Optimization Theory

Tutorial 7

Wang Xia

2018/3/5

Table of Contents

Table of Contents

Existence of Optimal Solutions

The set of **minima** of a real-valued function f over a nonempty set X, call is X^*, is equal to the intersection of X and the level sets of f that have a common points with X :

$$X^* = \cap_{k=0}^{\infty}\{x \in X | f(x) \leq \gamma_k\},$$

where $\{\gamma_k\}$ is any scalar sequence with $\gamma_k \downarrow inf_{x \in X} f(x)$.

Existence of Optimal Solutions

Theorem

Weierstrass' Theorem *Consider a closed proper function*

$$f \to (-\infty, \infty],$$

and assume that any one of the following three conditions holds:

(1) $dom(f)$ *is bounded.*

(2) *There exists a scalar* $\bar{\gamma}$ *such that the level set*

$$\{x | f(x) \leq \bar{\gamma}\}$$

is nonempty and bounded.

(3) f *is coercive.*

Then the set of minima of f *over* $\Re^n$ *is nonempty and compact.*

Partial Minimization of convex functions

Theorem

Consider a function $F : \Re^{n+m} \to (-\infty, \infty]$ *and the function* $f : \Re^n \to [-\infty, \infty]$ *defined by*

$$f(x) = inf_{z \in \Re^m F(x,z)}.$$

Then:

(a) *If* F *is convex, then* f *is also convex.*

(b) *We have*

$$P(epi(F)) \subset epi(f) \subset cl(P(epi(F))),$$

where $P(\cdot)$ *denotes projection on the space of* (x, w)*, i.e., for any subset* S *of* $\Re^{n+m+1}$*,* $P(S) = (x, w)|(x, z, w) \in S$.

Saddle Point and Minimax Theory

Theorem

Saddle Point: *A pair of vectors* $x^* \in X$ *and* $z^* \in Z$ *is called a saddle point of* ϕ *if*

$$\phi(x^*, z) \leq \phi(x^*, z^*) \leq \phi(x, z^*), \forall x \in X, \forall z \in Z.$$

minimax equality:

$$sup_{z\in Z} inf_{x\in X} \phi(x,z) = inf_{x\in X} sup_{z\in Z} \phi(x,z).$$

Saddle Point and Minimax Theory

Theorem

A pair (x^*, z^*) is a saddle point of ϕ if and only if the minimax equality holds, and x^* is an optimal solution of the problem:

$$\text{minimize} \quad \sup_{z \in Z} \phi(x, z), \text{subject to} \quad x \in X,$$

while z^* is an optimal solution of the problem

$$\text{maximize} \quad \inf_{x \in X} \phi(x, z), \text{subject to} \quad z \in Z$$

Saddle Point and Minimax Theory

Lemma 2.6.1: Let X be a nonempty convex subset of $\Re^n$, let Z be a nonempty subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $\phi(\cdot, z) : X \mapsto \Re$ is convex. Then the function p of Eq. (2.33) is convex.

Saddle Point and Minimax Theory

Lemma 2.6.2: Let X be a nonempty subset of $\Re^n$, let Z be a nonempty convex subset of $\Re^m$, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $x \in X$, the function $-\phi(x, \cdot) : Z \mapsto \Re$ is closed and convex. Then the function $q : \Re^m \mapsto [-\infty, \infty]$ given by

$$q(\mu) = \inf_{(u,w)\in \text{epi}(p)} \{w + u'\mu\}, \qquad \mu \in \Re^m,$$

where p is given by Eq. (2.33), satisfies

$$q(\mu) = \begin{cases} \inf_{x\in X} \phi(x, \mu) & \text{if } \mu \in Z, \\ -\infty & \text{if } \mu \notin Z. \end{cases} \tag{2.36}$$

Furthermore, we have $q^* = w^*$ if and only if the minimax equality (2.26) holds.

