AM205 Final Project

Owen Schafer, Arthur Young

14 December 2020

1 INTRODUCTION

The field of compressed sensing gained widespread attention from academics following the results of Candès et al. (2006) and the contemporary work by Donoho (2006). Briefly, compressed sensing refers to the practice of determining solutions to undetermined linear systems representing a signal processing problem with limited information. Historically, the task of signal reconstruction was constrained by the Nyquist-Shannon theorem of periodic signal sampling, which essentially asserts that the perfect reconstruction of a continuous, periodic signal of maximum frequency B is possible provided samples are observed at a rate twice that of B. In other words, if one is given a periodic time series with a minimum period of $T = 1/B$, then one needs two samples per period $S = 1/(2B)$, where S is samples per second, in order to exactly reconstruct said signal. The results by Candès et al. and Donoho demonstrated that under certain conditions of incoherence, sparse signals may be exactly reconstructed from far fewer samples than that suggested by the Nyquist-Shannon theorem. A variety of applications immediately arose from their novel results, the famous Netflix problem being one such example. In our study, we are interested in the impacts of the powerful algorithms made possible by these findings on the task of image recognition, where our application of convex optimization methods borne from the period of 2006-2007 are used as a pre-processor for a variety of image recognition techniques. We seek to determine the feasibility of combining compressed sensing techniques with both Bayesian and frequentist methods, particularly in the circumstance of noised data.

2 BASIS PURSUIT PROBLEM

The findings by Candès et al. and Donoho rely on the restatement of the following basis pursuit problem:

$$\min_{s.t. \mathbf{A}\vec{x}=\vec{b}} \{||\vec{x}||_0\} \tag{1}$$

where $\vec{x} \in \mathbb{R}^n$ is an observed signal with authentic value known to be sparse, $\mathbf{A} \in \mathbb{R}^{m \times n}$ is an operator mapping the observation to constraint $\vec{b} \in \mathbb{R}^m$, and $||.||_0$ denotes the informally defined ℓ_0 norm. We can understand the above problem statement in plain english as follows:

"What is the *simplest* value of $\vec{x}$ that still satisfies a requirement of $\mathbf{A}\vec{x} = \vec{b}$?"

We begin by first clarifying the meaning of the ℓ_0 norm: the ℓ_0 norm is an improper Lebesque norm because it is not homogenous, i.e. $||\alpha\vec{x}||_0 \neq \alpha||\vec{x}||_0$ for some $\alpha \in \mathbb{R}$. This is because the ℓ_0 norm $||.||_0$: $\mathbb{R}^n \to \mathbb{R}$ is mathematically defined as follows:

$$||\vec{x}||_0 = \sum_{i=1}^{n} \ell(x_i),\ \ell(\vec{x}_i) := \begin{cases} 1, & \text{if} \quad \vec{x}_i \neq 0 \\ 0, & \text{o.w.} \end{cases} \tag{2}$$

In other words, the ℓ_0 norm is equal to the number of the nonzero elements of $\vec{x}$. Therefore, equation 1 seeks to maximise the sparsity of vector $\vec{x}$, or maximise the "simplicity" of the solution to $\mathbf{A}\vec{x} = \vec{b}$. While this may seem to be an unusual and even incongruous objective, recall the prior statement that the task of *sparse* signal reconstruction is characterized by the pursuit of a low-dimensional sought-after signal. Hence, the minimization of the ℓ_0 norm, in a variety of signal processing tasks, is the appropriate objective.

Now, suppose that the sought after signal $\vec{x}$ is not a vector but a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$. In such a case, the equivalent objective to 1 is:

$$\min_{s.t. \mathcal{A}\mathbf{X} = \vec{b}} \{\text{rank}(\mathbf{X})\} \tag{3}$$

where $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^p$, and $\vec{b} \in \mathbb{R}^p$. This is the basis pursuit problem we are concerned with in our study, since if an image is given as a matrix $\mathbf{X}$, then we seek a rank minimal representation of $\mathbf{X}$ subject to the constraint $\mathcal{A}\mathbf{X} = \vec{b}$. The principle underlying this problem statement is that images typically have a low rank representation that captures all the crucial information in the image, but if the image is subject to noise, then the image is likely to be numerically full rank. Therefore, the task of "denoising" the image is one that concerns minimizing the rank whilst preserving the crucial information in the image by asserting $\mathcal{A}\mathbf{X} = \vec{b}$. Below, we discuss the expression of this problem in the context of convex optimization.

3 RANK MINIMIZATION WITH THE NUCLEAR NORM

In 2002, Maryam Fazel introduced the *nuclear norm* heuristic for the affinely constrained rank minimization problem [Fazel (2002)]:

$$\min_{\mathbf{X}} \{Rank(\mathbf{X}) : \mathcal{A}(\mathbf{X}) = \vec{b}\} \tag{4}$$

Where $\mathbf{X} \in \mathbb{R}^{m \times n}$ is the decision variable and the linear mapping $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^p$ is given. The vector $\vec{b} \in \mathbb{R}^p$ is called the *observations* of $\mathbf{X}$ and is given. Given the singular value decomposition of $\mathbf{X}$:

$$\begin{aligned} &\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \\ &\text{where } \mathbf{U} \in \mathbb{R}^{m \times r},\ \mathbf{\Sigma} \in \mathbb{R}^{r \times r},\ \mathbf{V} \in \mathbb{R}^{n \times r}, \\ &Rank(\mathbf{X}) = r,\ Diag\,(\mathbf{\Sigma}) = \vec{\sigma} \in \mathbb{R}^r \end{aligned} \tag{5}$$

Where $Diag(.)$ is the common shorthand for the resultant vector formed from the diagonal of an input matrix $Diag : \mathbb{R}^{r \times r} \to \mathbb{R}^r$, e.g. $Diag(\mathbf{X}) = \vec{x},\ s.t.\ \vec{x}_i = \mathbf{X}_{i,i}$. We note that one can easily imagine that the inverse of $Diag$ is an operation $Diag^{-1} : \mathbb{R}^r \to \mathbb{R}^{r \times r}$, e.g. $Diag^{-1}(\vec{x}) = \mathbf{X}$. These shorthands come in handy below for the description of various algorithms.

