TOPICS ON

CONVEX (AND NONCONVEX)

ANALYSIS AND OPTIMIZATION

Dimitri P. Bertsekas
Massachusetts Institute of Technology

A followup to the book "Convex Analysis and Optimization," Athena Scientific, 2003

Topics

(1) Sensitivity analysis under very general conditions.

(2) Nonemptiness of closed set intersections. Unification of conditions for existence of an optimal solution, absence of a duality gap, and min-max=max-min.

PART I: SENSITIVITY ANALYSIS

- Classical NLP sensitivity analysis:
  - Requires 2nd order sufficiency conditions, etc
- Convex programming sensitivity analysis:
  - Assumes no duality gap
  - Considers the directional derivative of the optimal cost under (straight line) constraint perturbations
- We present a more general framework:
  - We allow a duality gap
  - We consider sensitivity under curved constraint perturbations

- We show that the dual optimal solution of minimum norm determines the steepest descent rate of the optimal cost.

- The analysis is based on an extended version of the Fritz John conditions, which are of independent interest (paper by Bertsekas, Tseng, Ozdaglar).

MULTIPLIERS AND DUALITY

- Consider the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad x \in X,\ g_1(x) \le 0, \ldots, g_r(x) \le 0 \end{aligned}$$

assuming that its optimal value f^* is finite.

- A vector $\mu^* = (\mu_1^*, \ldots, \mu_r^*)$ is said to be a geometric (or G-) multiplier if $\mu^* \ge 0$ and

$$f^* = \inf_{x \in X} L(x, \mu^*) \equiv f(x) + \mu' g(x),$$

- The dual problem is

$$\begin{aligned} &\text{maximize} \quad q(\mu) \equiv \inf_{x \in X} L(x, \mu) \\ &\text{subject to} \quad \mu \ge 0, \end{aligned}$$

and its optimal value q^* satisfies $q^* \le f^*$.

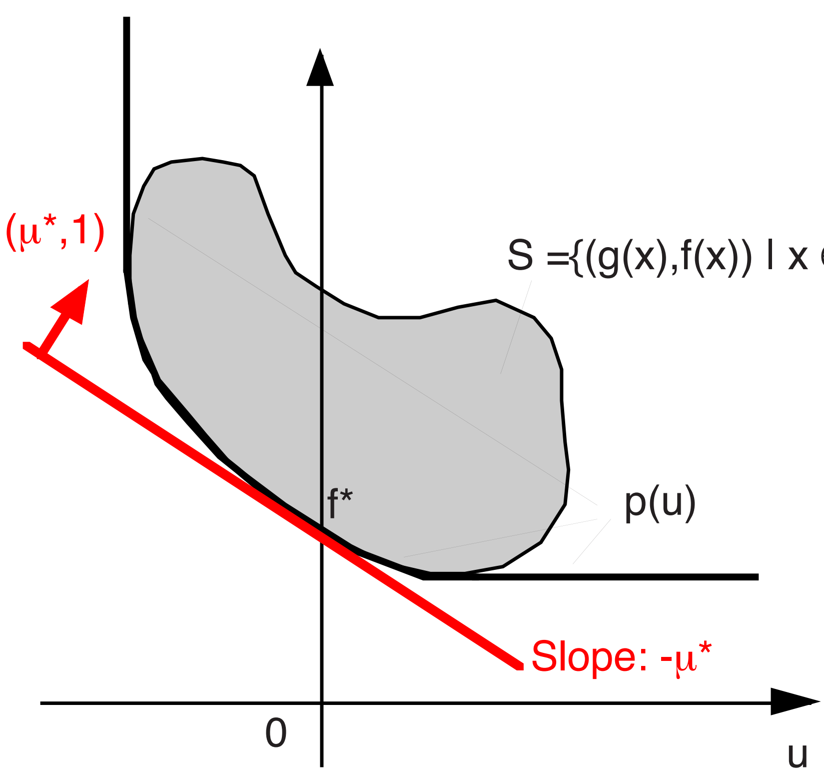
THE PRIMAL FUNCTION

- The primal function is the perturbed optimal value

$$p(u) = \inf_{\substack{x \in X \\ g(x) \leq u}} f(x)$$

- μ^* is a G-multiplier iff $-\mu^*$ is a subgradient of p at 0 (assuming that $p(u) > -\infty$ for all u).

