Compressed Sensing

May 17, 2020

1 Preliminaries

We define useful mappings used in convex optimization, namely the proximal mapping

$$\mathrm{Prox}_f^{\lambda}(y) = \arg\min_x \left(f(x) + \frac{1}{2\lambda} ||x - y|| \right)$$

2 Overview

We are given an incomplete matrix M with known entries $(i, j) \in \Omega$. In the general form matrix completion reads

$$\min ||X||_* \quad \text{s.t.} \quad \mathcal{A}(X) = B$$

where $\mathcal{A}$ is a linear map $\mathcal{A} : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{|\Omega|}$

and B encodes our knowlege of M. In our case we choose $\mathcal{A} = \mathcal{P}_\Omega$ with

$$[\mathcal{P}_\Omega(X)]_{i,j} = \begin{cases} \mathrm{vec}(X)_{j+mi} & \text{if } (i, j) \in \Omega \\ 0, & \text{otherwise} \end{cases}$$

We denote $A \in \mathbb{R}^{|\Omega|\times mn}$ the matrix corresponding to the mapping $\mathcal{P}_\Omega$. With this we can rewrite the objective above:

$$\min ||X||_* \quad \text{s.t.} \quad \mathcal{P}_\Omega(X) = A\mathrm{vec}(X) = B = \mathcal{P}_\Omega(M)$$

3 Douglas- Rachford Splitting

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

$$z^{(k+1)} = z^{(k)} + \mathrm{Prox}_g^{\gamma} \left(2\, \mathrm{Prox}_f^{\gamma}(z^{(k)}) - z^{(k)} \right) - \mathrm{Prox}_g^{\gamma}(z^{(k)})$$

To do this, we express our objective as a sum of two functions $f+g$ where $f(X)=\delta_C(X)$ with $C=\{X|AX=B\}$ and $g(X)=||X||_*$. We then obtain

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

To express the proximal mapping of the nuclear norm, we define the singular value thresholding operator $\mathcal{D}_\gamma$. For a matrix Y with the SVD $Y=U\Sigma V^T=U\,\mathrm{diag}(\sigma(Y))V^T$ it is expressed by $\mathcal{D}_\gamma(Y)=U\,\mathrm{diag}(\sigma(Y)-\lambda)_+V^T$. Thus we get

$$\begin{aligned}\mathrm{Prox}_g^\gamma(x)&=U(\sigma(x)-\gamma)_+V^T\\&=\mathcal{D}_\gamma(x)\end{aligned}$$

4 FISTA

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

$$\min_X \lambda||X||_*+\frac{1}{2}||AX-B||_2^2 \tag{4.1}$$

For some $\lambda>0$. 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

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

In our case we have $f(x)=\frac{1}{2}||AX-B||_2^2=\frac{1}{2}\langle AX-B,AX-B\rangle$ and $g(x)=\lambda||X||_*$ .By simple calculation we get

$$\nabla f(X)=A^TAX-A^TB=A^T(AX-B) \tag{4.2}$$

$$L(f)=2\lambda_{\max}(A^TA) \tag{4.3}$$

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** $k=1,...$ **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.

$$
\begin{aligned}
Q_L(x,y) &= f(y) + \langle x-y, \nabla f(y)\rangle + \frac{L(f)}{2}||x-y||^2 + g(x) \\
p_L(y) &= \arg\min_x Q_L(x,y) \\
&= \arg\min_x \left(g(x) + \frac{L(f)}{2} \left\| x - \left(y - \frac{1}{L}\nabla f(y) \right) \right\|^2 \right) \\
&= \arg\min_x \left(\frac{g(x)}{\lambda} + \frac{L(f)}{2\lambda} \left\| x - \left(y - \frac{1}{L}\nabla f(y) \right) \right\|^2 \right) \\
&= \mathrm{Prox}_{g/\lambda}^{\lambda/L} \left(y - \frac{1}{L}\nabla f(y) \right)
\end{aligned}
$$

Since we are trying to find the proximal mapping for the nuclear norm, we can apply the thresholding operator just like in our iteration step for DRS. We get

$$
\begin{aligned}
\mathrm{Prox}_{g/\lambda}^{\lambda/L} \left(y - \frac{1}{L}\nabla f(y) \right) &= U \,\mathrm{diag} \left(\sigma \left(y - \frac{1}{L}\nabla f(y) \right) - \frac{\lambda}{L} \right)_+ V^T \\
&= \mathcal{D}_{\lambda/L} \left(y - \frac{1}{L}\nabla f(y) \right) \\
&= \mathcal{D}_{\lambda/L} \left(y - \frac{1}{L} A^T (Ay - B) \right)
\end{aligned}
$$

I'm not sure if this is correct