Algorithms for Big Data Analysis : Homework #3

Due on March 28, 2017 at 23:59pm

Wen Zaiwen

Hu Tao

Problem 1

1. For each of the following function on $\mathbb{R}^n$, explain how to calculate a subgradient at a given x.
   A reference on subgradients is
   `http://bicmr.pku.edu.cn/~wenzw/bigdata/subgradients.pdf`
   (a) $f(x) = \|Ax - b\|_2 + \|x\|_2$ where $A \in \mathbb{R}^{m\times n}$ and $x \in \mathbb{R}^n$.
   (b) $f(x) = \inf_y \|Ay - x\|_\infty$ where $A \in \mathbb{R}^{n\times m}$ and $x \in \mathbb{R}^n$.

Solution

(a). 对于$||x||_2$，它的一个次梯度为$\frac{x}{||x||_2}$,因此对于$||Ax-b||_2$,它的一个次梯度为$\frac{A^T(Ax-b)}{||Ax-b||_2}$
因此原问题的一个次梯度是:

$$\frac{x}{||x||_2} + \frac{A^T(Ax-b)}{||Ax-b||_2}$$

(b)
//TODO

Problem 2

2. Give a formula or simple algorithm for evaluating the proximal operator

$$\operatorname{prox}_f(x) = \arg\min_u \left(f(u) + \frac{1}{2}\|u - x\|_2^2 \right).$$

   (a) $f(x) = \|x\|_1$ with domain $\operatorname{dom}(f) = \{x \mid \|x\|_\infty \le 1\}$
   (b) $f(x) = \max_k x_k$
   (c) $f(x) = \|Ax - b\|_1$ where $AA^\top = D$ and D is a diagonal matrix whose diagonal elements are positive.

Solution

(a).
如果没有x的定义域限制，原问题就是一个经典的lasso问题。它的解为Shrinkage。

$$\begin{gathered}x \le -1, prox_f(x) = x + 1;\\ -1 \ge x \le 1, prox_f(x) = 0;\\ x \le 1, prox_f(x) = x - 1;\end{gathered}$$

接下来再考虑x的定义域，由于无穷范数球的各个分量是独立的，因此他的projection如下:

$$\begin{gathered}if|x_i| \le 1, P_C(_i) = x_i;\\ elseP_C(x_i) = sign(x_i)\end{gathered}$$

做映射以后得到最终的结果如下:

$$\begin{gathered}x \le -2, , prox_f(x) = 2;\\ -2lex \le -1, prox_f(x) = x;\\ -1 \le x \le 1, prox_f(x) = 0;\end{gathered}$$

Problem 2 continued on next page. . .

$$1 \le x \le 2, prox_f(x) = x;$$
$$x \ge 2, prox_f(x) = 2;$$

(b).
TODO
(c).
如果C是affine set,$C = x|Ax = b, A \in R^{pxn}$,那么我们有在C上的projection:

$$P_C(X) = x + A^T(AA^T)^{-1}(b - Ax),$$

当p远小于n，或者$AA^T = I$时,上述映射计算量 inexpensive.
在本问题中，我们可以先将问题转换为:

$$\begin{aligned} min_{u,y}\ & f(y) + \tfrac{1}{2}||u - x||_2^2 \\ S.t\ & Au + b = y \end{aligned}$$

u是我们要求的变量，我们可以先将u映射到Au=y-b这个affine set，得到:

$$\begin{aligned} u &= x + A^T(AA^T)^{-1}(y - b - Ax) \\ &= x + A^T D^{-1}(y - b - Ax) \\ &= (I - A^T D^{-1} A)x + A^T D^{-1}(y - b) \end{aligned}$$

同时也可以得到:
$u - x = (A^T D^{-1})(y - b - Ax)$
将u带入原问题可以将u消去，原问题变成了:

$$\begin{aligned} & min_y\ \ f(y) + \tfrac{1}{2}(u - x)^T(u - x) \\ =\ & min_y\ \ f(y) + \tfrac{1}{2}(y - b - Ax)^T(A^T D^{-1})^T(A^T D^{-1})(y - b - Ax) \\ =\ & min_y\ \ f(y) + \tfrac{1}{2}(y - b - Ax)^T(D^{-1})^T(y - b - Ax) \\ =\ & min_y\ \ f(y) + \tfrac{1}{2}(y - b - Ax)^T(D^{-1})(y - b - Ax) \end{aligned}$$

此处可以将中间的D^{-1}对角矩阵提取出来:

$$\begin{aligned} & min_y\ \ f(y) + \tfrac{1}{2det(D)}(y - b - Ax)^T(y - b - Ax) \\ =\ & min_y\ \ f(y) + \tfrac{1}{2det(D)}||y - b - Ax||^2 \end{aligned}$$

因此，y可以求出来:

$$y = prox_{\frac{f}{det(D)}}(Ax + b)$$

将y带入上式，可以得到:

$$u = (I - A^T D^{-1} A)x + A^T D^{-1}(prox_{\frac{f}{det(D)}}(Ax + b) - b)$$

其中f函数的proximal是shrinkage，很方便求解

Problem 3

3. Given $w \in \mathbb{R}^n$, $\alpha, \sigma > 0$, write down an algorithm for solving the problem

$$\min_{t,y} \quad \phi(t,y),$$

where

$$\phi(t,y) := t + \frac{1}{(1-\alpha)n}\sum_{i=1}^{n}(y_i - t)_+ + \frac{\sigma}{2}\|y-w\|_2^2,$$

where $x_+ := \max(x,0)$.

Solution

对于$x_+ := max(x,0)$,其次梯度(记为g)为:

$$\begin{gathered} x<0, g=0; \\ x=0, g\in[0,1]; \\ x>0, g=1; \end{gathered}$$

对y_i求偏导得:

$$\frac{\partial\phi}{\partial y_i} = \frac{1}{(1-a)^n} g(y_i - t) + \sigma(y-w)$$

对t求偏导得:

$$\frac{\partial\phi}{\partial t} = 1 - \frac{1}{(1-a)^n}\sum_{i=1}^{n} g(y_i - t)$$

基于次梯度法采用Jacobi或者Gaussian Sidiel迭代即可求出原问题解。

Problem 4

4. Let $S^n = \{X \in \mathbb{R}^{n\times n} \mid X^\top = X\}$ and $S^n_{++} = \{X \in \mathbb{R}^{n\times n} \mid X^\top = X, X \text{ is positive definite }\}$. Find the proximal operator of the function $f(X) = -\log\det X$ where $X \in S^n$ and $\mathrm{dom} f = S^n_{++}$. Here, the proximal operator is defined as

$$\mathrm{prox}_f(X) = \arg\min_U \left(f(U) + \frac{1}{2}\|U-X\|_F^2\right),$$

where $\|\cdot\|_F$ is the Frobenius norm.

Solution

从matrix cookbook可知:

$$\begin{gathered} \frac{\partial}{\partial X}||X||_F^2 = \frac{\partial}{\partial x}tr(XX^H) = 2x \\ \frac{\partial}{\partial X}logdetX = X^{-1} \end{gathered}$$

因此对于题目中的proximal内部求导可以得到:

$$U^{-1} + U - X = 0$$

$$I + U^2 - UX = 0$$

所以求得U为:

$$U = \frac{X \pm (X^T X - 4I)^{0.5}}{2}$$