The nuclear norm, also known as the Schatten 1-norm or the Ky Fan norm, of $\mathbf{X}$ is defined as:

$$||\mathbf{X}||_* = \sum_{i=1}^{r} Diag(\mathbf{\Sigma})_i = \sum_{i=1}^{r} \vec{\sigma}_i \tag{6}$$

In this form, it is easy to recognize the analogous functionality of the nuclear norm and the $L1$-norm, particularly given their respective application to the rank minimization problem, which minimizes $\mathrm{rank}(\mathbf{X})$ and the cardinality minimization problem, which minimizes $||\vec{x}||_0$. The major obstacle to minimizing these two objectives is the fact that both of them are np-hard to minimize. However, the seminal findings by Candès et al. (2006) and Donoho (2006) essentially found that the geometry of the ℓ_0, while non-convex, has a convex hull given by the ℓ_1 norm. Thus, by recasting the cardinality minimization problem using the ℓ_1 norm, we may arrive at optimal solutions for $\min||\vec{x}||_0$ by minimizing $||\vec{x}||_1$, for which there are many polynomial time techniques. This significant result by Candès et al. and Donoho was the starting point for a broad swath of findings extending and applying these initial results.

One such significant subsequent finding was published in 2009 by Recht, Fazel, and Parrilo [Recht et al. (2006)]. Fazel, who first introduced the heuristic in her 2002 PhD dissertation, worked with Recht and Parrilo to provide similar rigour to the nuclear norm heuristic that Candes et al. developed for the $L1$-norm. Recht et al.'s work was inspired by the geometric intuition at the core of the findings in "Robust Uncertainty Principles" and elaborates on the description of Restricted Isometry Property (RIP), the satisfaction of which guarantees minimum rank solutions when applying nuclear norm minimization. Qualitatively, matrices with RIP are *nearly* orthonormal, particularly when operating on sparse vectors, and may be generated by sampling from certain random probability distributions of particular characteristics. The key outcome of Recht et al.'s additions to compressed sensing is a coherent path to the exact recovery of low-rank matrices employing the state of the art algorithms emerging from the field of compressed sensing. Recht et al. provided with their results an alternative basis pursuit problem to equation 4 that has polynomial time, convex algorithms for resolution:

$$\min_{\mathbf{X}} \{||\mathbf{X}||_* : \mathcal{A}(\mathbf{X}) = \vec{b},\ ||\mathbf{X}||_2 \leq 1\} \tag{7}$$

This basis pursuit problem is constrained by the following characteristics:

1. The *operator norm* of a matrix $\mathbf{X}$ is equal to the maximum singular value of $\mathbf{X}$, and is indicated by the symbol $||\mathbf{X}||_2$. In order for the nuclear norm to act as the convex envelope of the $Rank(.)$ function, the decision variable $\mathbf{X}$ must be normalised so that its operator norm is less than or equal to 1.

2. The mapping of $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^p$ is defined by Fazel to be equivalent to "vectorising" $\mathbf{X}$ by stacking the columns of $\mathbf{X}$ on top of each other and then multiplying the resultant vector $\vec{x} \in \mathbb{R}^{m.n}$ by a matrix $\mathbf{A} \in \mathbb{R}^{p \times m.n}$:
   $\mathcal{A}(\mathbf{X}) := \mathbf{A}vec(\mathbf{X}) = \mathbf{A}\vec{x}$, where $vec(.) : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m.n}$

3. The matrix $\mathbf{A}$ is populated by i.i.d. random entries sampled PDF's with zero mean and finite fourth moment. This is to ensure that the projection $\mathbf{A}\vec{x}$ is nearly isometric as the dimensions $m, n \rightarrow \infty$. Recht et al. gives three key conditions for a matrix $\mathbf{X}$ to be considered nearly isometric, and two of the three are instantly satisfied if $\mathbf{A}$ has entries:
   $\mathbf{A}_{i,j} \sim \mathcal{N}(0, 1/p)$
   Yin et al. proved in 1988 that the maximum singular value of a sufficiently large matrix sampled as such will be nearly $1 + \sqrt{mn/p}$ [Yin et al. (1988)], and El Karoui proved that the norm of a random projection is exactly $\sqrt{mn/p}$ [Karoui (2004)].

4. Finally, Recht et al. showed that there exists positive constants c_1, c_0 independent of $m, n, p,$ and r such that $p \geq c_0 r(m+n)\log(mn)$ observations are sufficient to yield close-enough isometry with probability greater than $1 - e^{-c_1 p}$. Recht et al. then state that for practical treatment, $p = O(r(m+n)\log(mn))$ is a conservative enforcement, and demonstrated that $p \in [2r(m+n-r),\ 4r(m+n-r)]$ observations was in fact sufficient in their proof of concept cases for exact recovery.

   For comparison, Candes et al. showed that RIP held for the cardinality minimization problem with observation $p = O(nr\log(n))$, where n is the larger dimension of the matrix.

