Optimization of functionals

Department of mathematics, Cinvestav-IPN

Problem set 2. Lower semicontinuous functions

DUE: October 3, 2024

1. Let X be a finite-dimensional vector space. Suppose $f : X \to (-\infty, +\infty]$ is lsc, bounded from below, and not identically $+\infty$. Let $\varepsilon > 0$ and $x_\varepsilon \in X$ sstisfy $f(x_\varepsilon) \leq \inf f + \varepsilon$. Prove (without using Ekeland variational principle!) that there exists $\overline{x} \in X$ such that

   (a) $f(\overline{x}) \leq f(x_\varepsilon)$,

   (b) $\|\overline{x} - x_\varepsilon\| \leq \sqrt{\varepsilon}$, and

   (c) $f(\overline{x}) \leq f(x) + \sqrt{\varepsilon}\|x - \overline{x}\|$ for all $x \in X$.

   Hint: Show that $x \mapsto f(x) + \sqrt{\varepsilon}\|x - \overline{x}\|$ is coercive.

2. Let $f, g : S \to \mathbb{R}$ be convex functions, where $S \subseteq \mathbb{R}^n$ is convex. Show the following:

   (a) If c is a nonnegative real number, then $f + cg$ is convex.

   (b) If $F : \mathbb{R} \to \mathbb{R}$ is convex and increasing, then $F \circ f$ is convex.

   (c) The set of minimizers (which could be empty) of any convex function is convex. Prove also that strictly convex functions have at most one global minimizer.

   (d) If f is differentiable and $\nabla f(x^*) = 0$, then $f(x^*) \leq f(x)$ for every $x \in S$.

3. **(Least squares)** Let $A \in \mathcal{M}_{m\times n}$, with $m > n$, and $b \in \mathbb{R}^m$. The system $Ax = b$ usually does not have a solution $x \in \mathbb{R}^n$, then an alternative is to find the *least-squares solution* $\hat{x}$—if it exists—, that is,

   $$|A\hat{x} - b|^2 = \min_{x\in\mathbb{R}^n} |Ax - b|^2.$$

   Assume $\operatorname{rank}(A) = n$ and prove that there exists a unique global minimizer $\hat{x}$, given by

   $$\hat{x} = (A^\top A)^{-1} A^\top b.$$

4. (**The Fundamental Theorem of Algebra**) Let $p(z) = a_n z^n + \ldots + a_1 z + a_0$ be a polynomial with complex coefficients, $a_n \neq 0$ and $n \geq 2$. Define the function $f(z) := |p(z)|$ for each $z \in \mathbb{C}$.

   (a) Prove that f has a global minimizer.

   Hint: Show that f is coercive.

   (b) Find explicitly one (there could be more) global minimizer of f when (i) $p(z) = a_1 z + \ldots + a_n z^n$, that is $a_0 = 0$, and (ii) $p(z) = a_0 + a_n z^n$.

(c) Let $z_0 \in \mathbb{C}$. Explain why there exist complex numbers $c_0, c_1, \ldots, c_n$ such that

$$p(z) = c_0 + c_1(z - z_0) + \ldots + c_n(z - z_0)^n.$$

Hint: Write $p(z) = p((z - z_0) + z_0)$.

Further, prove that, for some $k = 1, \ldots, n$,

$$p(z) = c_0 + c_k(z - z_0)^k + (z - z_0)^{k+1} q(z),$$

where $c_k \neq 0$ and q is a polynomial.

(d) Let z_0 be a global minimizer of f, $t \in (0, 1)$, and $w \in \mathbb{C}$ satisfies $c_0 + c_k w^k = 0$. Suppose $f(z_0) > 0$, that is, $c_0 \neq 0$. Show that

$$f(z_0 + tw) \leq |c_0|(1 - t^k) + |tw|^{k+1}|q(z_0 + tw)|$$

and

$$t|w^{k+1} q(z_0 + tw)| < |c_0|$$

for some t small enough.

(e) Prove the Fundamental Theorem of Algebra.

