Multi-objective optimization

Jitkomut Songsiri

Department of Electrical Engineering
Faculty of Engineering
Chulalongkorn University

CUEE

February 26, 2024

Outline

Generalized inequality

Convex cone

conic (nonnegative) combination of any two points x_1 and x_2 takes the form

$$x = \theta_1 x_1 + \theta_2 x_2, \quad \text{with} \quad \theta_1, \theta_2 \geq 0$$

convex cone is a set that contains all conic combination of points in the set

examples: ✎ these sets are convex cone

- **non-negative orthant:** $\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x \succeq 0\}$
- **norm cone:** is the set described by $\{(x, t) \mid \|x\| \leq t\}$
- **positive definite cone:** $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0\}$ is the set of positive semidefinite matrices (with $\mathbf{S}^n_{++}$ as the set of positive definite matrices)

Examples

which of the following is a convex cone ?

1. a line
2. a half space
3. a slab
4. $\mathbf{R}^n$
5. a unit-norm ball
6. $S = \{x \in \mathbf{R}^n \mid a^T x = 0 \}$
7. $S = \{x \in \mathbf{R}^n \mid x = ay, \ \ a \geq 0\}$ for some fixed $y \in \mathbf{R}^n$
8. orthogonal complement of $\mathbf{S}^n_+$

Proper cone

a convex cone $K \subseteq \mathbf{R}^n$ is a **proper cone** if

1. K is closed (contains its boundary)
2. K is solid (has non-empty interior)
3. K is pointed (contains no line,*i.e.*, if $x \in K$, then $-x$ cannot be in K)

examples: ✎ these are proper cones

- **non-negative orthant:** $\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x \succeq 0 \}$
- **positive definite cone:** $\mathbf{S}^n_+ = \{X \in \mathbf{S}^n \mid X \succeq 0 \}$ is the set of positive semidefinite matrices (with $\mathbf{S}^n_{++}$ as the set of positive definite matrices)

Generalized inequality

a **generalized inequality** defined by a proper cone K:

$$x \preceq_K y \quad \Longleftrightarrow \quad y - x \in K, \quad x \prec_K y \quad \Longleftrightarrow \quad y - x \in \mathbf{int}\, K$$

examples:

- component-wise inequality: $K = \mathbf{R}^n_+$

$$x \preceq_{\mathbf{R}^n_+} y \quad \Longleftrightarrow \quad x_i \leq y_i, \quad i = 1, 2, \ldots, n$$

- matrix inequality: $K = \mathbf{S}^n_+$

$$X \preceq_{S^n_+} Y \quad \Longleftrightarrow \quad Y - X \text{ is positive semidefinite}$$

a lot of times we can drop the subscript K in the generalized inequality of interest

properties: many properties of $\preceq_K$ are similar to $\leq$ in $\mathbf{R}$

$$x \preceq_K y, \; u \preceq_K v \quad \Longrightarrow \quad x + u \preceq_K y + v$$

Minimum and minimal elements

Minimum and minimal elements

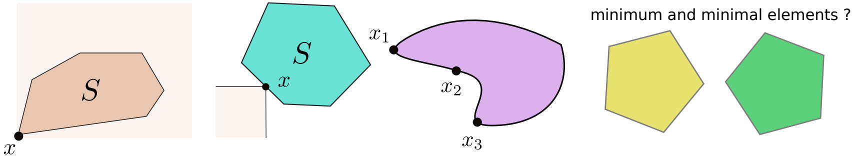
$\preceq_K$ is not general **linear ordering**: we can have $x \not\preceq_K y$ and $y \not\preceq_K$

- $x \in S$ is **the minimum element** of S with respect to $\preceq_K$ if

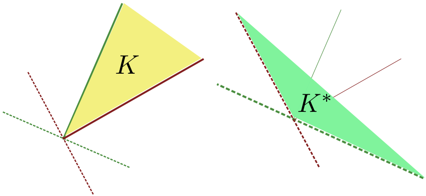
$$y \in S \quad \Longrightarrow x \preceq_K y$$

