Assignment 1

Convex Optimization

1 Convex sets

a. Closed sets and convex sets.

i. Show that a polyhedron $\{x \in \mathbb{R}^n : Ax \leq b\}$, for some $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$, is both convex and closed.

ii. Show that if $S_i \subseteq \mathbb{R}^n$, $i \in I$ is a collection of convex sets, then their intersection $\cap_{i \in I} S_i$ is also convex. Show that the same statement holds if we replace "convex" with "closed".

iii. Given an example of a closed set in $\mathbb{R}^2$ whose convex hull is not closed.

iv. Let $A \in \mathbb{R}^{m \times n}$. Show that if $S \subseteq \mathbb{R}^m$ is convex then so is $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$, which is called the preimage of S under the map $A : \mathbb{R}^n \to \mathbb{R}^m$. Show that the same statement holds if we replace "convex" with "closed".

v. Let $A \in \mathbb{R}^{m \times n}$. Show that if $S \subseteq \mathbb{R}^n$ is convex then so is $A(S) = \{Ax : x \in S\}$, called the image of S under A.

vi. Give an example of a matrix $A \in \mathbb{R}^{m \times n}$ and a set $S \subseteq \mathbb{R}^n$ that is closed and convex but such that $A(S)$ is not closed.

(b. Polyhedra.

i. Show that if $P \subseteq \mathbb{R}^n$ is a polyhedron, and $A \in \mathbb{R}^{m \times n}$, then $A(P)$ is a polyhedron. Hint: you may use the fact that

$$P \subseteq \mathbb{R}^{m+n} \text{ is a polyhedron } \Rightarrow \{x \in \mathbb{R}^n : (x, y) \in P \text{ for some } y \in \mathbb{R}^m\} \text{ is a polyhedron.}$$

ii. Show that if $Q \subseteq \mathbb{R}^m$ is a polyhedron, and $A \in \mathbb{R}^{m \times n}$, then $A^{-1}(Q)$ is a polyhedron.

2 Convex functions

(a.) Prove that the *entropy function*, defined as

$$f(x) = -\sum_{i=1}^{n} x_i \log(x_i),$$

with $\mathrm{dom}(f) = \{x \in \mathbb{R}^n_{++} : \sum_{i=1}^n x_i = 1\}$, is strictly concave.

(b.) Let f be twice differentiable, with $\mathrm{dom}(f)$ convex. Prove that f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0,$$

for all x, y. This property is called *monotonicity* of the gradient ∇f.

(c.) Give an example of a strictly convex function that does not attain its infimum.

(d.) A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *coercive* provided that $f(x) \to \infty$ as $\|x\|_2 \to \infty$. A key fact about coercive functions is that they attain their infimums. Prove that a twice differentiable, strongly convex function is coercive and hence attains its infimum.

(e.) Prove that the maximum of a convex function over a bounded polyhedron must occur at one of the vertices. Hint: you may use the fact that a bounded polyhedron can be represented as the convex hull of its vertices.

3 Partial optimization with ℓ_2 penalties

Consider the problem

$$\min_{\beta,\, \sigma \geq 0} \; f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} g(\beta_i, \sigma_i), \tag{1}$$

for some convex f with domain $\mathbb{R}^n$, $\lambda \geq 0$, and

$$g(x, y) = \begin{cases} x^2/y + y & \text{if } y > 0 \\ 0 & \text{if } x = 0,\ y = 0 \\ \infty & \text{else.} \end{cases}$$

In other words, the problem (1) is just the weighted ℓ_2 penalized problem

$$\min_{\beta,\, \sigma \geq 0} \; f(\beta) + \frac{\lambda}{2} \sum_{i=1}^{n} \Big(\frac{\beta_i^2}{\sigma_i} + \sigma_i \Big),$$

but being careful to treat the ith term in the sum as zero when $\beta_i = \sigma_i = 0$.

(a.) Prove that g is convex. Hence argue that (1) is a convex problem. Note that this means we can perform partial optimization in (1) and expect it to return another convex problem. Use the definition of convexity.

(b.) Argue that $\min_{y \geq 0} g(x, y) = 2|x|$.

(c.) Argue that minimizing over $\sigma \geq 0$ in (1) gives the ℓ_1 penalized problem

$$\min_{\beta} \; f(\beta) + \lambda \|\beta\|_1.$$

4 Lipschitz gradients and strong convexity

Let f be convex and twice continuously differentiable.

(a.) Show that the following statements are equivalent.

i. ∇f is Lipschitz with constant L;

ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \leq L\|x-y\|_2^2$ for all x, y;

iii. $\nabla^2 f(x) \preceq LI$ for all x;

iv. $f(y) \leq f(x) + \nabla f(x)^T(y-x) + \frac{L}{2}\|y-x\|_2^2$ for all x, y.

Your solution should have 5 parts, where you prove i $\Rightarrow$ ii, ii $\Rightarrow$ iii, iii $\Rightarrow$ iv, iv $\Rightarrow$ ii, and iii $\Rightarrow$ i.

(b.) Show that the following statements are equivalent.

i. f is strongly convex with constant m;

ii. $(\nabla f(x) - \nabla f(y))^T(x-y) \geq m\|x-y\|_2^2$ for all x, y;

iii. $\nabla^2 f(x) \succeq mI$ for all x;

iv. $f(y) \geq f(x) + \nabla f(x)^T(y-x) + \frac{m}{2}\|y-x\|_2^2$ for all x, y.

Your solution should have 4 parts, where you prove i $\Rightarrow$ ii, ii $\Rightarrow$ iii, iii $\Rightarrow$ iv, and iv $\Rightarrow$ i.