5. Let $\overline{x}$ be a local minimizer of f subject to $g(x) = 0$, where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are differentiable, $m < n$. Assume that $\nabla g_1(\overline{x}), \ldots, \nabla g_m(\overline{x})$ are linearly independent. Use the Fritz John's conditions to prove the existence of multipliers $\lambda_1, \ldots, \lambda_m$ such that

$$\nabla f(\overline{x}) + \lambda_1 \nabla g_1(\overline{x}) + \ldots + \lambda_n \nabla g_n(\overline{x}) = 0.$$

6. Let $a \in \mathbb{R}^n$, $a \neq 0$. Find the unique solution to the problem

$$\max_{x \in \mathbb{R}^n} \{a^\top x : |x|^2 = 1\}.$$

Compare with Cauchy-Schwarz inequality.

7. Let $a \in \mathbb{R}$, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^k$. Assume $a \geq 0$ and $c \geq 0$. Find the unique solution to the problem

$$\max_{(\lambda_0, \lambda, \mu)} \{a\lambda_0 + b^\top \lambda + c^\top \mu : |(\lambda_0, \lambda, \mu)| = 1,\ \lambda_0 \geq 0,\ \mu \geq 0\}.$$

8. **(Spectral theorem).** Let $A \in \mathcal{M}_n(\mathbb{R})$ be a symmetric matrix.

(a) Use Lagrange multipliers to show that there exists $\lambda_1 \in \mathbb{R}$ and $u_1 \in \mathbb{R}^n$, $|u_1| = 1$, such that

$$Au_1 = \lambda_1 u_1$$

and

$$x \in \mathbb{R}^n,\ |x| = 1 \quad \Rightarrow \quad x^\top A x \geq \lambda_1.$$

(b) Show that λ_1 is the smallest eigenvalue of A.

(c) Show that exists $\lambda_2 \in \mathbb{R}$ and $u_2 \in \mathbb{R}^n$, $|u_2| = 1$, such that

$$Au_2 = \lambda_2 u_2$$

and

$$u_2^\top u_1 = 0.$$

Hint: Consider $W_1 = \{x \in \mathbb{R}^n \mid x^\top u_1 = 0\}$, *verify that* $Ax \in W_1$ *for every* $x \in W_1$, *and find a minimizer* u_2 *of* $x^\top Ax$ *in some compact subset of* W_1.

(d) Prove that there exist an orthonormal basis $\{u_1, \dots, u_n\}$ of $\mathbb{R}^n$ and a vector $(\lambda_1, \dots, \lambda_n)^\top$ such that

$$Au_j = \lambda_j u_j, \quad 1 \le j \le n.$$

9. Let $A \in \mathcal{M}_n(\mathbb{R})$ be a symmetric matrix. Use the Spectral theorem to show the following:

(a) Let $(\lambda_1, \dots, \lambda_n)^\top$ be as in Exercise 8(d). Show that

$$\mathrm{tr}(A) := \sum_{j=1}^{n} A_{jj} = \sum_{j=1}^{n} \lambda_j.$$

Hint: Show first that $AU = U\Lambda$, *where* Λ *is diagonal.*

(b) A is positive semidefinite if and only if its eigenvalues are nonnegative.

(c) A is positive definite if and only if its eigenvalues are positive.

10. Let $h : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function. Consider the problem

$$\inf\{h(x) \mid x \in \mathbb{R}^n\}.$$

In numerical analysis, a vector $v \in \mathbb{R}^n \setminus \{0\}$ is said to be a *descent direction* of f at a if $D_v h(a) < 0$. If $\nabla h(a) \neq 0$, then $-\nabla h(a)$ is called the *steepest-descent direction* of h at a. Justify these names by proving the following:

(a) If $D_v h(a) < 0$, then there exists $t_0 > 0$ such that

$$h(a + tv) < h(a) \qquad \forall t \in (0, t_0].$$

(b) There exists a solution to

$$\min_{v \in \mathbb{R}^n}\{D_v h(a) : |v| = 1\},$$

and such a solution is given by $-Dh(a)/|Dh(a)|$.