Homework 1

Zuobai Zhang
July 19, 2019

1 Convex Sets

(a) **Closed and convex sets.**

i. For any $y_1, y_2 \in A(S)$, according to the definition of $A(S)$, there exist $x_1, x_2 \in S$ satisfying $A(x_1) = y_1$ and $A(x_2) = y_2$. Since S is a convex set, then for any $t \in (0,1)$ we have $tx_1 + (1-t)x_2 \in S$. Hence, we have

$$ty_1+(1-t)y_2 = tA(x_1)+(1-t)A(x_2) = tAx_1+(1-t)Ax_2 = A(tx_1+(1-t)x_2) \in A(S), \tag{1}$$

which suggests that $A(S)$ is convex.

ii. For any $y_1, y_2 \in A^{-1}(S)$, according to the definition of $A^{-1}(S)$, there exist $x_1, x_2 \in S$ satisfying $A(y_1) = x_1$ and $A(y_2) = x_2$. Since S is a convex set, then for any $t \in (0,1)$ we have $tx_1 + (1-t)x_2 \in S$. Hence, we have

$$A(ty_1+(1-t)y_2) = tAy_1+(1-t)Ay_2 = tA(y_1)+(1-t)A(y_2) = tx_1+(1-t)x_2 \in S, \tag{2}$$

which implies $ty_1 + (1-t)y_2 \in A^{-1}(S)$. This follows the definition of convex set.

iii. The statement holds true given the fact: If f is a continuous function, then the preimage of a closed set is closed. This can be understood since the linear transformation $A(S)$ is always continuous.

iv. Consider the following case. Let

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad S = \left\{(x,y) \in \mathbb{R}^2 | y \geq e^x \right\}. \tag{3}$$

It turns out that S is a closed set, while the image of S under A is $\{0\} \times (0, +\infty)$, which is not closed.

(b) **Polyhedra.**

i. Assume that $P = \{x \in \mathbb{R}^n : Ax \preceq b, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m\}$ is a polyhedron. We will prove the closeness and convexity of set P.

Closeness. Given a point $x \notin P$, then we have $Ax \npreceq b$. There exists an integer $i \in \{1, 2, ..., m\}$, satisfying $a_i^\top x > b$ where $a_i^\top$ denotes the i-th row of A. Let $\epsilon = \frac{1}{2n \max_{j=1,2,\ldots,n}\{|a_{ij}|\}}(a_i^\top x - b)$, then for any point $x' \in O(x, \epsilon)$, we have $a_i^\top x' > b$. Thus, $Ax \npreceq b$, which suggests the complement set of P is open. This reflects the closeness of set P.

Convexity. For $x_1, x_2 \in P$, we have $Ax_1 \preceq b$ and $Ax_2 \preceq b$. Then, for any $t \in (0,1)$, we have

$$A(tx_1 + (1-t)x_2) = tAx_1 + (1-t)Ax_2 \preceq tb + (1-t)b = b. \tag{4}$$

Hence, $tx_1 + (1-t)x_2 \in P$, which demonstrates the convexity of set P.

ii. The statement tells us that a projection map takes polyhedra to polyhedra. Obviously, $Q = \{(x, y) : x \in P, y = Ax\}$ is a polyhedron. Considering the projection $(x, y) \to y$ taking Q to $A(P)$, it turns out that $A(P)$ is a polyhedron.

Bonus. Similar to (b, ii.), just consider the set $Q = \{(x, y) : x \in P, Ay = x\}$ and take the projection map $(x, y) \to y$.

(c) Let S denote $\{Z \in \mathbb{S}^n : 0 \preceq Z \preceq I,\ \mathrm{tr}(Z) = k\}$ and for any $X_1, X_2 \in S$ and $t \in (0, 1)$, it suffices to prove $0 \preceq tX_1 + (1-t)X_2 \preceq I$ and $\mathrm{tr}(tX_1 + (1-t)X_2) = k$.

Since $X_1, X_2 \in S$, we have $0 \preceq tX_1 \preceq tI$ and $0 \preceq (1-t)X_2 \preceq (1-t)I$. After adding these two inequalities, we will get the first equality. And the second equality can be derived by

$$\mathrm{tr}(tX_1 + (1-t)X_2) = t\ \mathrm{tr}(X_1) + (1-t)\,\mathrm{tr}(X_2) = tk + (1-t)k = k, \tag{5}$$

which completes the proof. Also, we can recognize that the *Fantope* is the convex hull of $\{Z \in \mathbb{S}^n : \lambda_i(Z) \in \{0, 1\}, i = 1, 2, ..., n,\ \mathrm{tr}(Z) = k\}$ where $\lambda_i(Z)$ denotes the eigenvalues of Z.

