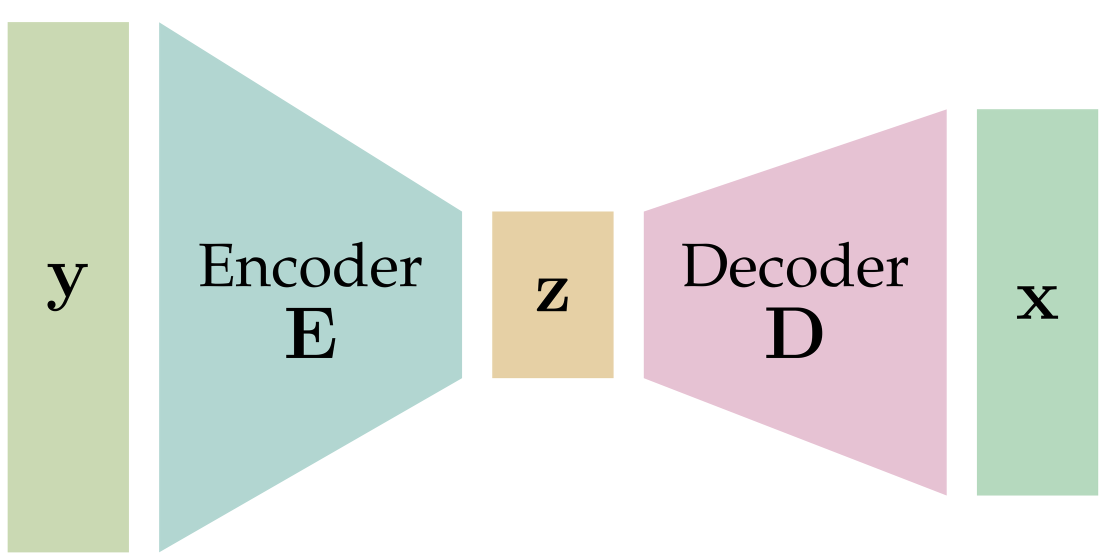
Inverse Problem: Given $\mathbf{y} = \mathbf{F}\mathbf{x} + \varepsilon$, recover the parameters $\mathbf{x}$ given observations $\mathbf{y}$ and forward operator $\mathbf{F}$.

→ Classical inversion methods struggle with ill-posed and high dimensional problems.

→ Data-driven methods are robust to noise, scalable, and adaptable.

Goal: Use encoder-decoder architectures to solve inverse problems via a Bayes risk approach.

Encoder-Decoder Architecture



- Lower dimensional representation of data improves robustness to noise and efficiency
- Bayesian approach incorporates underlying probability distributions and better quantifies uncertainty

Theoretical Results

Background

- X is a random variable with second moment $\mathbf{\Gamma}_X = \mathbf{L}_X\mathbf{L}_X^\top$, covariance $\mathbf{S}_X = \mathbf{K}_X\mathbf{K}_X^\top$, and mean $\boldsymbol{\mu}_X$
- $\mathcal{E}$ is a random noise variable with second moment $\mathbf{\Gamma}_\mathcal{E} = \mathbf{L}_\mathcal{E}\mathbf{L}_\mathcal{E}^\top$
- $\mathbf{F}$ is a known linear forward operator
- Y is a random variable such that $Y = \mathbf{F}X + \mathcal{E}$, with second moment $\mathbf{\Gamma}_Y = \mathbf{L}_Y\mathbf{L}_Y^\top$
- Let $(\mathbf{A})_r = \mathbf{U}_{\mathbf{A},r}\mathbf{\Sigma}_{\mathbf{A},r}\mathbf{V}_{\mathbf{A},r}^\top$ denote the rank-r truncated SVD of a matrix $\mathbf{A}$ [1]