- $x \in S$ is **a minimal element** of S with respect to $\preceq_K$ if

$$y \in S, \;\; y \preceq_k x \quad \Longrightarrow \quad y = x$$



Dual cone and generalized inequalities



dual cone of a cone K:

$$K^* = \{y \mid y^T x \geq 0, \quad \forall x \in K\}$$

examples: ✎ we can show that

- $K = \mathbf{R}^n_+$: $K^* = \mathbf{R}^n_+$
- $K = \mathbf{S}^n_+$: $K^* = \mathbf{S}^n_+$
- $K = \{(x,y) \mid \|x\|_2 \leq t\}$: $K^* = \{(x,y) \mid \|x\|_2 \leq t\}$
- $K = \{(x,y) \mid \|x\|_1 \leq t\}$: $K^* = \{(x,y) \mid \|x\|_\infty \leq t\}$

when $K = K^*$, it is called **self dual cone**

Dual cone of non-negative orthant

$K = \mathbf{R}^n_+$, by definition of dual cone

$$K^* = \{y \mid y^T x \geq 0, \quad \forall x \in \mathbf{R}^n_+ \}$$

- if $y \succeq 0$, then it's obvious that $y^T x \geq 0$ for all $x \succeq 0$ – this shows that $\mathbf{R}^n_+ \subseteq K^*$
- if $y \in K^*$, we examine how the vector y should be
- since $y^T x \geq 0$ must hold for all $x \succeq 0$, it holds for when x is a standard unit vector
- when $x = e_1$, we have $y_1 \geq 0$, and when $x = e_k$, we obtain $y_k \geq 0$ – equivalently, if $y \in K^*$ then $y \succeq 0$ – this shows that $K^* \subseteq \mathbf{R}^n_+$
- from the above results, $K^* = \mathbf{R}^n_+$

Properties of dual cone

dual cones satisfy several properties:

1. K^* is closed and convex ✎ exercise
2. $K_1 \subseteq K_2$ implies $K_2^* \subseteq K_1^*$ ✎ exercise
3. if K has nonempty interior, then K^* is pointed
4. if the closure of K is pointed then K^* has nonempty interior

✎ from these properties, if K is a proper cone, then so is its dual K^*

Dual generalized inequalities

K^* defines a generalized inequality

$$x \preceq_K y \quad \text{if and only if} \quad \lambda^T x \leq \lambda^T y, \quad \text{for all } \lambda \succeq_{K^*} 0$$
$$x \prec_K y \quad \text{if and only if} \quad \lambda^T x < \lambda^T y, \quad \text{for all } \lambda \succ_{K^*} 0, \lambda \neq 0$$

- the property is just a re-statement of the relationship between a proper cone K and its dual K^*

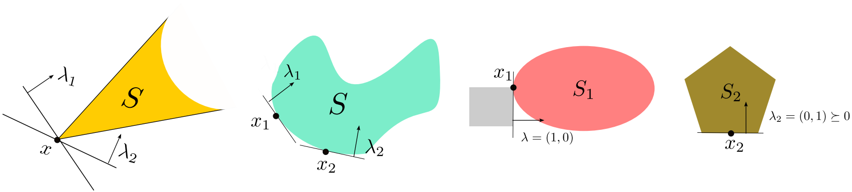
$$\lambda \succeq_{K^*} 0 \Leftrightarrow \lambda \in K^* \Leftrightarrow \lambda^T (y - x) \geq 0, \quad \text{for all } y - x \in K$$

- for a specific example,

$$\lambda \succeq_{K^*} 0 \quad \Longleftrightarrow \quad \lambda^T z \geq 0, \ \forall z \in K$$