- Classical sensitivity theory revolves around the directional derivative of p at 0.



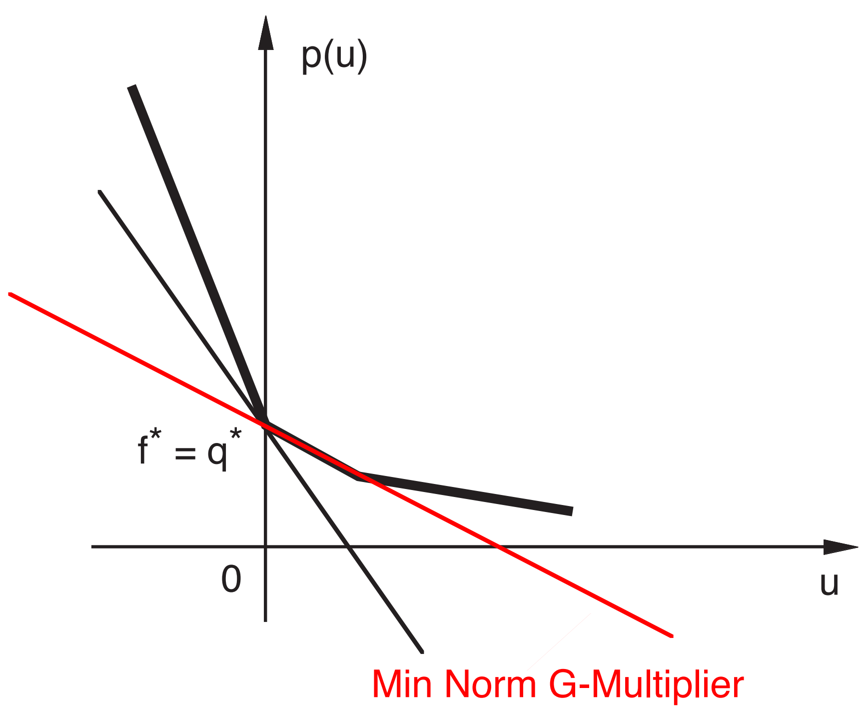
CLASSICAL SENSITIVITY THEORY

- Assume that $p(u) > -\infty$ for all u and 0 belongs to $\mathrm{ri}\big(\mathrm{dom}(p)\big)$. Then:

  (a) The set of G-multipliers is nonempty.

  (b) If μ^* is the G-multiplier of minimum norm and $\mu^* \neq 0$:

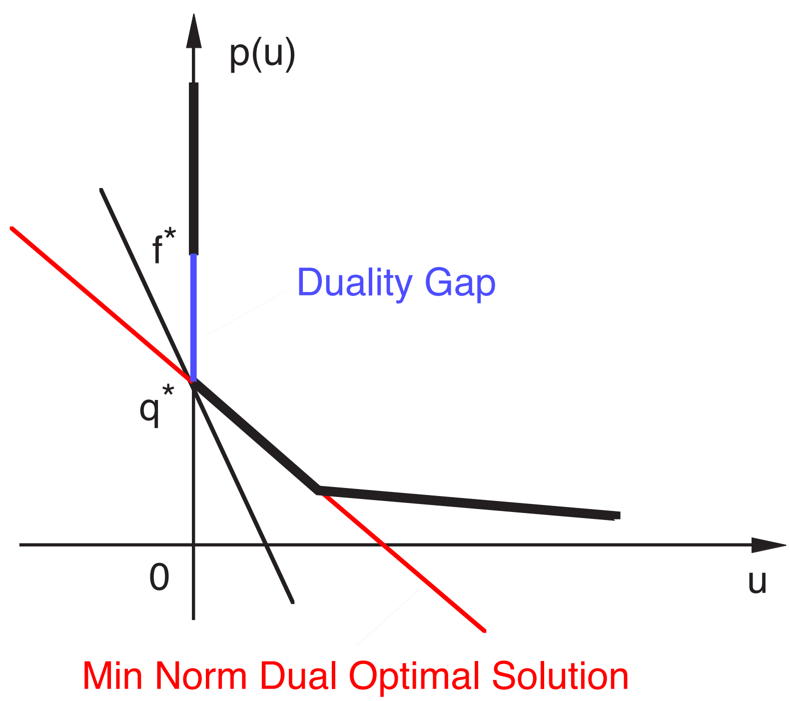
  - The direction of steepest descent of p at 0 is $\mu^*/\|\mu^*\|$
  - The rate of steepest descent (per unit norm of constraint violation) is $\|\mu^*\|$.