Having identified these constraints to basis pursuit problem of equation 7, and acknowledging that the constraint $\mathcal{A}(\mathbf{X}) = \vec{b}$ must be relaxed if we are to obtain a rank deficient solution, we are prepared to render the question into a computationally tractable problem statement in Legrangian form:

$$\min_{\mathbf{X}}\{\mu||\mathbf{X}||_* + \frac{1}{2}||\mathbf{A}\vec{x} - \vec{b}||_2^2\} \tag{8}$$

where μ is a positive parameter for the relative weighting of the nuclear and L_2-norm. However, for our particular study, we are concerned the the matrix completion problem, which is given as follows:

$$\min_{\mathbf{X}}\{\mu||\mathbf{X}||_* + \frac{1}{2}||\mathbf{X}_{i,j} - \mathbf{M}_{i,j}||_2^2 : (i,j) \in \Omega\} \tag{9}$$

where Ω denotes a subset of index pairs corresponding to known elements of the matrix $\mathbf{X}$, stored in the sparse matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$. Thus, the formulation of the matrix $\mathbf{A}$ is sparse and populated by ones and zeros, and is considered instead a *sampling* matrix. There is only a single value of one per row of $\mathbf{A}$ in this case, and it resides at the column coordinate of $\mathbf{A}$ corresponding to the row coordinate of a known value element of the vectorised matrix $vec(\mathbf{X}) = \vec{x}$. It is easy to see that the definition of such a matrix $\mathbf{A}$ satisfies the restricted isometry property, and we give explanation as follows:

Suppose one has an image $\mathbf{X} \in \mathbb{R}^{m \times n}$ that is corrupted by noise so that only p indices Ω:

$$\Omega = \{(i_1, j_1), (i_2, j_2), ..., (i_p, j_p)\} \tag{10}$$

have known values in the set $\vec{b} \in \mathbb{R}^p$. We begin by vectorising $\mathbf{X}$ as follows:

$$\vec{x} = \text{vec}\left(\begin{bmatrix} X_{1,1} & X_{1,2} & X_{1,3} & ... & X_{1,n} \\ X_{2,1} & X_{2,2} & X_{2,3} & ... & X_{2,n} \\ X_{3,1} & X_{3,2} & X_{3,3} & ... & X_{3,n} \\ \vdots & & & & \\ X_{m,1} & X_{m,2} & X_{m,3} & ... & X_{m,n} \end{bmatrix}\right) = \begin{bmatrix} X_{1,1} \\ X_{2,1} \\ X_{3,1} \\ \vdots \\ X_{m,1} \\ X_{1,2} \\ X_{2,2} \\ \vdots \\ X_{m,2} \\ \vdots \\ X_{m,n} \end{bmatrix} \tag{11}$$

Now suppose that the our observed indices included $\{(1,1),(3,1),(m,1),(1,2),(m,2)\} \subset \Omega$. Thus, the corresponding observation vector b would have form:

$$\mathbf{A}\vec{x} = \vec{b} = \begin{bmatrix} X_{1,1} \\ X_{3,1} \\ \vdots \\ X_{m,1} \\ X_{1,2} \\ \vdots \\ X_{m,2} \\ \vdots \\ b_p \end{bmatrix} \in \mathbb{R}^p \tag{12}$$

We may devise a matrix $\mathbf{A} \in \mathbb{R}^{p \times m.n}$ that samples from the vector $\vec{x}$ as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & ... & ... & ... & ... & ... & ... & 0 \\ 0 & 0 & 1 & ... & ... & ... & ... & ... & ... & 0 \\ 0 & 0 & 0 & ... & 1 & 0 & ... & ... & ... & 0 \\ 0 & 0 & 0 & ... & 0 & 1 & ... & ... & ... & 0 \\ 0 & 0 & 0 & ... & 0 & 0 & ... & 1 & ... & 0 \\ \vdots & & & & & & & & & \end{bmatrix} \tag{13}$$

Notice how each row of the matrix $\mathbf{A}$ contains a single value of 1, which corresponds to the vector coordinate of the appropriate observation of $\mathbf{X}$. Therefore, the observations vector $\vec{b}$ is a list of known element values of $\mathbf{X}$, and hence the rank minimization problem becomes a matrix *completion* problem.

Now we can see that the since the matrix $\mathbf{A}$ has values of one at a unique column coordinate in each row, then the matrix is clearly row orthonormal. Since the restricted isometry property is used to generate approximately orthogonal (row and column orthonormal) matrices for the projection $\mathbf{A}\vec{x} = \vec{b}$, we can see that this formulation of $\mathbf{A}$ is clearly full rank. This is important, since the projection acts as a constraint to the rank minimization problem, and therefore we require matrix equation $\mathbf{A}\vec{x} = \vec{b}$ be uniquely solved by $\vec{x}$ in order for it to serve as a constraint.

Obviously, the storage and full operation of the $\mathbf{A}$ in this form is inappropriate, and so in practice, we instead defined our operation $\mathbf{A}\vec{x}$ to be defined by sampling indices Ω directly from $\mathbf{X}$ to generate $\vec{b}$. An important caveat is that the operation $\mathbf{A}^T\vec{b}$, which we use in our algorithm, is similarly defined by assigning the values of $\vec{b}$ to their appropriate coordinate in the vectorised $\mathbf{X}$, with zeros elsewhere. We can see this using our above example as follows:

$$\mathbf{A}^T\vec{b} = \mathbf{A}^T \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ \vdots \end{bmatrix} = \begin{bmatrix} b_1 \\ 0 \\ b_2 \\ \vdots \\ b_3 \\ b_4 \\ \vdots \\ b_5 \\ 0 \\ \vdots \end{bmatrix} = \begin{bmatrix} X_{1,1} \\ 0 \\ X_{3,1} \\ \vdots \\ X_{m,1} \\ X_{1,2} \\ \vdots \\ X_{m,2} \\ 0 \\ \vdots \end{bmatrix} \tag{14}$$

In their publication, Recht et al. proceed to solve equation 8 by means of semi-definite programming using solvers such as SDPT3 or SeDuMi. However, in 2009, following successes with Bregman iterative algorithms and fixed point programming methods for $L1$-regularized problems, Shiqian Ma, Donald Goldfarb, and Lifeng Chen proposed several fixed point programming based algorithms for equation 8 that exhibit considerable performance advantages over semi-definite programming methods described by Recht [Ma et al. (2011)]. Ma et al. showed that their algorithms can solve the matrix completion problem equation 9 for matrix dimensions and rank exceeding what SeDuMi and SDPT3 can solve, as well as recovering matrices from only 20% of the elements.

