DATA130026.01 Optimization
Assignment 12
Due time: at the beginning of the class, Jun. 15, 2021

1. Compute the projection of a given point x to the second order cone $Q = \{(x,t) \in \mathbb{R}^{n+1} \mid \|x\| \leq t\}$.

2. Show that the projection onto the set $C = [l, u] = \{x \in \mathbb{R}^n | l \leq x \leq u\}$ can be written as

$$P_C(x)_i = \begin{cases} l_i & \text{if } x_i \leq l_i, \\ x_i & \text{if } l_i \leq x_i \leq u_i, \\ u_i & \text{if } x_i \geq u_i. \end{cases}$$

3. Compute the proximal mapping prox_{tf} $(t > 0)$ for the following function f.
   - quadratic function ($A \succeq 0$): $f(x) = \frac{1}{2}x^T Ax + b^T x + c$.
   - Euclidean norm: $f(x) = \|x\|_2$.
   - logarithmic barrier: $f(x) = -\sum_{i=1}^n \log x_i$.

4. *You are not required to submit the solutions for this question. Please think about this question.

   (a) Reformulate the following problem as a semidefinite programming (SDP) problem

$$\max_{\|B\| \leq 1} A \bullet B,$$

   where $A \bullet B$ denotes the inner product of two $m \times n$ $(m \leq n)$ matrixes A and B, $\|\cdot\|$ is the spectral norm (largest singular value) of a matrix. Derive the dual problem of the SDP problem.

   (b) Let X be rank r $m \times n$ matrix with $1 \leq r \leq m \leq n$. Suppose that

$$X = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T$$

   is a singular value decomposition of X, where $\Sigma = \text{Diag}(\sigma_1(X), \ldots, \sigma_r(X)) \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing the nonzero singular values of X. Let $\|X\|_* = \sum_{i=1}^r = \sigma_i$ denote the nuclear norm of X. Show that

$$\partial \|X\|_* = \left\{ U \begin{pmatrix} I_r & 0 \\ 0 & W \end{pmatrix} V^T : \|W\| \leq 1 \right\},$$

   where I_r is the $r \times r$ identity matrix and $\|W\|$ is the spectral norm (largest singular value) of W.

   Hint: You may first prove the dual norm of the nuclear norm is the spectral norm and then use subgradient calculus rule to prove the result. Hint for proving the dual norm: you may first prove that dual norm of spectral norm "$\geq$" nuclear norm and then construct a dual problem of $\max_{\|B\| \leq 1} A \bullet B$, and show that dual norm of spectral norm "$\leq$" nuclear norm.

(c) Use the subdifferential $\partial\|X\|_*$ to compute the proximal mapping of $t\|X\|_*$.

5. Solve the following quadratic programming problem:

$$\begin{aligned} \min \quad & \frac{1}{2}x^T Q x - b^T x \\ \text{s.t.} \quad & 1 \le x_i \le 2, i = 1, \dots, n. \end{aligned}$$

Use the following Matlab code to generate the data:

```
n = 500; xbar=randn(n,1);
Q = randn(n,n);Q=Q*Q';Q=Q+Q'+eye(n);
b=Q*xbar;e=ones(n,1);
```

Choose initial point `e=ones(n,1)`. Terminate your code after 1000 iterations. Use fixed step size $1/L$, where L is the Lipschitz constant for the smooth part. Implement both the proximal gradient method and the FISTA with i) fixed step size 1/L, and ii) line search (i.e., you need to implement 4 methods). Plot the results (use $f(x_k) - f^*$ as the y-axis, where f^* can be computed by CVX).

6. Consider the Lasso problem

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \tau \|x\|_1$$

where $\tau = 1$ is a weighting parameter, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ are given data. Use the following Matlab code to generate the data:

```
m = 100; n = 500; s = 50;
A = randn(m,n);
xs = zeros(n,1); picks = randperm(n); xs(picks(1:s)) = randn(s,1);
b = A*xs;
```

Choose $x = 0$ as the starting point. Terminate your code after 1000 iterations. Use fixed step size $1/L$, where L is the Lipschitz constant for the smooth part. Implement both the proximal gradient method and the FISTA with i) fixed step size 1/L, and ii) line search (i.e., you need to implement 4 methods). Plot the results (use $f(x_k) - f^*$ as the y-axis, where f^* can be computed by CVX).