BREAKDOWN OF CLASSICAL THEORY

- If 0 does not belong to $\mathrm{ri}\big(\mathrm{dom}(p)\big)$, sensitivity theory breaks down because:

  (1) There may exist a duality gap and no G-multiplier.



  (2) Even if there is no duality gap and there exists a G-multiplier, the formula

  (Dir. derivative of p along $\mu^*/\|\mu^*\|) = -\|\mu^*\|$

  may not hold.

EXTENDED SENSITIVITY THEORY

Proposition: Assume that the primal function p is convex, and that $-\infty < q^* \leq f^* < \infty$. If μ^* is a dual optimal solution of minimum norm and $\mu^* \neq 0$, then for all infeasible $x \in X$

$$\frac{q^* - f(x)}{\|g^+(x)\|} \leq \|\mu^*\|,$$

where $g^+(x) \in \Re^r$ has components $\max\{0, g_j(x)\}$. Furthermore, the inequality is sharp, i.e., there exists a sequence $\{x_k\} \subset X$ such that

$$\frac{q^* - f(x_k)}{\|g^+(x_k)\|} \to \|\mu^*\|, \qquad \|g^+(x_k)\| \to 0.$$

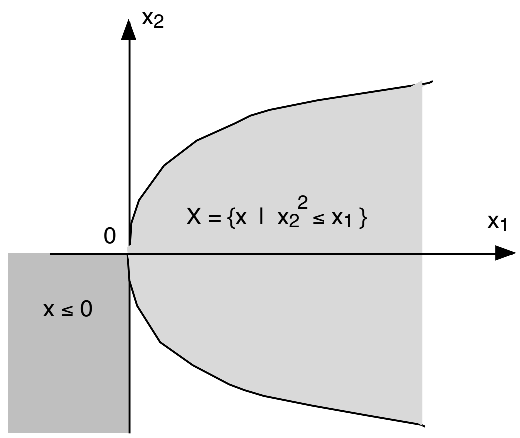
- **Note:** The sequence $g^+(x_k)$ may have to go to 0 along a curve.

EXAMPLE

- Consider the 2-dimensional problem

minimize $\quad -x_2$

subject to $\quad x \in X = \{x \mid x_2^2 \le x_1\}, \quad g(x) = x \le 0.$



- Then $f^* = q^* = 0$, and the set of G-multipliers is $\{\mu \ge 0 \mid \mu_2 = 1\}$.
- However, the min norm G-multiplier, $\mu^* = (0, 1)$, is not a steepest descent direction; along μ^*, we have

$$p'(0; \mu^*) = 0.$$

- The steepest descent rate is $\|\mu^*\|$, but can be obtained only by approaching 0 along a curve.

PART II: CLOSED SET INTERSECTIONS

- Given a sequence of nonempty closed sets $\{S_k\}$ in $\Re^n$ with $S_{k+1} \subset S_k$ for all k, when is $\cap_{k=0}^{\infty} S_k$ nonempty?

- Set intersection theorems are significant in at least three major contexts:

  – Existence of optimal solutions

  – Duality gap issue, i.e., equality of optimal values of the primal convex problem

$$\text{minimize}_{x \in X,\, g(x) \leq 0}\, f(x)$$

  and its dual

$$\text{maximize}_{\mu \geq 0}\, q(\mu) \equiv \inf_{x \in X} \{ f(x) + \mu' g(x) \}$$

