NUMERICAL OPTIMIZATION

Sheet 4: Convexity continued

Exercise one We will continue our work with $g(x_1, x_2, \ldots, x_n) = \log(\sum_{i=1}^{n} e^{x_i})$, trying to prove its convexity. Let H denote its Hessian. First, we try to use a direct approach to show that H is positive semidefinite.

1. Let A be a symmetric matrix and let c^2 be some non-negative constant or parameter (in this case, it must be non-negative because it is a square). Show that A/c^2 is positive semidefinite if and only if A is positive semidefinite.
2. What is a good choice for c^2 to simplify the Hessian matrix H?
3. After dividing by c^2, try to use the Sylvester criterion (checking determinants of principal minors) to check positive semidefiniteness of H. To make it simple for you, I only ask you to show it for a 3×3 version, i.e., with only the vector (x_1, x_2, x_3).

~~**Exercise two**~~ Let us try to give a different argument for the positive semidefiniteness of H.

For this exercise, we will make good use of the notation

$$s_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}.$$

1. Express both the gradient and the Hessian of $g(x_1, x_2, \ldots, x_n) = \log(\sum_{i=1}^{n} e^{x_i})$ using only the substituted variables $\vec{s} = s_1, \ldots, s_n$.
2. Next, write the Hessian H as a difference of two matrices $H = D - K$, where D is a diagonal matrix (a matrix with non-zero elements only on the diagonal). Can you also write K as a product of some vectors like $\vec{s}$?
3. Finally, let us try to prove the positive semidefiniteness directly. That means that we wish to check that for any non-zero vector y, we have $y^T H y \geq 0$. This means $y^T(D - K)y \geq 0$ needs to be proven. Try to prove that inequality.

We already know plenty of convex sets – all linear subspaces, all affine subspaces including hyperplanes, halfspaces, polyhedra (finite intersections of halfspaces), open and closed balls of any dimensions.

We also know a basic rule that convexity is preserved on arbitrary intersections.

Let us learn a few more:

- image and counterimage of convex set by affine function is convex
- in particular scaling and translation of convex set is convex
- Cartesian product of convex sets is convex
- complex sum $A + B = \{x + y : x \in A, y \in B\}$ is convex for convex A and B

Some examples (borrowing from previous year's lecture notes):

Example: When A is a fixed linear operator and a is a fixed vector than

$$\{x : Ax = a\}$$

is a convex set as counterimage of one point set.

Example: Set of $\{x \in \mathbb{R}^n | \vec{x} \geq \vec{0}\}$ is a convex set as it is the Cartesian product of convex sets $\{x_i : x_i \geq 0\} \subset \mathbb{R}$.

Example: For fixed $b \in \mathbb{R}^n$ the set

$$\{x : x \geq b\}$$

is convex as translation of the set from the previous example.

Example: When B is a linear operator, and b is a fixed vector, then the set

$$\{x : Bx \geq b\}$$

is a convex set as counterimage of convex set from previous example.

EXERCISE THREE An *ellipsoid*, a multidimensional analogue of the ellipse, can be formally defined in multiple dimensions as follows:

Given a center a, and a positive definite matrix A, the ellipsoid $E(a, A)$ is defined as $\{x \in \mathbb{R}^n : (x-a)^T A^{-1}(x -$

Prove that this object is always convex. $\{x \in \mathbb{R}^n \mid (x-a)^T A^{-1} (x-a) \leq 1\}$

Hint: You can start from the fact that $\{x \in \mathbb{R}^n : (x-a)^T(x-a) \leq 1\}$ is a ball and hence convex.

EXERCISE FOUR

Prove or disprove the convexity of the following sets:

1. A slab, i.e., a set of the form $\{x \in \mathbb{R}^n : \alpha \leq \langle a, x\rangle \leq \beta\}$.
2. A wedge, i.e., $\{x \in \mathbb{R}^n : \langle a_1, x\rangle \leq b_1, \langle a_2, x\rangle \leq b_2\}$.
3. The set of points closer to a given point than a given set, i.e.,

$$\{x : \|x - x_0\|_2 \leq \|x - y\|_2 \text{ for all } y \in S\}, \qquad \text{where } S \subset \mathbb{R}^n.$$

EXERCISE FIVE Prove or disprove the convexity of the following sets:

1. The set of points closer to one set than another, i.e., $\{x : \operatorname{dist}(x, S) \leq \operatorname{dist}(x, T)\}$, where $S, T \subset \mathbb{R}^n$, and $\operatorname{dist}(x, S) = \inf_{z \in S} \|xz\|_2$.
2. The set $\{x : x + S_1 \subset S_2\}$, where $S_1, S_2 \subset \mathbb{R}^n$ with S_2 convex.

EXERCISE SIX Prove the convexity of the following set:

1. The set of points whose distance to a does not exceed a fixed fraction θ of the distance to b, i.e., the set $\{x : \|x - a\|_2 \leq \theta\|x - b\|_2\}$. You can assume $a \neq b$ and $0 \leq \theta \leq 1$.