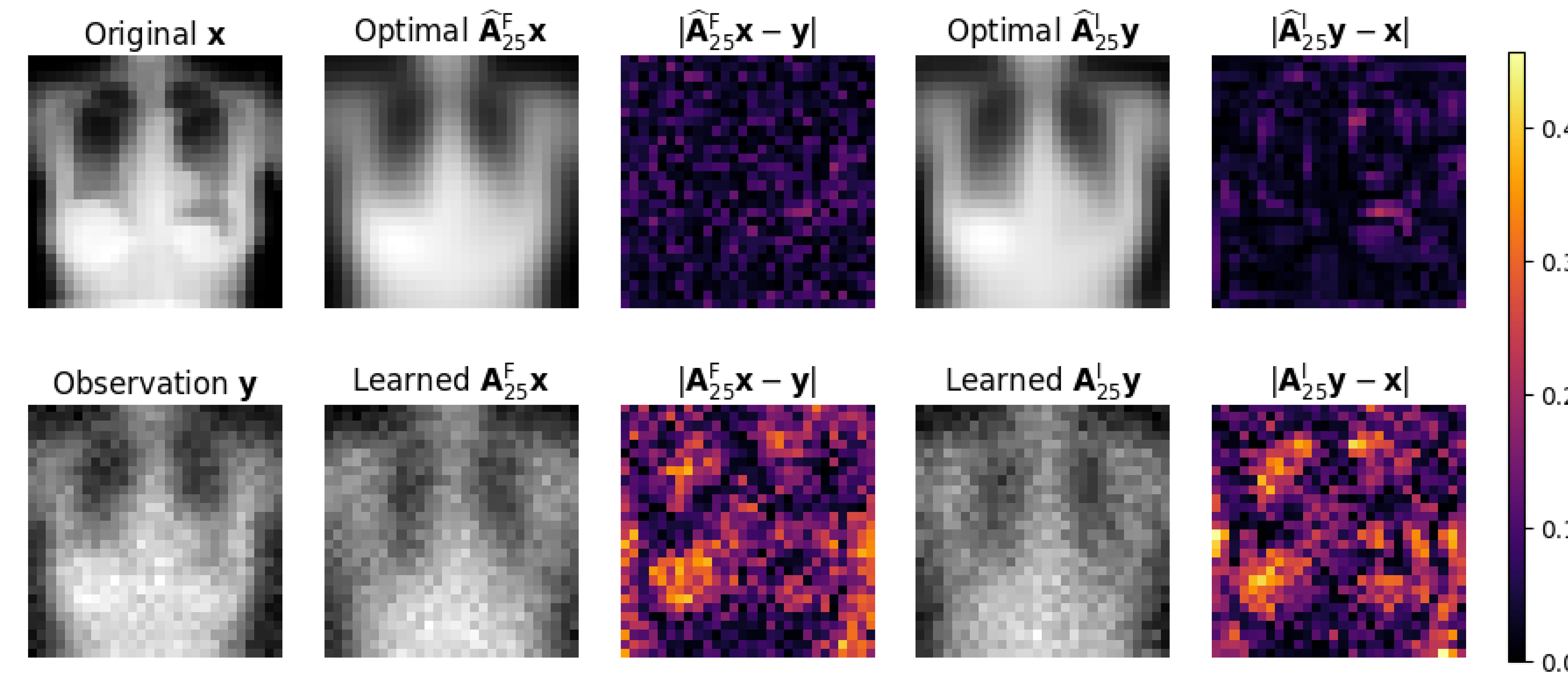
Minimization Problem	Result
$\min_{\mathrm{rank}(\mathbf{A})\le r} \mathbb{E}\,\|\mathbf{A}Y - X\|_2^2$	$\widehat{\mathbf{A}} = (\mathbf{\Gamma}_X\mathbf{F}^\top\mathbf{L}_Y^{-\top})_r\mathbf{L}_Y^{-1}$
$\min_{\mathrm{rank}(\mathbf{A})\le r} \mathbb{E}\,\|\mathbf{A}(X+\mathcal{E}) - X\|_2^2$	$\widehat{\mathbf{A}} = \left(\mathbf{\Gamma}_X\mathbf{L}_Y^{-\top}\right)_r\mathbf{L}_Y^{-1}$
$\min_{\mathrm{rank}(\mathbf{A})\le r} \mathbb{E}\,\|\mathbf{A}X - Y\|_2^2$	$\widehat{\mathbf{A}} = (\mathbf{F}\mathbf{L}_X)_r\mathbf{L}_X^\dagger$
$\min_{\mathrm{rank}(\mathbf{A})\le r} \mathbb{E}\,\|\mathbf{A}X + \mathbf{b} - X\|_2^2$	$\widehat{\mathbf{A}} = \mathbf{U}_{\mathbf{K}_X,r}\mathbf{U}_{\mathbf{K}_X,r}^\top$ and $\widehat{\mathbf{b}} = (\mathbf{I} - \mathbf{A})\mu_X$
$\min_{\mathrm{rank}(\mathbf{A})\le r} \mathbb{E}\,\|\mathbf{A}X - X\|_2^2$	$\widehat{\mathbf{A}} = \mathbf{U}_{\mathbf{L}_X,r}\mathbf{U}_{\mathbf{L}_X,r}^\top$