Minimum and minimal elements via dual inequalities

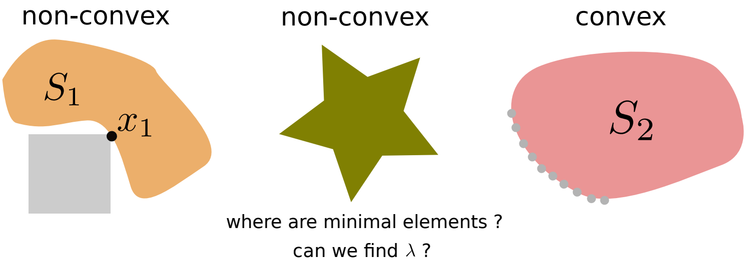
minimum element w.r.t. $\preceq_K$: x is the minimum element of S if and only if all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over S



minimal element w.r.t. $\preceq_K$

- fact: if x minimizes $\lambda^T z$ over S for some $\lambda \succ_{K^*} 0$, then x is minimal
- (the converse is not true) e.g., $x_1 \in S_1$ is minimal but is not a minimizer of $\lambda^T z$ for $\lambda \succ 0$
- $x_2 \in S_2$ is *not minimal* but it does minimize $\lambda^T z$ over $z \in S_2$ for $\lambda = (0,1) \succeq 0$

Minimal elements of a convex set



- S_1 is non-convex; we see that $x_1 \in S_1$ is minimal but there exists no λ for which x minimizes $\lambda^T z$ over $z \in S_1$
- if x is a minimal element of a **convex set** S, then there exists a nonzero $\lambda \succeq_{K^*} 0$ such that x minimizes $\lambda^T z$ over S

Pareto optimal frontier

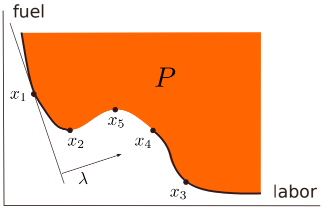
example: a product requires n resources (labor, electricity, gas, water) collected as a resource vector, x

- the **production set** $P \subseteq \mathbf{R}^n$ is defined as the set of all resource vectors x that correspond to some production method
- production methods with resource vectors that are **minimal elements** of P (w.r.t. $\preceq$) are called **Pareto optimal** or efficient
- the set of minimal elements of P is called the **efficient production frontier**
- one production method with resource vector x is *better* than another, with resource vector y if $x \preceq y$, $x \neq y$
- if cost $= \lambda^T x = \lambda_1 x_1 + \lambda_2 x_2 + \cdots + \lambda_n x_n$ then λ_i is the price of resource i

find Pareto optimal production methods by minimizing

$$\lambda^T x = \lambda_1 x_1 + \cdots + \lambda_n x_n$$

over the set P using any λ that $\lambda \succ 0$



- x_1, x_2, x_3 are efficient (Pareto optimal)
- x_4 is not efficient (since x_2 corresponds to a production method that uses less labor while no more fuel)
- x_5 is not efficient (since x_2 is better)
- the point x_1 is efficient and is also the minimum cost production method for the price vector λ (which is positive)
- the point x_2 is efficient but cannot be found by minimizing the total cost $\lambda^T x$ for any $\lambda \succeq 0$

K-convex functions

let $K \subseteq \mathbf{R}^m$ be a proper cone with generalized inequality $\preceq_K$

we say $f : \mathbf{R}^n \to \mathbf{R}^m$ is K-**convex** if for all x, y, and $0 \leq \theta \leq 1$

$$f(\theta x + (1 - \theta)y) \preceq_K \theta f(x) + (1 - \theta)f(y)$$

example: $K = \mathbf{R}^m_+$

- $f : \mathbf{R}^n \to \mathbf{R}^m$ is convex w.r.t. $\mathbf{R}^m_+$ if and only if

$$f(\theta x + (1 - \theta)y) \preceq \theta f(x) + (1 - \theta)f(y)$$

- each component f_i is a convex function

Vector optimization

Introduction

setting: minimizing $f_0 : \mathbf{R}^n \to \mathbf{R}^m$ (vector-valued function) over a feasible set