For the scope of this research, our proof of concept case makes use of the fixed point continuation technique described by Ma et al., applied in conjunction with a Bregman two-step formulation formulated by Yin et al. in 2008 [YIN et al. (2008)].

4 BREGMAN-WRAPPED FIXED POINT CONTINUATION (FPC)

Ma et al. propose a composite Bregman iterative and Fixed Point method for the resolution of the Equation 8 (Ma et al. (2011)). At its core, the algorithm performs a two component minimization on the decision variable $\mathbf{X}$'s Euclidean distance from its noised state and the nuclear norm of said matrix, weighted by the Legrangian parameter μ. It accomplishes this by simultaneously minimizing the L_2 norm $||\mathbf{A}\vec{x} - \vec{b}||_2^2$ via the widely applied Landweber iteration and minimizing the Nuclear norm $||\mathbf{X}||_*$ by employing Soft-Thresholding in a similar fashion to Yin et al.'s application of Soft-Thresholding to the L_1 norm (YIN et al. (2008)).

The modification to Yin et al.'s Soft-Thresholding formulation as it is applied to the cardinality of a vector decision variable is defined by applying what Ma et al. refer to as the *matrix shrinkage* operator to the intermediate matrix that follows Landweber iteration, which we call $\mathbf{X}_{k+1/2}$. Following Landweber iteration and thus the generation of $\mathbf{X}_{k+1/2}$, we determine the matrix $(\mathbf{X}_{k+1/2})^T(\mathbf{X}_{k+1/2})$'s spectral matrix $\mathbf{\Lambda}_{\mathbf{X}^2_{k+1/2}} = \left(\mathbf{\Sigma}_{k+1/2}\right)\left(\mathbf{\Sigma}_{k+1/2}\right)$, which has the squared singular values of $\mathbf{X}_{k+1/2}$ along its diagonal, which we call:

$$\vec{\sigma}_{k+1/2} := \mathrm{Diag}\left(\mathbf{\Sigma}_{k+1/2}\right) \tag{15}$$

We denote this *matrix shrinkage* operator by the symbol $\mathcal{S}_\nu(.)$, and it is defined as follows:

$$\mathcal{S}_\nu\left(\mathbf{X}_{k+1/2}\right) := \left(\mathbf{U}_{k+1/2}\right)\left[\mathrm{Diag}^{-1}\left(s_\nu(\vec{\sigma}_{k+1/2})\right)\right]\left(\mathbf{V}^T_{k+1/2}\right) \tag{16}$$

where $s_\nu(.)$ refers to Ma et al.'s *vector shrinkage* operator, and applies to the vector $\vec{\sigma}_{k+1/2}$ exactly as soft-thresholding applies to a vector:

$$s_\nu(\vec{x}) := \bar{x}, \text{with } \bar{x}_i = \begin{cases} x_i - \nu, & \text{if} \quad x_i - \nu > 0 \\ 0, & o.w. \end{cases} \tag{17}$$

where ν is the product of the two parameters μ and τ, the Legrangian weight and Landweber relaxation parameter, respectively.

Together, the Landweber iteration and matrix-shrinkage operator comprise Ma et al.'s *fixed point iterative* (FPI) scheme. The *fixed point continuation* (FPC) scheme is accomplished by wrapping the fixed point iteration with another loop that incrementally reduces the value of the Legrangian weightμ by multiplying it by a reduction parameter η, which, following the work by Yin wet al., we set to be equal to 0.25. Following the work of Yin et al., Ma et al. wrap their fixed point continuation method with a Bregman iterative update step. This implementation amounts to an additional loop around the fixed point continuation, where the observations vector $\vec{b}$ is updated using the previous Bregman iteration's results. The Bregman-wrapped fixed point continuation scheme is thus expressed as follows:

Bregman Iterative Algorithm:

- $\vec{b}_0 \leftarrow \mathbf{0}, \mathbf{X}_0 \leftarrow \mathbf{0}$
- **for** $k = 1, 2, ...$ **do** $\{$
- $\quad \vec{b}_{k+1} \leftarrow \vec{b} + \left(\vec{b}_k - \mathcal{A}(\mathbf{X}_k)\right)$
- $\quad \mathbf{X}_{k+1} \leftarrow \text{argmin}_{\mathbf{X}}\{\mu||\mathbf{X}||_* + \frac{1}{2}||\mathcal{A}(\mathbf{X}) - \vec{b}_{k+1}||_2^2\}$
- $\}$

The latter step is solved with the fixed point continuation scheme described above, and therefore FPC executes once for every Bregman iteration called. Bregman iterations converge very rapidly, and therefore throughout the course of our own work, we fix our number of Bregman iterations to be equal to 3.

We are now prepared to describe the algorithm employed in the current scope of work. This is a Bregman iterative wrapped fixed point method to be employed on the problem statement from Equation 8. This algorithm will yield an optimal matrix $\mathbf{X}^*$ satisfying the optimality conditions of Equation 8.