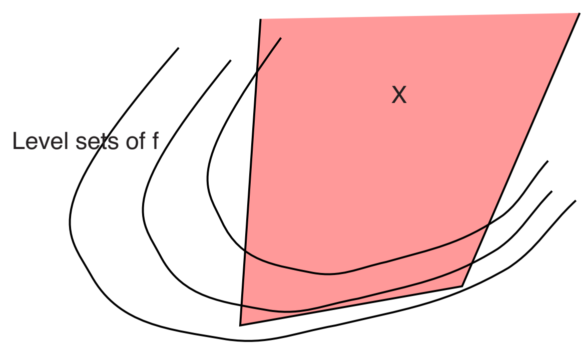
  – min-max = max-min issue, i.e., equality in

$$\min_x \max_z \phi(x, z) = \max_z \min_x \phi(x, z),$$

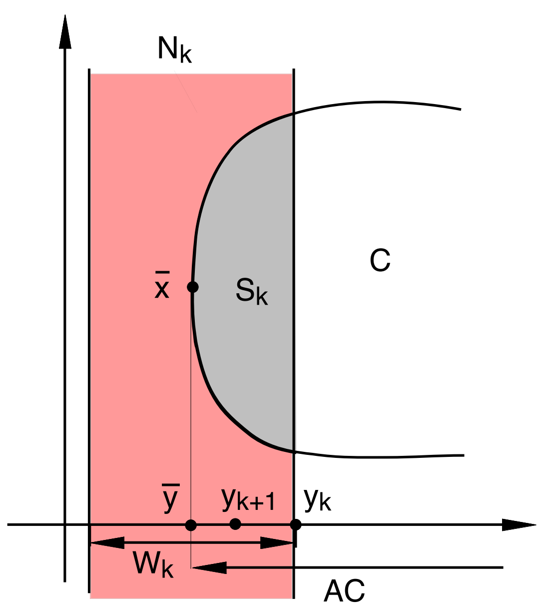
  where ϕ is convex in x and concave in z

SOME SPECIFIC CONTEXTS I

- Does a function $f : \Re^n \mapsto (-\infty, \infty]$ attain a minimum over a set X?
  - This is true iff the intersection of the nonempty sets $\{x \in X \mid f(x) \leq \gamma\}$ is nonempty



- If C is closed and A is a matrix, is $A\,C$ closed?



  - Many interesting special cases, e.g., if C_1 and C_2 are closed, is $C_1 + C_2$ closed?

SOME SPECIFIC CONTEXTS II

- If $F(x,u)$ is closed, is $p(u) = \inf_x F(x,u)$ closed?
  - Critical question in the duality gap issue, where

$$F(x,u) = \begin{cases} f(x) & \text{if } x \in X,\ g(x) \leq u, \\ \infty & \text{otherwise} \end{cases}$$

    and p is the primal function.
  - Critical question regarding min-max=max-min where

$$F(x,u) = \begin{cases} \sup_{z \in Z} \{\phi(x,z) - u'z\} & \text{if } x \in X, \\ \infty & \text{if } x \notin X. \end{cases}$$

    We have min-max=max-min if

$$p(u) = \inf_{x \in \Re^n} F(x,u)$$

    is closed.
  - Can be addressed by using the relation

$$\text{Proj}\big(\text{epi}(F)\big) \subset \text{epi}(p) \subset \text{cl}\Big(\text{Proj}\big(\text{epi}(F)\big)\Big)$$