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in \mathcal{C} \end{array}$$

a **vector optimization** has a **vector-valued** objective function

- example: $f_0(x) = (\text{fuel,time})$ the energy used and time spent of a vehicle parameter x
- require a generalized inequality definition for comparing any two vectors of $f_0(x)$

$$\begin{bmatrix} 5 \\ 2 \end{bmatrix} \preceq \begin{bmatrix} 10 \\ 3 \end{bmatrix} \quad \text{but} \quad \begin{bmatrix} 5 \\ 2 \end{bmatrix} \not\preceq \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

here, for $f_0(x) \in \mathbf{R}^n$, we typically use the **non-negative orthant** to define $\preceq$

General vector optimization

a vector optimization problem is defind as

$$\begin{array}{ll} \text{minimize (w.r.t. } K) & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, 2, \ldots, m \\ & h_i(x) = 0, \quad i = 1, 2, \ldots, p \end{array}$$

where

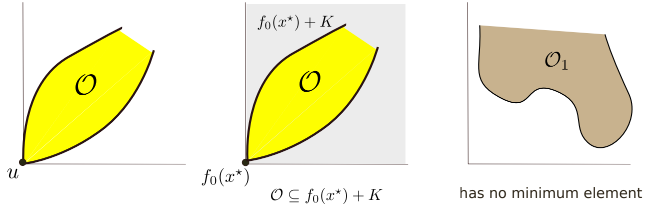
- $f_0 : \mathbf{R}^n \to \mathbf{R}^q$ (vector-valued function)
- $K \subseteq \mathbf{R}^q$ is a proper cone
- f_i's and h_i's are inequality and equality constraint functions

here, f_0 takes value in $\mathbf{R}^q$ and we use K to compare objective values

definition: we say the problem is **convex vector optimization** if f_0 is K-convex

Optimal points and values

define $\mathcal{O} = \{ f_0(x) \mid \exists x \in \mathcal{D}, \;\; x \in \mathcal{C} \}$ the set of **acheivable objective values**



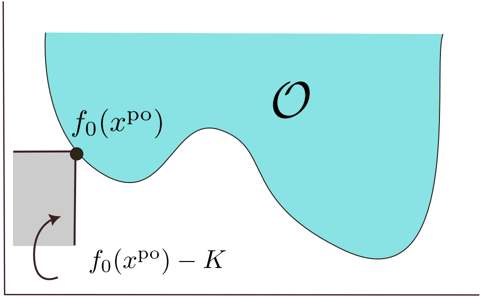
- u is said to be the **minimum** element of $\mathcal{O}$ if $u \preceq v$, for every $v \in \mathcal{O}$
- if $\mathcal{O}$ has a minimum point (then it is unique) and

$$\exists \text{ feasible } x^\star \quad \text{such that} \quad f_0(x^\star) \preceq f_0(y), \quad \text{for all feasible } y$$

  then we say $x^\star$ is **optimal**
- a point $x^\star$ is **optimal** if and only if it is feasible and $\mathcal{O} \subseteq f_0(x^\star) + K$ ($f_0(x^\star)$ is 'better' which is below and the left of other $f_0(y)$)

Pareto optimal points

consider when $\mathcal{O}$ does not have a minimum element (often occur in most vector opt)



$f_0(x) - K$ is the set of values that are better than or equal to $f_0(x)$

a problem can have many Pareto optimal points

- x is called **Pareto optimal** (or efficient) if $f_0(x)$ is a minimal element of $\mathcal{O}$
- recall: u is said to be a **minimal** element of $\mathcal{O}$ if $v \in \mathcal{O}$, $v \preceq u$ only if $v = u$
- a point x is **Pareto optimal** if and only if it is feasible and

$$(f_0(x) - K) \cup \mathcal{O} = \{f_0(x)\}$$

(the only achievable value better than or equal to $f_0(x)$ is $f_0(x)$ itself)

Scalarization

scalarization is a standard technique for finding Pareto optimal points

by choosing $\lambda \succeq_{K^*} 0$ (positive in the dual generalized inequality),