(d) Consider the polyhedron $S = \{Ax : x \geq 0\}$. If $b \in S$, then the first statement holds true. And for any $y \in \mathbb{R}^m$, we have $b^\top y = x^\top A^\top y$. If $A^\top y \geq 0$, then $b^\top y \geq 0$ since $x \geq 0$. This conflicts with the second statement.

On the other hand, if $b \notin S$, then the first statement is false.

2 Convex functions

(a) Notice that $f(x) = \min_\sigma \sum_{i=1}^{n-1} |x_{\sigma(i)} - x_{\sigma(i+1)}| = x_{\max} - x_{\min}$. Here σ is the permutation that permute x in non-decreasing or non-increasing order. Thus, for any $x_1, x_2 \in \mathbb{R}^n$ and $t \in (0,1)$, we have

$$f(tx_1+(1-t)x_2) = tx_{1\,\max} - tx_{1\,\min} + (1-t)x_{2\,\max} - (1-t)x_{2\,\min} = tf(x_1)+(1-t)f(x_2). \quad (6)$$

(b) Taking the second derivative of f gives us

$$\nabla^2 f(x) = \mathrm{diag}([\frac{1}{x_1^2}, \frac{1}{x_2^2}, ..., \frac{1}{x_n^2}]) \succ 0. \quad (7)$$

The Hessian matrix of f is positive definite, which suggests f is strictly convex.

(c) If the $(\nabla f(x) - \nabla f(y))^\top (x-y) \geq 0$ holds true, then $\nabla f(x)$ is monotonically increasing. Combining with the condition that f is twice differentiable, we conclude that $\nabla^2 f(x) \geq 0$ for any $x \in \mathrm{dom}(f)$. This satisfies the second-order characterization of convex function and indicates f is convex.

Similarly, if f is convex, then $\nabla^2 f \geq 0$, which implies that $\nabla f(x)$ is monotonically increasing. This is equivalent to say $(\nabla f(x) - \nabla f(y))^\top (x-y) \geq 0$.

(d) $f(x) = \frac{1}{x}, \quad x > 0$

(e)

(f) For any $(x_1, t_1), (x_2, t_2) \in \mathrm{dom}(g)$ and any $\alpha \in (0,1)$, we have

$$\begin{aligned}
& g(\alpha x_1 + (1-\alpha)x_2, \alpha t_1 + (1-\alpha)t_2) \\
&= (\alpha t_1 + (1-\alpha)t_2) \cdot f\left(\frac{\alpha x_1 + (1-\alpha)x_2}{\alpha t_1 + (1-\alpha)t_2}\right) \\
&= (\alpha t_1 + (1-\alpha)t_2) \cdot f\left(\frac{\alpha t_1}{\alpha t_1 + (1-\alpha)t_2}\frac{x_1}{t_1} + \frac{1-\alpha}{\alpha t_1 + (1-\alpha)t_2}\frac{x_2}{t_2}\right) \\
&\leq (\alpha t_1 + (1-\alpha)t_2) \cdot \left(\frac{\alpha}{\alpha t_1 + (1-\alpha)t_2} f(x_1) + \frac{1-\alpha}{\alpha t_1 + (1-\alpha)t_2} f(x_2)\right) \\
&= \alpha t_1 \cdot f\left(\frac{x_1}{t_1}\right) + (1-\alpha)t_2 \cdot f\left(\frac{x_2}{t_2}\right) \\
&= \alpha \cdot g(x_1, t_1) + (1-\alpha) \cdot g(x_2, t_2),
\end{aligned}$$

where the inequality is due to the convexity of f.

3 Lipschitz gradients and strong convexity

(a) **ii.** $\Rightarrow$ **i.** ∇f is Lipschitz with constant L
$\Rightarrow \|\nabla f(x) - \nabla f(y)\|_2 \leq L\|x-y\|_2 \quad \forall x, y$
$\Rightarrow (\nabla f(x) - \nabla f(y))^\top (x-y) \leq \|\nabla f(x) - \nabla f(y)\|_2 \|x-y\|_2 \leq L\|x-y\|_2^2 \quad \forall x, y$

i. $\Rightarrow$ **iii.**

iii. $\Rightarrow$ **iv.** According to the Lagrange Expansion, for any x, y, there exists a $t \in [0,1]$ such that

$$f(y) = f(x) + \nabla f(x)^\top (y-x) + \frac{1}{2}(y-x)^\top \nabla^2 f\left((1-t)x + ty\right)(y-x) \tag{8}$$