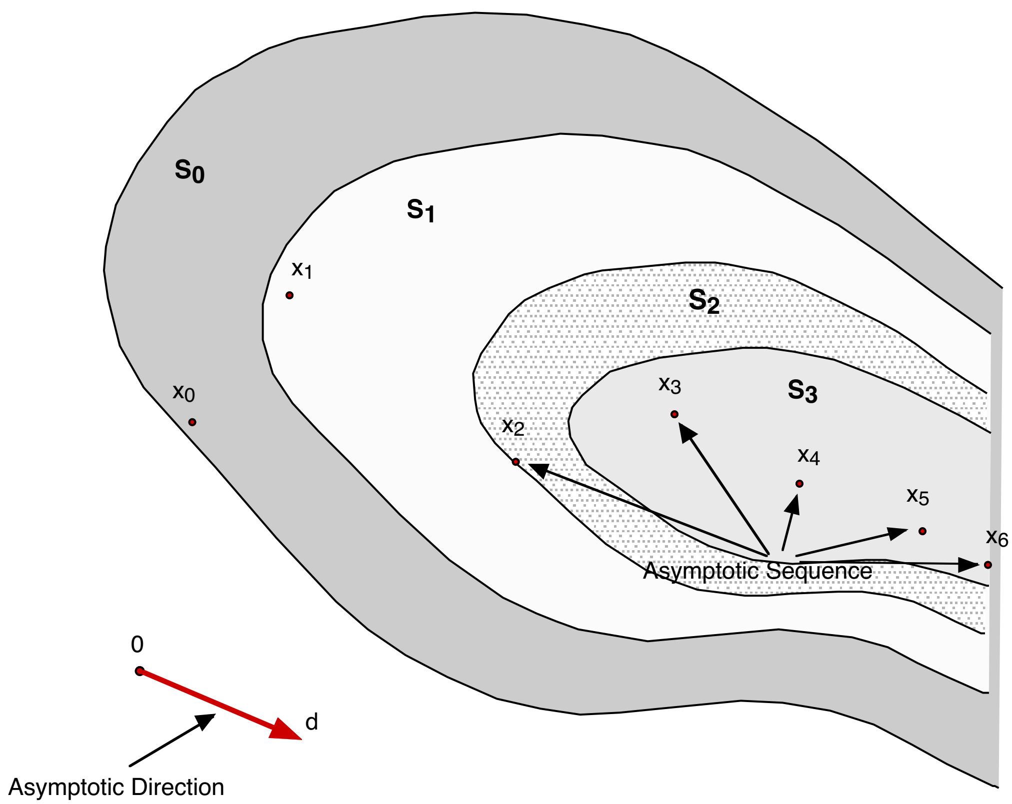
ASYMPTOTIC DIRECTIONS

- Given a sequence of nonempty nested closed sets $\{S_k\}$, we say that a vector $d \neq 0$ is an asymptotic direction of $\{S_k\}$ if there exists $\{x_k\}$ s. t.

$$x_k \in S_k, \qquad x_k \neq 0, \qquad k = 0, 1, \ldots$$

$$\|x_k\| \to \infty, \qquad \frac{x_k}{\|x_k\|} \to \frac{d}{\|d\|}.$$

- A sequence $\{x_k\}$ associated with an asymptotic direction d as above is called an asymptotic sequence corresponding to d.

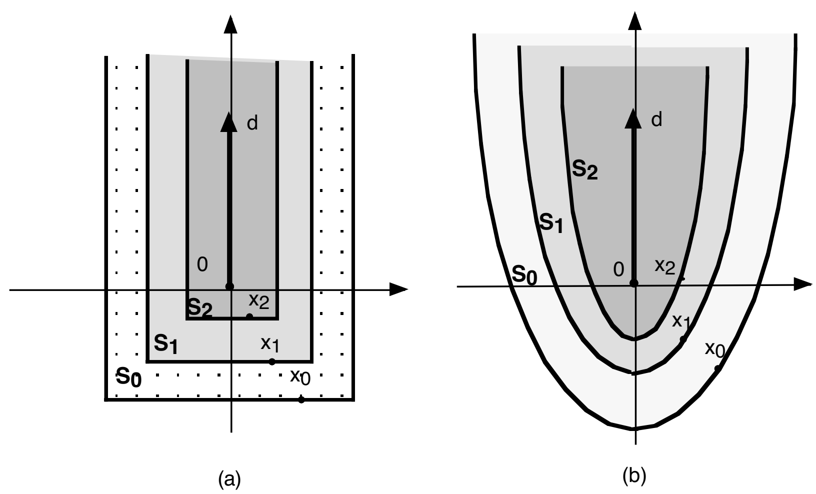


RETRACTIVE ASYMPTOTIC DIRECTIONS

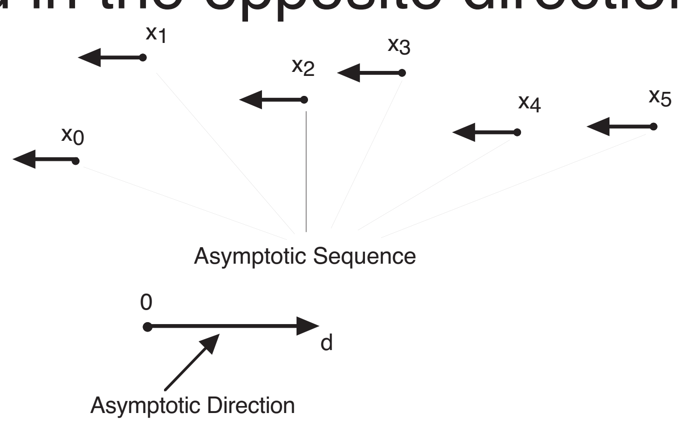
- An asymptotic sequence $\{x_k\}$ and corresponding asymptotic direction are called retractive if for every $\overline{\alpha} > 0$ there exists $\overline{k}$ such that