Acknowledgments

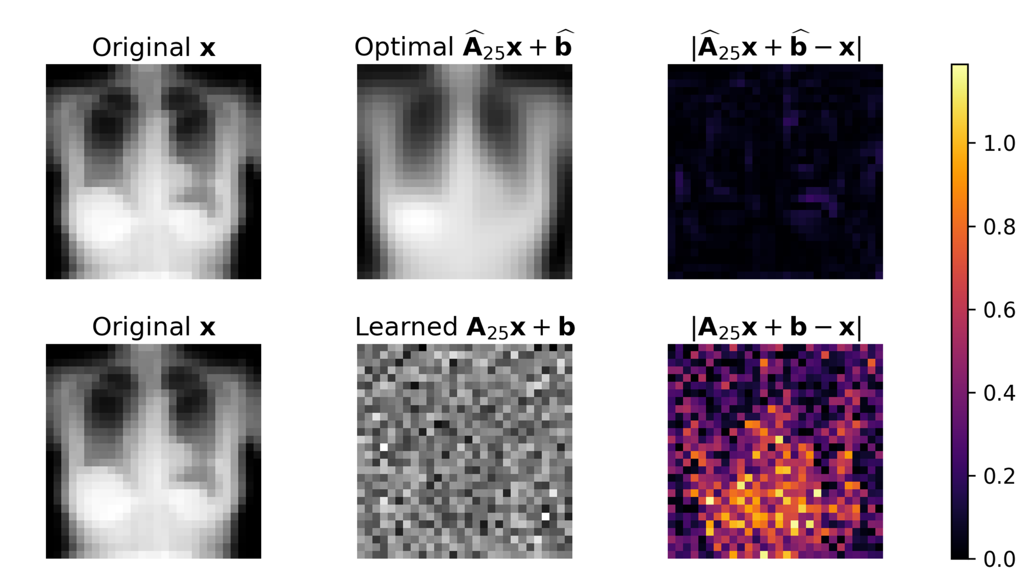
This work is supported by National Science Foundation grant number NSF DMS-2349534.

Numerical Results

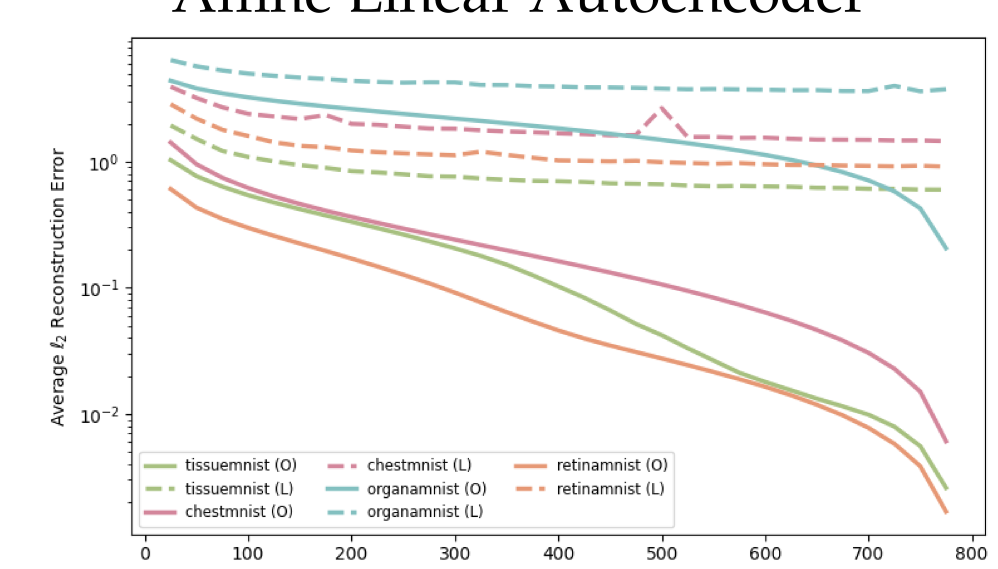
Optimal and Learned Forward and Inverse Mappings applied on ChestMNIST data with Gaussian forward operator