$\mathbf{FPC}(\mathbf{G})$:

- define $\tau, \mu, \mu_{min}, \eta$
- $\nu \leftarrow \tau\mu$
- $\mathbf{G_{00}} \leftarrow \mathbf{G} \in \mathbb{R}^{m \times n}$
- // Matrix $\mathbf{G}$ is inputted as above
- $\mathrm{SVD}\left(\mathbf{G_{00}}\right) : \mathbf{G_{00}} = \mathbf{U_{00}}\mathbf{\Sigma_{00}}\mathbf{V_{00}^T}$
- $\sigma_1 \leftarrow \mathbf{\Sigma_{00}}_{1,1}$
- $\mathbf{G_0} \leftarrow \frac{1}{\sigma_1}\left(\mathbf{G_{00}}\right)$
- $\vec{g_0} \leftarrow vec(\mathbf{G_0})$
- $\mathbf{A} \leftarrow \mathrm{RIP}(m * n, p, \frac{1}{p}) \in \mathbb{R}^{m.n \times p},\ \mathbf{A}_{i,j} \sim \mathcal{N}(0, \frac{1}{p})$
- $\vec{b} \leftarrow \mathbf{A}\left(\vec{g_0}\right),\ \vec{b} \in \mathbb{R}^p$
- $\vec{b_l} \leftarrow \mathbf{0}, \mathbf{X}_k \leftarrow \mathbf{0},\ \vec{b_l} \in \mathbb{R}^p,\ \mathbf{X}_k \in \mathbb{R}^{m \times n}$
- **for** $l = 0, 1, 2, ...$ { // begin Bregman iteration
- $\quad \vec{b}_{l+1} = b + \left(\vec{b_l} - \mathcal{A}(\mathbf{X_k})\right)$
- $\quad k = 0$
- $\quad$ **while** $\mu > \mu_{min}$**do** { // Fixed Point Continuation
- $\quad\quad$ **while** NOT converged, **do** { // begin Fixed Point Iteration
- $\quad\quad\quad k++$
- $\quad\quad\quad$ compute $\mathbf{X}_{k+1/2} = \mathbf{X}_k + \tau\mathcal{A}^T\left(\mathcal{A}(\mathbf{X}_k) - \vec{b}_{l+1}\right)$ // Landweber iteration
- $\quad\quad\quad$ compute $\mathrm{SVD}\left(\mathbf{X}_{k+1/2}\right) :\ \mathbf{X}_{k+1/2} = \mathbf{U}_k\mathbf{\Sigma}_k\mathbf{V}_k^T$
- $\quad\quad\quad \mathbf{X}_{k+1} = \mathcal{S}_\nu\left(\mathbf{X}_{k+1/2}\right) = \mathbf{U}_k\left[\mathrm{Diag}^{-1}\left(s_\nu(\mathbf{\Sigma}_k)\right)\right]\mathbf{V}_k^T$ // soft-thresholding
- $\quad\quad$ } // end **while**
- $\quad\quad \mu \leftarrow \eta\mu$
- $\quad$ } // end **while**
- } // end **for**
- **return** $\mathbf{X}_{k+1}$

As a proof of concept, we demonstrate the efficacy of this algorithm below on the reconstruction of the Harvard Paulson School of Engineering logo. When the associated png is converted to greyscale, one obtains a 283 by 246 matrix of values ranging from 0 to 256. This matrix is numerically full rank, but in the spirit of compressed sensing, we truncate the rank of this image by setting the last 205 singular values to zero. This results in a grainier image, but importantly, is certainly still recognizable as the logo. This is the key assumption behind the working of the compressed sensing algorithms: the authentic signal has a low dimensional representation, but with noise, the signal is observed to be of high dimensionality. Therefore, using a dimensionality reduction method such as our matrix shrinkage operator $\mathcal{S}_\nu$ alongside a constraint enforcement such as an ℓ_2 penalty on the distance from the original image ensures that we are reducing the dimensionality of our system without compromising the integrity of the authentic signal.

In the case of the logo problem, we mask 50 percent of the pixels at random by setting them equal to zero, with the indices of the remaining unmasked pixels defining the set Ω. This allows us to define our $\mathcal{A}$ operator as above. We normalize the masked image by its operator norm and then feed it to the Bregman-wrapped fixed point continuation technique. The results are given in figure 1, below.

5 EXPERIMENT

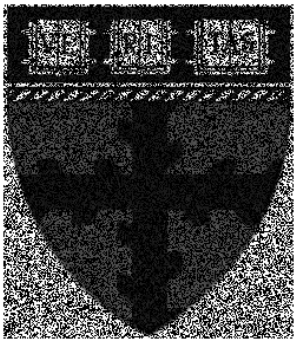
We aim to evaluate the quality of the matrix reconstructions yielded by the above methods. For simplicity and reproducibility, we work with the MNIST data set and various sets of linear logistic regression classifiers, each with the goal of correctly classifying all ten digits represented in the data, which is shown to the model as a flattened matrix. As with many frequentist approaches, these classification methods can suffer from over-fitting to training data when such samples are scarce, which can be observed as a larger difference between the model's training accuracy and its validation score as compared to the asymptotic difference observed as the number of training samples increases. As such, we use this difference as a simple over-fitting measure for each of the sets of classifiers described below.