Saddle Point and Minimax Theory

Proposition 2.6.2: (Minimax Theorem I) Let X and Z be nonempty convex subsets of $\Re^n$ and $\Re^m$, respectively, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $\phi(\cdot, z) : X \mapsto \Re$ is convex, and for each $x \in X$, the function $-\phi(x, \cdot) : Z \mapsto \Re$ is closed and convex. Assume further that

$$\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty.$$

Then, the minimax equality holds, i.e.,

$$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z),$$

if and only if the function p of Eq. (2.33) is lower semicontinuous at $u = 0$, i.e., $p(0) \leq \liminf_{k \to \infty} p(u_k)$ for all sequences $\{u_k\}$ with $u_k \to 0$.

Saddle Point and Minimax Theory

Proposition 2.6.3: (Minimax Theorem II) Let X and Z be nonempty convex subsets of $\Re^n$ and $\Re^m$, respectively, and let $\phi : X \times Z \mapsto \Re$ be a function. Assume that for each $z \in Z$, the function $\phi(\cdot, z) : X \mapsto \Re$ is convex, and for each $x \in X$, the function $-\phi(x, \cdot) : Z \mapsto \Re$ is closed and convex. Assume further that

$$-\infty < \inf_{x \in X} \sup_{z \in Z} \phi(x, z),$$

Saddle Point and Minimax Theory

and that 0 lies in the relative interior of the effective domain of the function p of Eq. (2.33). Then, the minimax equality holds, i.e.,

$$\sup_{z\in Z} \inf_{x\in X} \phi(x,z) = \inf_{x\in X} \sup_{z\in Z} \phi(x,z),$$

and the supremum over Z in the left-hand side is finite and is attained. Furthermore, the set of $z \in Z$ attaining this supremum is compact if and only if 0 lies in the interior of the effective domain of p.

Table of Contents

EX 1

Saddle Points in Two Dimensions

Consider a function ϕ of two real variables x and z taking values in compact intervals of X and Z, respectively. Assume that for each $z \in Z$, the function $\phi(\cdot, z)$ is minimized over X at a unique point denoted $\hat{x}(z)$. Similarly, assume that for each $x \in X$, the function $\phi(x, \cdot)$ is maximized over Z at a unique point denoted $\hat{z}(x)$. Assume further that the functions $\hat{x}(z)$ and $\hat{z}(x)$ are continuous over Z and X, respectively. Show that ϕ has a saddle point (x^*, z^*). Use this to investigate the existence of saddle points of $\phi(x, z) = x^2 + z^2$ over $X = [0, 1]$ and $Z = [0, 1]$.

Solution 1

We consider a function ϕ of two real variables x and z taking values in compact intervals X and Z, respectively. We assume that for each $z \in Z$, the function $\phi(\cdot, z)$ is minimized over X at a unique point denoted $\hat{x}(z)$, and for each $x \in X$, the function $\phi(x, \cdot)$ is maximized over Z at a unique point denoted $\hat{z}(x)$,

$$\hat{x}(z) = \arg\min_{x \in X} \phi(x, z), \qquad \hat{z}(x) = \arg\max_{z \in Z} \phi(x, z).$$

Consider the composite function $f : X \mapsto X$ given by

$$f(x) = \hat{x}\big(\hat{z}(x)\big),$$

which is a continuous function in view of the assumption that the functions $\hat{x}(z)$ and $\hat{z}(x)$ are continuous over Z and X, respectively. Assume that the compact interval X is given by $[a, b]$. We now show that the function f has a fixed point, i.e., there exists some $x^* \in [a, b]$ such that

$$f(x^*) = x^*.$$

Solution 1

Define the function $g : X \mapsto X$ by

$$g(x) = f(x) - x.$$

Assume that $f(a) > a$ and $f(b) < b$, since otherwise we are done. We have

$$g(a) = f(a) - a > 0,$$

$$g(b) = f(b) - b < 0.$$

Since g is a continuous function, the preceding relations imply that there exists some $x^* \in (a, b)$ such that $g(x^*) = 0$, i.e., $f(x^*) = x^*$. Hence, we have