Since $\nabla^2 f(x) \preceq LI$ for all x, we have

$$\begin{aligned} f(y) &\leq f(x) + \nabla f(x)^\top (y-x) + \frac{L}{2}(y-x)^\top (y-x) \\ &= f(x) + \nabla f(x)^\top (y-x) + \frac{L}{2}\|y-x\|_2^2 \end{aligned} \tag{9}$$

iv. $\Rightarrow$ **ii.** For any $x, y \in \mathrm{dom}(f)$, we have

$$f(y) \leq f(x) + \nabla f(x)^\top (y-x) + \frac{L}{2}\|y-x\|_2^2 \tag{10}$$

and

$$f(x) \leq f(y) + \nabla f(y)^\top (x-y) + \frac{L}{2}\|x-y\|_2^2. \tag{11}$$

Adding the two inequalities, we will obtain

$$(\nabla f(x) - \nabla f(y))^\top (x-y) \leq L\|x-y\|_2^2 \tag{12}$$

(b) **i.** $\Rightarrow$ **ii.**

ii. $\Rightarrow$ **iii.**

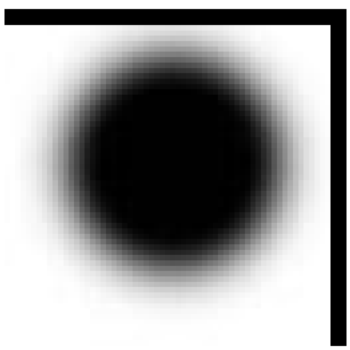
iii. $\Rightarrow$ **iv.**

iv. $\Rightarrow$ **i.**

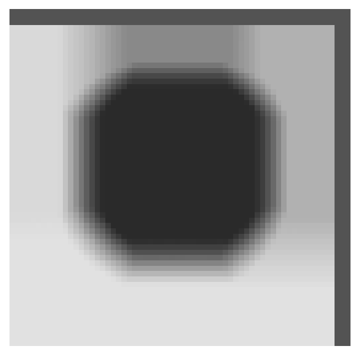
4 Solving optimization problems with CVX

(a) We have implemented the convex optimization algorithms in Julia with the package Convex.jl[1].

1. We first employ the CSV package to load the data provided in toy.csv and then use CVX to solve the problem given in the statement. The objective value obtained at the solution is 199.765. After that, we plot the original and solution data as images in Fig. 1. Observe that the shape of circle in the original image has been changed into a more angular one. This may be attributed to the penalty term in the lasso criterion function.



(a) Original Image



(b) Solution Image

Figure 1: Images for Problem (a, 1)

2. This time, we adapt the latter term in the criterion function to a 2-norm and obtain an "isotropic" total variation penalty. This gives us the optimal value 182.205 and the solution image displayed in Fig. 2.

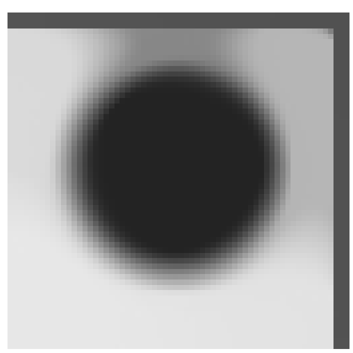


Figure 2: Solution Image for "isotropic" 2d lasso problem

[1] https://github.com/JuliaOpt/Convex.jl

It can be recognized that the solution image for 2-norm lasso looks like a circle more. In spite of this, it still has a darker background compared with the original image due to the penalty term.

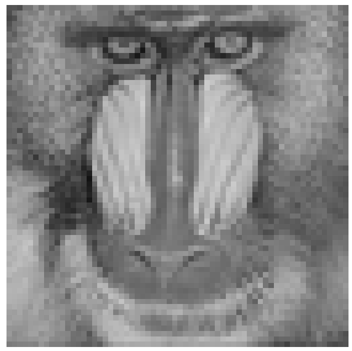


Figure 3: Original Images in baboon.csv

3.

(b) 1. Since $x > |y|$, we have $\frac{1}{x-y} > 0$. So $\frac{1}{x-y}$ and $\frac{1}{(x+y)^2}$ are both convex. Combining with an affine function z, these inequalities define a convex set. And this problem is equivalent to the following one:

$$\frac{1}{a} + \frac{1}{b^2} - z \leq 0,\ a > 0,\ b > 0.$$

Here, we let $a = x - y$ and $b = x + y$. Since $x > |y|$, we have $a = x - y > 0$ and $b = x + y > 0$. Thus, the two problems are equivalent.