Note that while logistic regression classifiers are **not** the most suitable method for non-binary classification of image data, they are easier to train on such high-dimensional data than other more appropriate methods and are likely to exhibit large approximately-asymptotic measures of over-fitting for finite amounts of training data, which is desirable for the comparisons demonstrated here. We train three sets of these classifiers:

- **Using the unmodified MNIST data**. We vary the number of samples n used to train these models in order to generate base measures of training accuracy and validation scores for this type of classifier. Validation scores are measured as classification accuracy on a set of 3000 unseen in-distribution MNIST samples. We expect to observe over-fitting for small n.

(a) original RGB Harvard Paulson school of engineering image, numerically full rank of 246

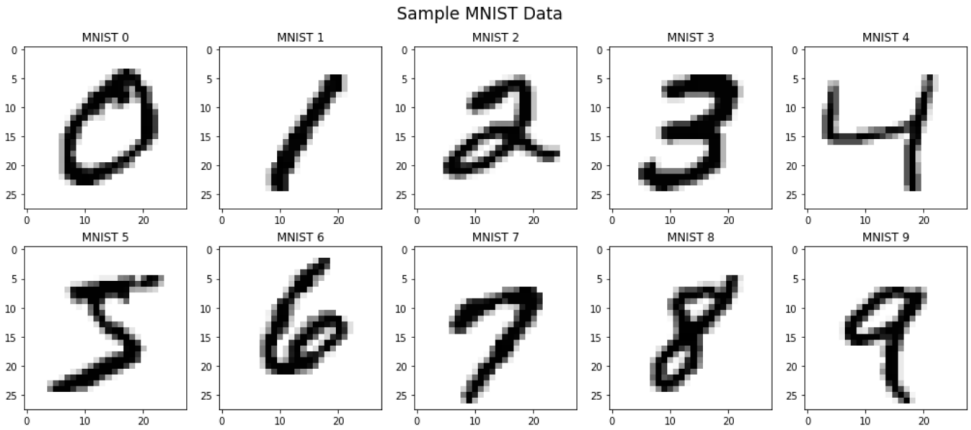
(b) image a) converted to greyscale and truncated to rank 41



(c) 50 percent of the pixels in image b) are masked, set to zero

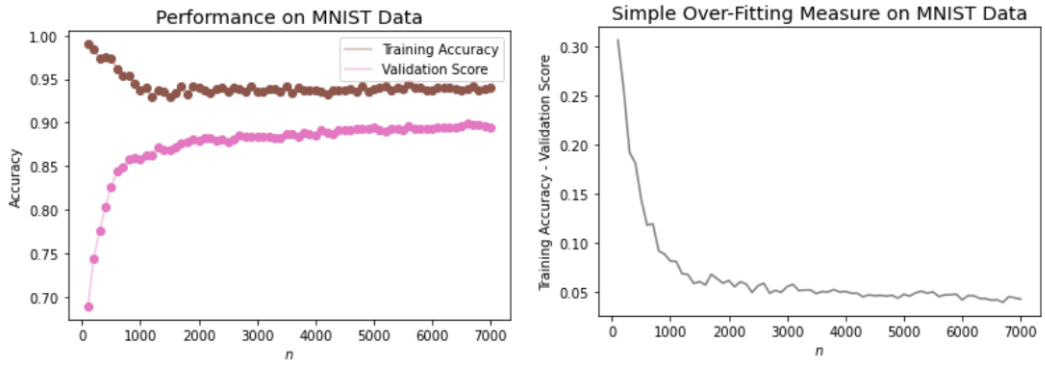
(d) reconstruction of image b) via Bregman-wrapped fixed point continuation. Note the general similarity to image b), but with minor defects at edges.

Figure 1: Proof of concept demonstration of the efficacy of fixed point methods on the task of matrix completion, using the Harvard Paulson school of engineering logo.

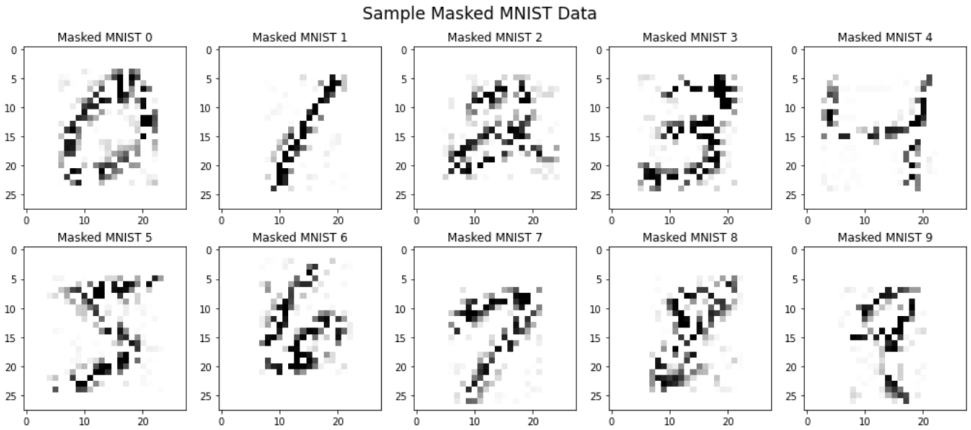


We find that for small n, the models exhibit higher training accuracy and lower validation scores, as expected. The models achieve an asymptotic difference between training