$$\hat{x}\big(\hat{z}(x^*)\big) = x^*.$$

Solution 1

Denoting $\hat{z}(x^*)$ by z^*, we get

$$x^* = \hat{x}(z^*), \qquad z^* = \hat{z}(x^*). \tag{2.24}$$

By definition, a pair $(\overline{x}, \overline{z})$ is a saddle point if and only if

$$\max_{z \in Z} \phi(\overline{x}, z) = \phi(\overline{x}, \overline{z}) = \min_{x \in X} \phi(x, \overline{z}),$$

or equivalently, if $\overline{x} = \hat{x}(\overline{z})$ and $\overline{z} = \hat{z}(\overline{x})$. Therefore, from Eq. (2.24), we see that (x^*, z^*) is a saddle point of ϕ.

We now consider the function $\phi(x, z) = x^2 + z^2$ over $X = [0, 1]$ and $Z = [0, 1]$. For each $z \in [0, 1]$, the function $\phi(\cdot, z)$ is minimized over $[0, 1]$ at a unique point $\hat{x}(z) = 0$, and for each $x \in [0, 1]$, the function $\phi(x, \cdot)$ is maximized over $[0, 1]$ at a unique point $\hat{z}(x) = 1$. These two curves intersect at $(x^*, z^*) = (0, 1)$, which is the unique saddle point of ϕ.

Ex 2

Saddle Points of Quadratic Functions

Consider a quadratic function $\phi : X \times Z \to \Re$ of the form

$$\phi(x, z) = x'Qx + x'Dz - z'Rz,$$

where Q and R are symmetric positive semidefinite $n \times n$ and $m \times m$ matrices, respectively, D is some $n \times m$ matrix, and X and Z are subsets of $\Re^n$ and $\Re^m$, respectively. Derive conditions under which ϕ has at least one saddle point.

Solution 2

Let X and Z be closed and convex sets. Then, for each $z \in Z$, the function $t_z : \Re^n \mapsto (-\infty, \infty]$ defined by

$$t_z(x) = \begin{cases} \phi(x,z) & \text{if } x \in X, \\ \infty & \text{otherwise,} \end{cases}$$

is closed and convex in view of the assumption that Q is a positive semidefinite symmetric matrix. Similarly, for each $x \in X$, the function $r_x : \Re^m \mapsto (-\infty, \infty]$ defined by

$$r_x(z) = \begin{cases} -\phi(x,z) & \text{if } z \in Z, \\ \infty & \text{otherwise,} \end{cases}$$

is closed and convex in view of the assumption that R is a positive semidefinite symmetric matrix. Hence, Assumption 2.6.1 is satisfied. Let also Assumptions 2.6.2 and 2.6.3 hold, i.e,

$$\inf_{x \in X} \sup_{z \in Z} \phi(x,z) < \infty,$$

Solution 2

and

$$-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z).$$

By the positive semidefiniteness of Q, it can be seen that, for each $z \in Z$, the recession cone of the function t_z is given by

$$R_{t_z} = R_X \cap N(Q) \cap \{y \mid y' Dz \leq 0\},$$

where R_X is the recession cone of the convex set X and $N(Q)$ is the null space of the matrix Q. Similarly, for each $z \in Z$, the constancy space of the function t_z is given by

$$L_{t_z} = L_X \cap N(Q) \cap \{y \mid y' Dz = 0\},$$

where L_X is the lineality space of the set X. By the positive semidefiniteness of R, for each $x \in X$, it can be seen that the recession cone of the function r_x is given by

$$R_{r_x} = R_Z \cap N(R) \cap \{y \mid x' Dy \geq 0\},$$

Solution 2

where R_Z is the recession cone of the convex set Z and $N(R)$ is the null space of the matrix R. Similarly, for each $x \in X$, the constancy space of the function r_x is given by

$$L_{r_x} = L_Z \cap N(R) \cap \{y \mid x'Dy = 0\},$$

where L_Z is the lineality space of the set Z.