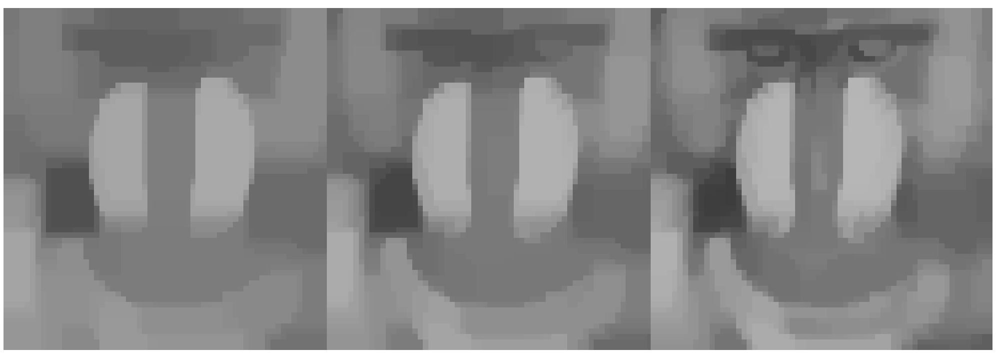
2. Here, let $y' = \log y$ and $z' = z^{1/4}$, we will obtain

$$x^2 + 2e^{-2y'} - 5z' \leq 0,\ y' > 0,\ z' \geq 0.$$

Since x^2, $e^{-2y'}$ and $-5z'$ are

Bonus.

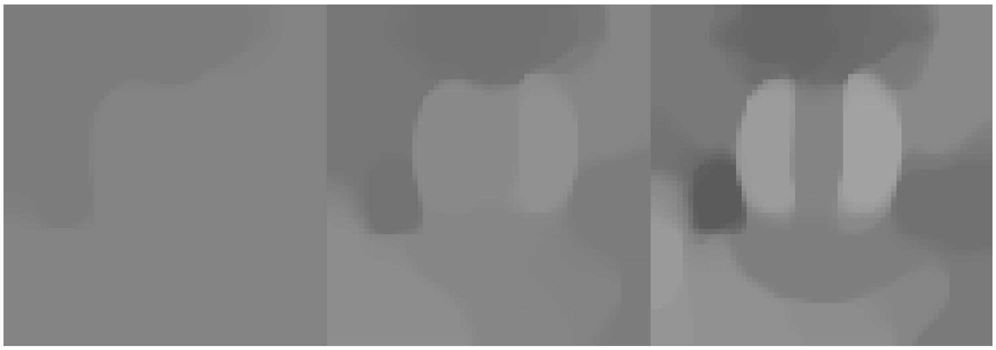
(a) $\lambda = 1, k = 0, \text{opt} = 36.6$ (b) $\lambda = 1, k = 1, \text{opt} = 34.7$ (c) $\lambda = 1, k = 2, \text{opt} = 29.2$



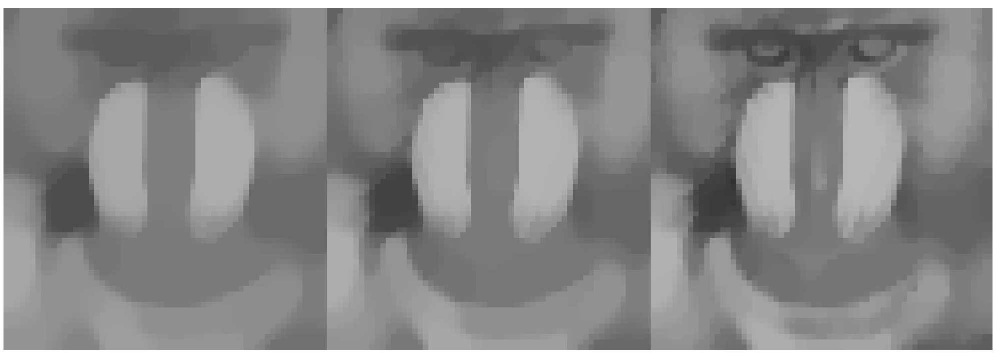
(d) $\lambda = 1, k = 3, \text{opt} = 22.6$ (e) $\lambda = 1, k = 4, \text{opt} = 16.7$ (f) $\lambda = 1, k = 5, \text{opt} = 36.6$

(g) $\lambda = 1, k = 6, \text{opt} = 11.9$ (h) $\lambda = 1, k = 7, \text{opt} = 8.20$ (i) $\lambda = 1, k = 8, \text{opt} = 3.45$

Figure 4: Images for Problem (a, 3)



(a) $\lambda = 2, k = 0, \mathrm{opt} = 36.4$ (b) $\lambda = 2, k = 1, \mathrm{opt} = 33.8$ (c) $\lambda = 2, k = 2, \mathrm{opt} = 27.7$



(d) $\lambda = 2, k = 3, \mathrm{opt} = 21.1$ (e) $\lambda = 2, k = 4, \mathrm{opt} = 15.3$ (f) $\lambda = 2, k = 5, \mathrm{opt} = 10.8$

(g) $\lambda = 2, k = 6, \mathrm{opt} = 7.26$ (h) $\lambda = 2, k = 7, \mathrm{opt} = 4.69$ (i) $\lambda = 2, k = 8, \mathrm{opt} = 2.91$

Figure 5: Images for Problem (a, 3)