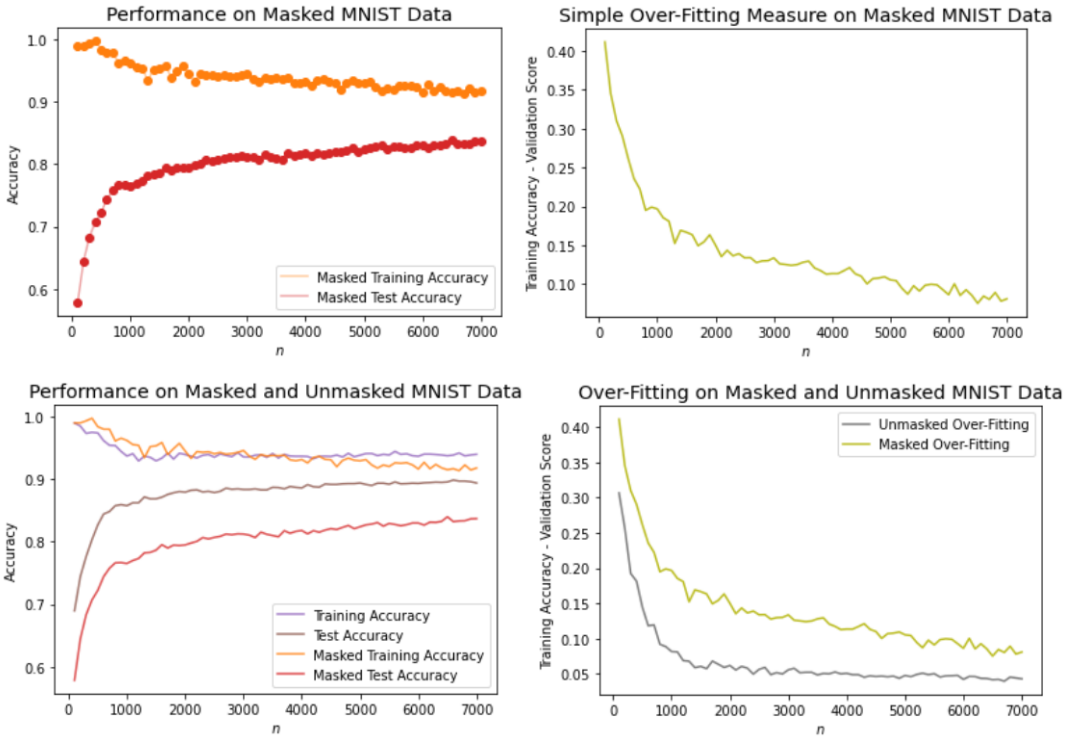
accuracy and validation score of approximately 0.05 for $n \geq 2000$. This will serve as our base case of comparison when discussing the remaining sets of classifiers.



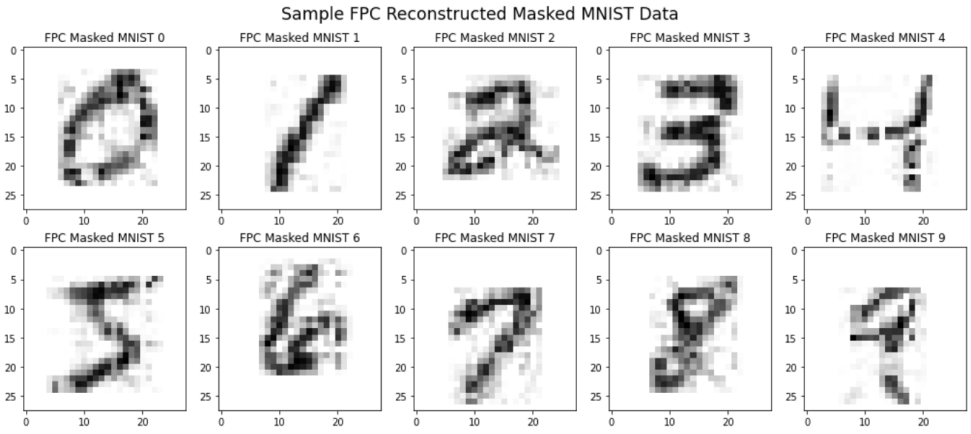
- **Using masked MNIST data**. The unmodified MNIST data consists of sets of 28×28 full-rank grey-scale images. This masked data is modified from the original data set in two ways: the rank of all images has been truncated such that they are rank 5 (where the remaining singular values are the highest of those corresponding to the original image), and each pixel of the resulting truncated-rank image was masked (set to 0) with a probability of 0.5. Validation scores are measured as classification accuracy on masked data corresponding to the validation set used in the previous set of classifiers. We expect these models to perform worse on this validation set as compared to the validation scores of the previous set of models, simply on account of their noisier distribution.



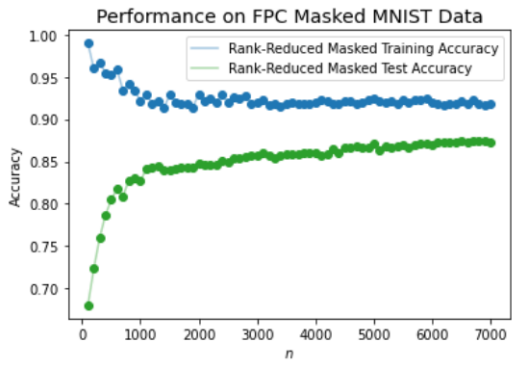
We find that these models exhibit high over-fitting for $n \leq 7000$, and as such the over-fitting measure does not converge to its asymptotic value in the plots displayed below. This is a result of a property of logistic regression classifiers: the fit of our models is highly sensitive to aleatoric noise in the data. The random masking of pixels in the data contributes to this aleatoric noise, which accounts for any non-zero singular values of the masked matrix (other than the largest 5). Although the unmasked data are full-rank and also have low-magnitude singular values corresponding in part to aleatoric noise, not all of these values account for aleatoric noise entirely. That is, not only is information lost in the rank truncation of our data (in the form of either in-distribution features or heteroskedastic noise) which contributes to the low validation scores of these models, this information is replaced with uninformative noise that increases over-fitting in our frequentist ensemble.