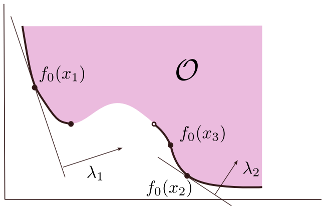
let x be an optimal point of the *scalar* optimization problem:

$$\begin{array}{ll} \text{minimize} & \lambda^T f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, 2, \ldots, m \\ & h_i(x) = 0, \quad i = 1, 2, \ldots, p \end{array}$$

then x is Pareto optimal for the vector problem on page 21

(this follows from dual characterization of minimal points)

$K = \mathbf{R}^q_+$**:** for convex vector optimization, we can find Pareto optimal points by solving a *convex scalar* optimization because $\lambda^T f_0(x)$ is non-negative sum of convex functions we can find Pareto optimal points by solving the scalarized problem by varying $\lambda \succ_{K^*} 0$



- all three points: $f_0(x_1), f_0(x_2), f_0(x_3)$ are Pareto optimal
- only $f_0(x_1)$ and $f_0(x_2)$ can be obtained by scalarization
- $f_0(x_1)$ minimizes $\lambda_1^T u$ over all $u \in \mathcal{O}$ and $f_0(x_2)$ minimizes $\lambda_2^T u$ where both $\lambda_1, \lambda_2 \succ 0$
- $f_0(x_3)$ cannot be found by scalarization

Multicriterion optimization

when $K = \mathbf{R}^q_+$, the vector optimization is called a **multicriterion** or **multi-objective** optimization (MOP)

$$\text{objective function} \quad f_0(x) = (F_1(x), F_2(x), \ldots, F_q(x))$$

an MOP is convex if F_i, for $i = 1, 2, \ldots, q$ are convex and the constraint set is convex

suppose x and y are both feasible

- $F_i(x) \leq F_i(y)$ means that x is at least as good as y, according to ith obj
- $F_i(x) < F_i(y)$ means that x is *better* than y, according to ith obj
- if $F_i(x) \leq F_i(y)$ for $i = 1, \ldots, q$ and for at least one j, $F_j(x) < F_j(y)$, we say x is *better than* y or x **dominates** y

Optimal solution of MOP

an optimal point $x^\star$ satisfies

$$F_i(x^\star) \leq F_i(y), \quad i = 1, 2, \ldots, q, \quad \text{for all feasible } y$$

in other words, $x^\star$ is simultaneously optimal for each scalar problem

$$\begin{array}{ll} \text{minimize} & F_j(x) \\ \text{subjec to} & f_i(x) \leq 0, \quad i = 1, 2, \ldots, m \\ & h_i(x) = 0, \quad i = 1, 2, \ldots, p, \end{array}$$

for $j = 1, 2, \ldots, q$

when there is an optimal point, we say that the objectives are **noncompeting**

Pareto optimal point $x^{\rm po}$ satisfies: if y is feasible and $F_i(y) \leq F_i(x^{\rm po})$ for $i = 1, 2, \ldots, q$ then $F_i(x^{\rm po}) = F_i(y)$, $i = 1, \ldots, q$

Scalarizing multicriterion problems

we form the weighted sum objective

$$\lambda^T f_0(x) = \lambda_1 F_1(x) + \cdots + \lambda_q F_q(x), \quad \lambda \succ 0$$

- interpret λ_i as the **weight** quantifying our desire to make F_i small
- the ratio λ_i/λ_j is the **relative weight** showing the relative importance of the ith objective compared to the jth objective
- recall that a weight vector $\lambda \succ 0$ yields the Pareto optimal point
- the set of Pareto optimal values is called the **optimal trade-off surface** or optimal trade-off curve for bicriterion

References

1. Chapter 2 in B. Stephen and L. Vandenberghe, *Convex optimization*, Cambridge Press
2. Lecture slide on convex set, Lieven Vandenberghe, *Convex Optimization*, EE236B, UCLA