If

$$\bigcap_{z \in Z} R_{t_z} = \{0\}, \quad \text{and} \quad \bigcap_{x \in X} R_{r_x} = \{0\}, \tag{2.25}$$

then it follows from the Saddle Point Theorem part (a), that the set of saddle points of ϕ is nonempty and compact. [In particular, the condition given in Eq. (2.25) holds when Q and R are positive definite matrices, or if X and Z are compact.]

Similarly, if

$$\bigcap_{z \in Z} R_{t_z} = \bigcap_{z \in Z} L_{t_z}, \quad \text{and} \quad \bigcap_{x \in X} R_{r_x} = \bigcap_{x \in X} L_{r_x},$$

then it follows from the Saddle Point Theorem part (b), that the set of saddle points of ϕ is nonempty.

Ex 3

Convex-concave functions and saddle points

We say the function $f : \Re^n \times \Re^m \to \Re$ is convex-concave if $f(x, z)$ is a concave function of z, for each fixed x, and a convex function of x, for each fixed z. We also require its domain to have the product form $\mathbf{dom} f = A \times B$, where $A \subset \Re^n$ and $B \subset \Re^m$ are convex.

(a) Give a second-order condition for a twice differentiable function $f : \Re^n \times \Re^m \to \Re$ to be convex-concave, interms of its Hessian $\nabla^2 f(x, z)$.

(b) Suppose that $f : \Re^n \times \Re^m \to \Re$ is convex-concave and differentiable, with $\nabla f(\hat{x}, \hat{z}) = 0$. Show that the saddle point property holds: for all x, z, we have

$$f(\hat{x}, z \leq f(\hat{x}, \hat{z})) \leq f(x, \hat{z}).$$

Show that this implies that f satisfies the strong max-min property:

$$sup_z inf_x f(x,z) = inf_x sup_z f(x,z)$$

(and their common value is $f(\hat{x}, \hat{z})$).

(c) Now suppose that $f : \Re^n \times \Re^m \to \Re$ is differentiable, but not necessarily convex-concave, and the saddle-point property holds at $\hat{x}$, $\hat{z}$:

$$f(\hat{x}, z \leq f(\hat{x}, \hat{z})) \leq f(x, \hat{z}).$$

for all x, z. Show that $\nabla f(\hat{x}, \hat{z}) = 0$.

Solution 3

(a) The condition follows directly from the second-order conditions for convexity and concavity: it is

$$\nabla^2_{xx} f(x,z) \succeq 0, \qquad \nabla^2_{zz} f(x,z) \preceq 0,$$

for all x, z. In terms of $\nabla^2 f$, this means that its $1,1$ block is positive semidefinite, and its $2,2$ block is negative semidefinite.

(b) Let us fix $\tilde{z}$. Since $\nabla_x f(\tilde{x}, \tilde{z}) = 0$ and $f(x, \tilde{z})$ is convex in x, we conclude that $\tilde{x}$ minimizes $f(x, \tilde{z})$ over x, *i.e.*, for all z, we have

$$f(\tilde{x}, \tilde{z}) \le f(x, \tilde{z}).$$

This is one of the inequalities in the saddle-point condition. We can argue in the same way about $\tilde{z}$. Fix $\tilde{x}$, and note that $\nabla_z f(\tilde{x}, \tilde{z}) = 0$, together with concavity of this function in z, means that $\tilde{z}$ maximizes the function, *i.e.*, for any x we have

$$f(\tilde{x}, \tilde{z}) \ge f(\tilde{x}, z).$$

(c) To establish this we argue the same way. If the saddle-point condition holds, then $\tilde{x}$ minimizes $f(x, \tilde{z})$ over all x. Therefore we have $\nabla f_x(\tilde{x}, \tilde{z}) = 0$. Similarly, since $\tilde{z}$ maximizes $f(\tilde{x}, z)$ over all z, we have $\nabla f_z(\tilde{x}, \tilde{z}) = 0$.