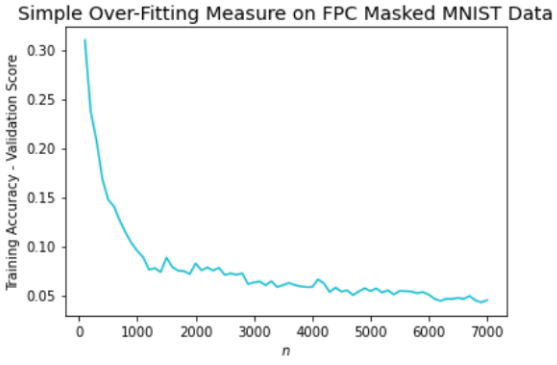


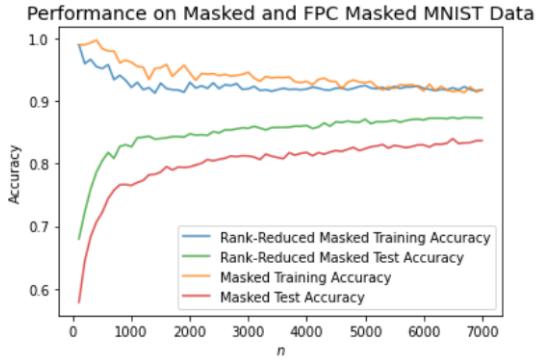
- **Using masked MNIST data reconstructed via FPC**. We apply the FPC algorithm to the masked data used in training and validating our second set of classifiers. Validation scores are measured as classification accuracy on masked data reconstructed via FPC corresponding to the same validation samples used for the above two sets of classifiers. We expect that these models will reduce over-fitting and exhibit higher validation scores as compared to the second ensemble described above, as the FPC algorithm will aim to minimize the rank of these images, thus resulting in a decrease in aleatoric noise. Because this is not an exact reconstruction of the original full-rank matrix, we also expect lower validation scores for these models as compared to those trained on unmasked data.

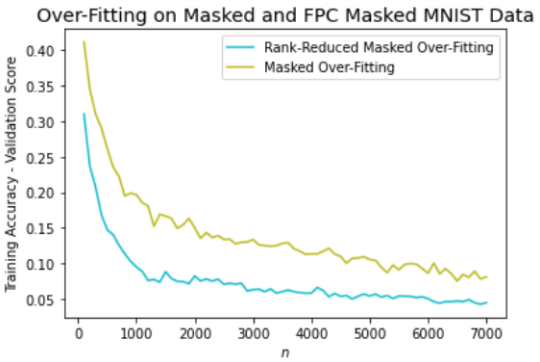


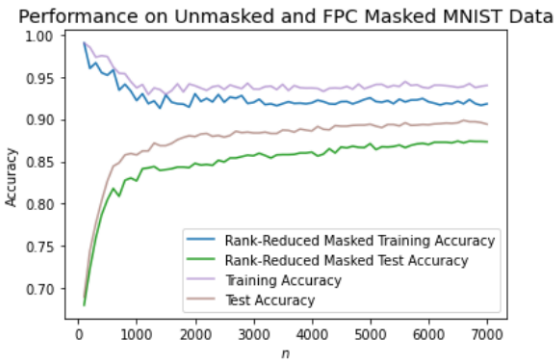
We find that not only does the FPC algorithm improve validation scores on masked data and reduce over-fitting, but the over-fitting exhibited in this ensemble nearly mirrors that of the ensemble trained on unmasked data. This suggests that the FPC algorithm was able to remove nearly all aleatoric noise added in the masking of the data, and that the difference between the validation scores of this ensemble and that of the ensemble trained on unmasked data is nearly entirely attributable to the information lost in the rank truncation of the original data.

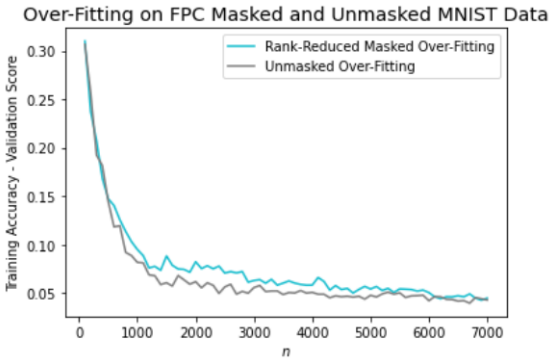












We have demonstrated how the FPC algorithm can be used to reduce over-fitting in a simple logistic regression classifier for image data. It is worth exploring if similar effects can be observed in ensembles of n-dimensional polynomial logistic regression. We may also be able to determine the effect of the FPC algorithm on a Bayesian logistic classifier's classification uncertainty.

REFERENCES

E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on Pure and Applied Mathematics*, 2006.

D. L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

M. Fazel. *Matrix Rank Minimization with Applications*. PhD thesis, Stanford University, 2002.

N. E. Karoui. *New results about random covariance matrices and statistical applications*. PhD thesis, Stanford University, 2004.

S. Ma, D. Goldfarb, and L. Chen. Fixed point and bregman iterative methods for matrix rank minimization. *Mathematical Programming Series A*, 128(1):321–353, 2011.

B. Recht, M. Fazel, and P. A. Parrilo. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

W. YIN, S. OSHER, D. GOLDFARB, and J. DARBON. Bregman iterative algorithms for l1-minimization with applications to compressed sensing. *Society for Industrial and Applied Mathematics*, 1(1):143–168, 2008.

Y. Yin, Z. Bai, and P. Krishnaiah. On the limit of the largest eigenvalue of the large dimensional sample covariance matrix. *Probability Theory and Related Fields*, 78:509–521, 1988.