$$x_k - \alpha d \in S_k, \qquad \forall\, \alpha \in [0, \overline{\alpha}],\ k \geq \overline{k}.$$

$\{S_k\}$ is called retractive if all its asymptotic sequences are retractive.



- **Important observation:** A retractive asymptotic sequence $\{x_k\}$ (for large k) gets closer to 0 when shifted in the opposite direction $-d$.



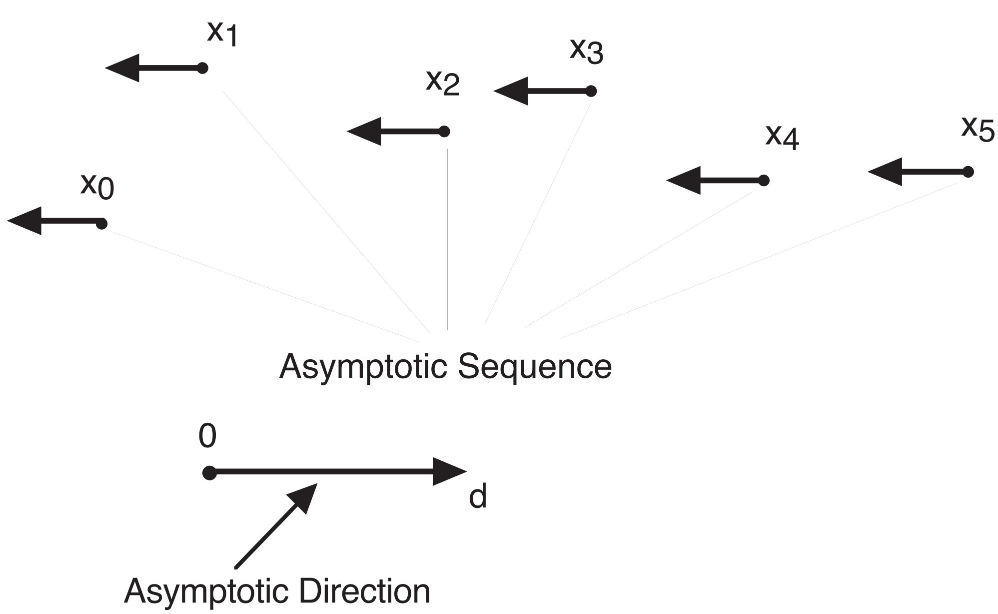
SET INTERSECTION THEOREM

Proposition: The intersection of a retractive nested sequence of closed sets is nonempty.

- Key proof ideas:

  (a) The intersection $\cap_{k=0}^{\infty} S_k$ is empty iff there is an unbounded sequence $\{x_k\}$ consisting of minimum norm vectors from the S_k.

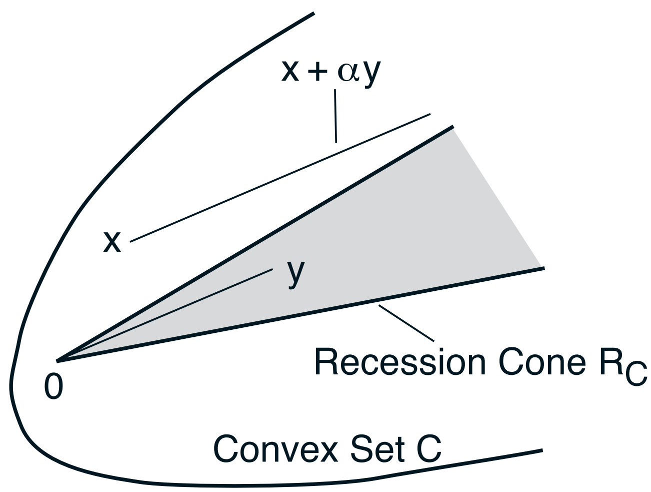
  (b) An asymptotic sequence $\{x_k\}$ consisting of minimum norm vectors from the S_k cannot be retractive, because $\{x_k\}$ eventually gets closer to 0 when shifted opposite to the asymptotic direction.



