MATH-329 Nonlinear optimization
Exercise session 8: Convex optimization

Instructor: Nicolas Boumal
TAs: Guifré Sánchez, Antoine Gonon

Document compiled on November 4, 2025

1. Sufficient optimality conditions. Mind the following: in Theorem 9.2, it is important to check *all three* of the following boxes:

(a) We are *minimizing* (and not maximizing).

(b) The cost function f is convex.

(c) The search space S is convex.

To really appreciate this fact, do the following:

1. Give examples of optimization problems which check two of the above but not all three of the above, and for which there exists a non-optimal stationary point.
2. If given a maximization problem, explain how you can get an equivalent minimization problem.
3. If the cost function is not convex, explain how you can get an equivalent problem with a convex cost function (you can even make the cost function linear).
4. If the constraint set is not convex, explain how you can get an equivalent problem with a convex search space (you can even make the problem unconstrained)—you will need to make the cost function a bit weird for this though (hint: "indicator function" with values in $\{0, \infty\}$).

For each of the above, explain how we should understand Theorem 9.2 against the modified problem, specifically to verify that, sadly, there is no free lunch.

2. Discontinuous projection. Show with a drawing that Proj_S may be discontinuous if S is non-empty and closed but fails to be convex. This reveals why the PGD iteration map $x \mapsto \mathrm{Proj}_S(x - \alpha \nabla f(x))$ could be discontinuous if S is not convex. It would be much harder to analyze the algorithm if we allowed that to happen.

3. Trust-region subproblem. Let $S = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$ be the unit norm ball. Give an expression for Proj_S. Given a symmetric matrix A of size n and $b \in \mathbb{R}^n$, let $f(x) = \frac{1}{2}x^\top Ax + b^\top x$. What is the Lipschitz constant of ∇f? Can you easily compute an upper bound for it? Implement projected gradient descent (see lecture notes) for this problem, with a proper choice of step-size. Do you expect that this method would converge to a global minimizer? Note: this is a fairly terrible algorithm for the trust-region subproblem but it has the merit of being simple.

4. Image and inverse image of affine function. Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be an affine function, that is, there exists $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ such that $f(x) = Ax + b$ for all $x \in \mathbb{R}^n$.

1. Let $S \subseteq \mathbb{R}^n$ be a convex set. Show that the image of S under f,

$$f(S) = \{f(x) \mid x \in S\},$$

   is convex.

2. Let $S \subseteq \mathbb{R}^m$ be a convex set. Show that the inverse image of S under f,

$$f^{-1}(S) = \{x \in \mathbb{R}^n \mid f(x) \in S\},$$

   is convex.

3. Let $g\colon \mathbb{R}^m \to \mathbb{R}^n$ be a convex function. Show that $g \circ f$ is convex.

Supplementary exercises

1. Convex combination. Let $C \subseteq \mathbb{R}^n$ be a convex set, $x_1, \dots, x_k \in C$ and $\theta_1, \dots, \theta_k \geq 0$ be non-negative coefficients such that $\theta_1 + \cdots + \theta_k = 1$. Show that the convex combination $\theta_1 x_1 + \cdots + \theta_k x_k$ is in C.

2. Intersection with a line. Show that a set is convex if and only if its intersection with any line is convex.

3. Sublevel sets. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a convex function. Show that for all $\alpha \in \mathbb{R}$ the sublevel set $\{x \in \mathbb{R}^n \mid f(x) \leq \alpha\}$ is convex.