Compressed Sensing

May 7, 2020

1 Overview

Short overwiew of matrix completion:

$$\min ||X||_* \quad s.t. \quad AX = B$$

We need to find the min rank X such that $AX = B$

2 Douglas- Rachford Splitting

In the noiseless case we can use DRS to solve this numerically via a fixed-point iteration:

$$z^{(k+1)} = z^{(k)} + Prox_{\gamma g}\Big(2Prox_{\gamma f}(z^{(k)}) - z^{(k)}\Big) - Prox_{\gamma g}(z^{(k)})$$

We know that for $f = \delta_C$ with $C = \{X|AX = B\}$ we obtain

$$\begin{aligned} Prox_{\gamma f}(x) = \Pi_C(x) &= x + A^+(b - Ax) \\ &= x + A^T(AA^T)^{-1}(b - Ax) \\ &= x + A^T(b - Ax) \end{aligned}$$

and via the SVD

$$\begin{aligned} Prox_{\gamma g}(x) &= US_\delta(\sigma(x))V^T \\ &= U(\sigma(x) - \gamma)_+V^T \end{aligned}$$

3 FISTA

In the noisy case we are trying to minimize the following objective (in the Lagrangian formulation):

$$\min_X \lambda||X||_* + \frac{1}{2}||AX - B||_2^2$$

The FISTA algorithm is formulated with respect to the following objective:

$$\min_X f(x) + g(x)$$

where we assume f and g to be sufficiently smooth, i.e. $f \in C^{1,1}(\mathbb{R}^n)$, which means

$$\exists L(f) : ||\nabla f(x) - \nabla f(y)|| \leq ||x - y|| \forall x, y \in \mathbb{R}^n$$

In our case we have $f(x) = \frac{1}{2}||AX - B||_2^2$ and $g(x) = \lambda||X||_*$ and by simple calculation: $L(f) = 2\lambda_{max}(A^T A)$

Algorithm 1 FISTA with constant step size

Input Lipschitz- constant $L(f)$ of ∇f, $y_1 = x_0 \in \mathbb{R}^n, t_1 = 1$

1: **for** $k = 1, \ldots$ **do**
2: $\quad x_k = p_L(y_k)$
3: $\quad t_{k+1} = \frac{1+\sqrt{1+4*t_k^2}}{2}$
4: $\quad y_{k+1} = x_k + \frac{t_k - 1}{t_{k+1}}(x_k - x_{k-1})$

In this algorithm we def.

$$\begin{aligned}
Q_L(x, y) &= f(y) + \langle x - y, \nabla f(y) \rangle + \frac{L(f)}{2}||x - y||^2 + g(x) \\
p_L(y) &= \arg\min_x Q_L(x, y) \\
&= \arg\min_x \Big(g(x) + \frac{L(f)}{2} \Big\| x - \Big(y - \frac{1}{L} \Big) \Big\|^2 \Big) \\
&= Prox_{1/Lg}(y - \frac{1}{L}\nabla f(y))
\end{aligned}$$

To simplfify this expression further, we examine the the SVD of the input matrix of p_L, which we call B for now. It turns out we arrive at a singular value thresholding step: Let $B = U\Sigma V^T = Udiag(\sigma(b))V^T$ and

$$Prox_{\lambda g}(B) = Udiag(\sigma(B) - \lambda)_+ V^T$$

which finally reveals $Prox_{1/Lg}\Big(y - \frac{1}{L}\nabla f(y)\Big) = Udiag\Big(\sigma(y - \frac{1}{L}\nabla f(y)) - \frac{1}{L}\Big)_+ V^T$

I'm not sure if this is correct