CALCULUS OF RETRACTIVE SEQUENCES

- Unions and intersections of retractive set sequences are retractive.

- Recall the recession cone R_C of a convex set C, and its lineality space $L_C = R_C \cap (-R_C)$.



If the S_k are convex, the set of asymptotic directions of $\{S_k\}$ is the set of nonzero common directions of recession of the S_k.

- The vector sum of a compact set and a polyhedral cone (e.g., a polyhedral set) is retractive.

- The level sets of a continuous concave function $\{x \mid f(x) \leq \gamma\}$ are retractive.

APPLICATION: EXISTENCE OF SOLUTIONS ISSUES

- Standard results on existence of minima of convex functions generalize with simple proofs using the set intersection theorems.

- **Example 1:** The set of minima of a closed convex function f over a closed set X is nonempty if there is no asymptotic direction of X that is a direction of recession of f.

- **Example 2:** The set of minima of a closed quasiconvex function f over a retractive closed set X is nonempty if

$$A \cap R \subset L,$$

where A: set of asymptotic directions of X,

$$R = \cap_{k=0}^{\infty} R_{\overline{S}_k}, \qquad L = \cap_{k=0}^{\infty} L_{\overline{S}_k},$$

$$\overline{S}_k = \{x \mid f(x) \leq \gamma_k\}$$

and $\gamma_k \downarrow f^*$.

LINEAR AND QUADRATIC PROGRAMMING

- **Frank-Wolfe Theorem:** Let X be polyhedral and

$$f(x) = x'Qx + c'x$$

where Q is symmetric (not necessarily positive semidefinite). If the minimal value of f over X is finite, there exists a minimum of f of over X.

- The proof is straightforward using the set intersection theorems.

- **Extensions** (based on the subsequent theory):
  - X can be the vector sum of a compact set and a polyhedral cone.
  - f can be of the form

$$f(x) = p(x'Qx) + c'x$$

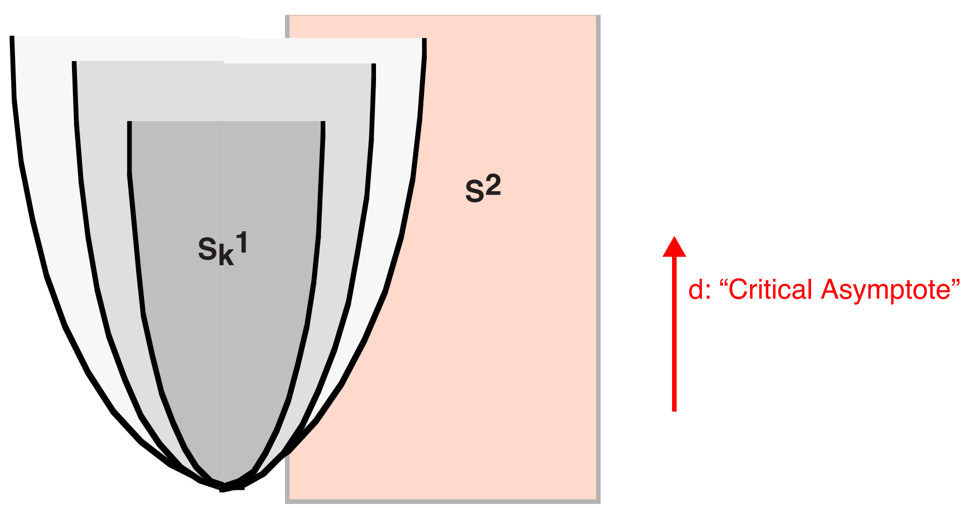
where Q is positive semidefinite and p is a polynomial.

ASYMPTOTIC INSIGHTS

- **Key question:** Given $\{S_k^1\}$ and $\{S_k^2\}$, each with nonempty intersection by itself, and with

$$S_k^1 \cap S_k^2 \neq \emptyset,$$

for all k, when does the intersection sequence $\{S_k^1 \cap S_k^2\}$ have an empty intersection?