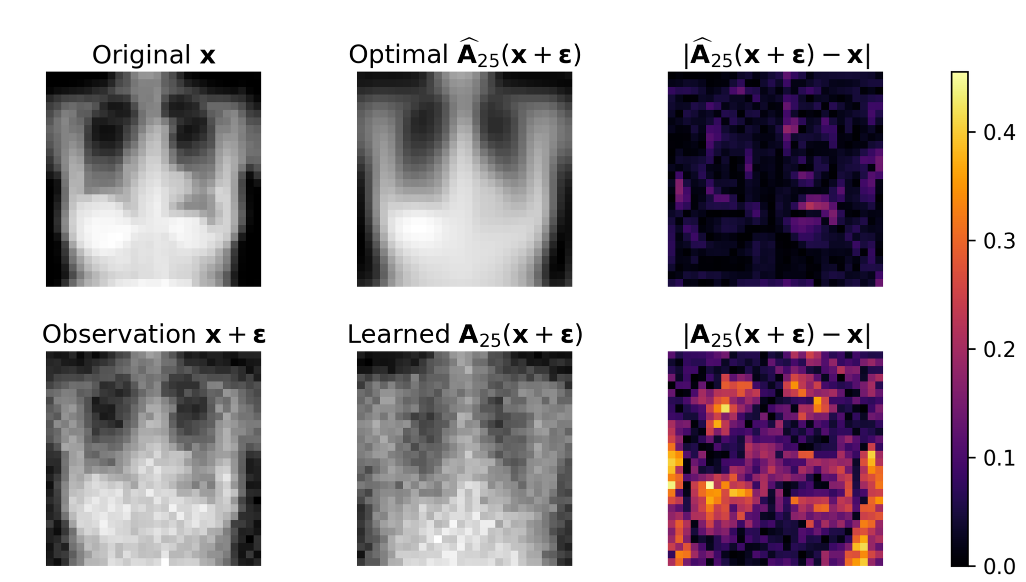
Affine Linear Autoencoder : ChestMNIST data



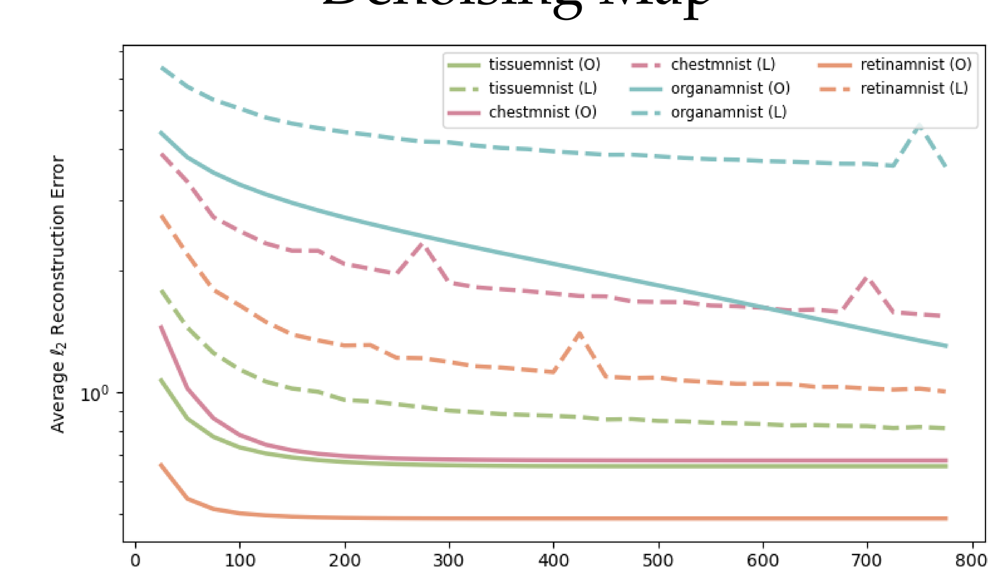
Comparing Learned vs Optimal : Affine Linear Autoencoder



Denoising Map : ChestMNIST data



Comparing Learned vs Optimal : Denoising Map



References

[1] Shmuel Friedland and Anatoli Torokhti. Generalized rank-constrained matrix approximations. *SIAM Journal on Matrix Analysis and Applications*, 29(2):656–659, 2007.

Animations

Inverse End-to-End
Left: ChestMNIST
Right: RetinaMNIST