- With a few examples, we see that the trouble lies with the existence of some “critical asymptotes”

- “Critical asymptotes” roughly are: Common asymptotic directions d such that starting at $\cap_k S_k^2$ and looking at the horizon along d, we do not meet $\cap_k S_k^1$ (and similarly with the roles of S_k^1 and S_k^2 reversed).

CRITICAL DIRECTIONS

- We say that an asymptotic direction d of $\{S_k\}$, with $\cap_k S_k \neq \emptyset$ is a horizon direction with respect to a set G if for every $x \in G$, we have $x + \alpha d \in \cap_k S_k$ for all α sufficiently large.

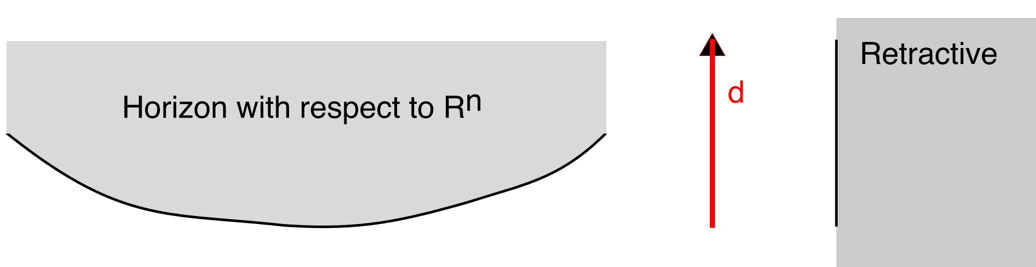
- We say that an asymptotic direction d of $\{S_k\}$ is noncritical with respect to a set G if it is either a horizon direction with respect to G or a retractive horizon direction with respect to $\cap_k S_k$. Otherwise, d is called critical with respect to G.

- **Example:** The as. directions of a vector sum S of a compact set and a polyhedral set are noncritical (are retractive horizon dir. with resp. to S).

- **Example:** The asymptotic directions of a level set sequence of a convex quadratic

$$S_k = \{x \mid x'Qx + c'x + b \leq \gamma_k\}, \qquad \gamma_k \downarrow 0,$$

are noncritical. (Extension: convex polynomials.)



CRITICAL DIRECTION THEOREM

- Roughly it says that: For the intersection of a set sequence $\{S_k^1 \cap S_k^2 \cap \cdots \cap S_k^r\}$ to be empty, some common asymptotic direction must be critical for one of the $\{S_k^j\}$ with respect to the others.

- **Critical Direction Theorem:** Consider $\{S_k^1\}$ and $\{S_k^2\}$, each with nonempty intersection by itself. If

$$S_k^1 \cap S_k^2 \neq \emptyset \quad \text{for all } k, \text{ and} \quad \cap_{k=0}^{\infty} (S_k^1 \cap S_k^2) = \emptyset,$$

there is a common asymptotic direction that is critical for $\{S_k^1\}$ with respect to $\cap_k S_k^2$ (or for $\{S_k^2\}$ with respect to $\cap_k S_k^1$).

- Extends to any finite number of sequences $\{S_k^j\}$.

- **Special Case:** The intersection of set sequences defined by convex polynomial functions

$$S_k^j = \{x \mid p_j(x) \leq \gamma_k^j,\ j = 1, \ldots, r\}, \qquad \gamma_k^j \downarrow 0,$$

is nonempty, assuming all the S_k^j are nonempty. (For example p_j may be convex quadratic.)

EXISTENCE OF SOLUTIONS THEOREMS

- **Convex Quadratic/Polynomial Problems:** For $j = 0, 1, \ldots, r$, let $f_j : \Re^n \mapsto \Re$ be polynomial convex functions. Then the problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad f_j(x) \leq 0, \qquad j = 1, \ldots, r, \end{aligned}$$

has at least one optimal solution if and only if its optimal value is finite.

- **Extended Frank-Wolfe Theorem:** Let

$$f(x) = x'Qx + c'x$$

where Q is symmetric, and let X be a closed set whose asymptotic directions are retractive horizon directions with respect to X. If the minimal value of f over X is finite, there exists a minimum of f over X.