# Multi Objective Optimization

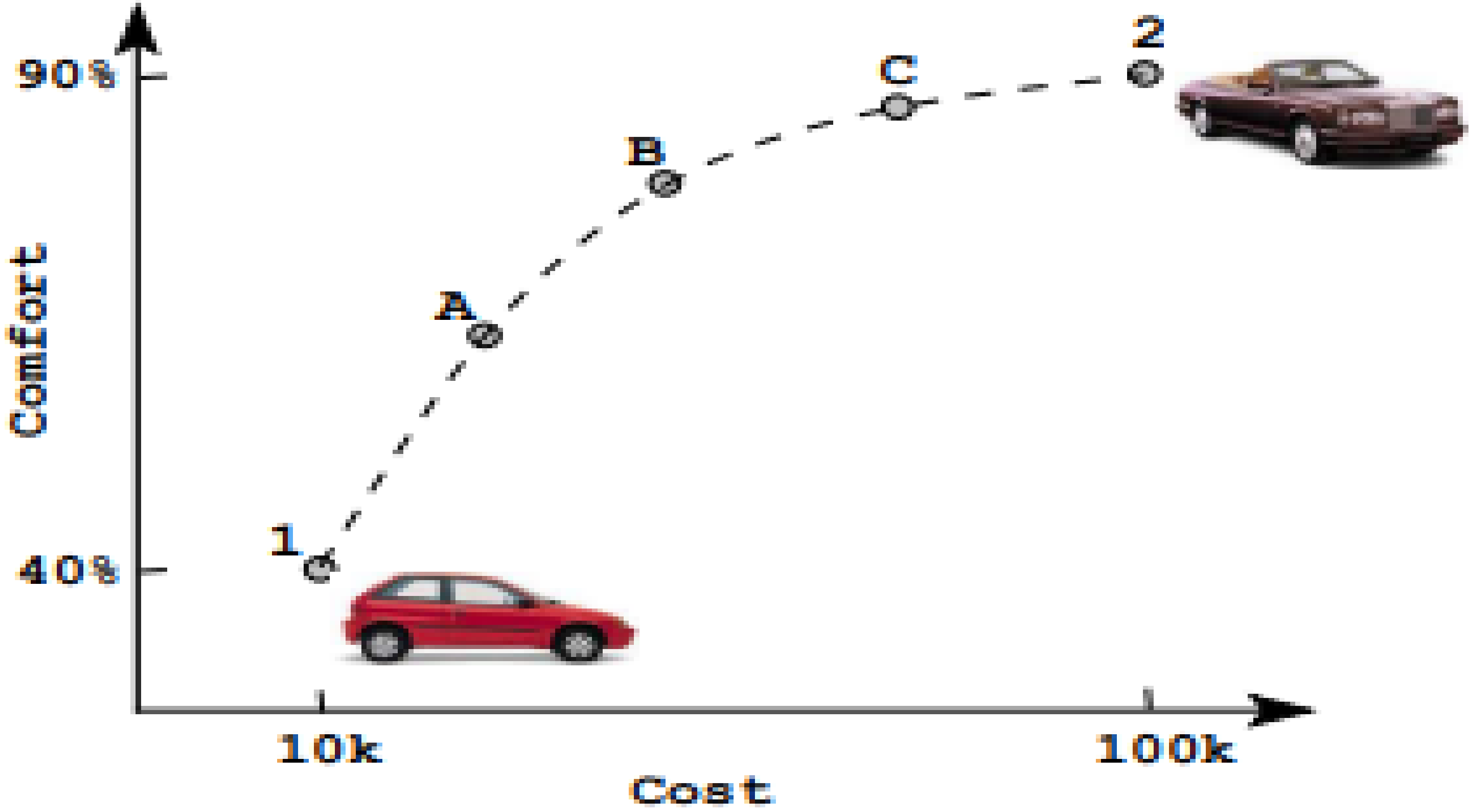
Arsha J K, AP, CSED,
Viswajyothi college of Engineering and Technology

# Multi Objective Optimization

- Multi objective optimization problem.
- Principles of Multi- objective optimization,
- Dominance and pareto-optimality.
- Optimality conditions

# Multi-Objective Optimization

- We often face them



# Multi objective optimization problem

- MOOP addresses scenarios with multiple objective functions, common in practical decision-making.
- Historically, due to limited solution methodologies, MOOPs were often treated as single-objective problems.
- Single and multi-objective optimization algorithms differ fundamentally in their approaches.
- Contrary to single-objective optimization where one optimal solution is sought, MOOP aims for optimal solutions considering multiple objectives.

# Multi objective optimization problem

- MOOP involves minimizing or maximizing multiple objective functions, subject to constraints.
- The problem's general form includes objective functions to be optimized, constraints, and variable bounds.
- Solutions in MOOP are vectors of decision variables, constrained within a decision space defined by variable bounds.

$$\left.\begin{array}{lll} \text{Minimize/Maximize} & f_m(\mathbf{x}), & m = 1, 2, \ldots, M; \\ \text{subject to} & g_j(\mathbf{x}) \geq 0, & j = 1, 2, \ldots, J; \\ & h_k(\mathbf{x}) = 0, & k = 1, 2, \ldots, K; \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)}, & i = 1, 2, \ldots, n. \end{array}\right\} \quad (2.1)$$

A solution $\mathbf{x}$ is a vector of $n$ decision variables: $\mathbf{x} = (x_1, x_2, \ldots, x_n)^T$. The last set of constraints are called variable bounds, restricting each decision variable $x_i$ to take a value within a lower $x_i^{(L)}$ and an upper $x_i^{(U)}$ bound. These bounds constitute a *decision variable space* $\mathcal{D}$, or simply the decision space. Throughout this book, we will use the terms *point* and *solution* interchangeably to mean a solution vector $\mathbf{x}$.

# Multi objective optimization problem

- Associated $g_j(x)$ and $h_k(x)$ em are J inequality and K equality constraints.
- The terms are called constraint functions.
- The inequality constraints are treated as 'greater-than-equal-to types, although a 'less-than-equal-to' type inequality constraint is also taken care of in the above formulation.
- In the latter case, the constraint must be converted into a 'greater-than-equal-to' type constraint by multiplying the constraint function by -1

# Multi objective optimization problem

- A solution x that does not satisfy *all* of the (J + K) constraints and *all* of the 2N variable bounds stated above is called an *infeasible solution.*
- On the other hand, if any solution x satisfies all constraints and variable bounds, it is known as a *feasible solution.*
- Therefore, we realize that in the presence of constraints, the entire decision variable space *D* need not be feasible.
- The set of all feasible solutions is called the *feasible region,* or *S.*

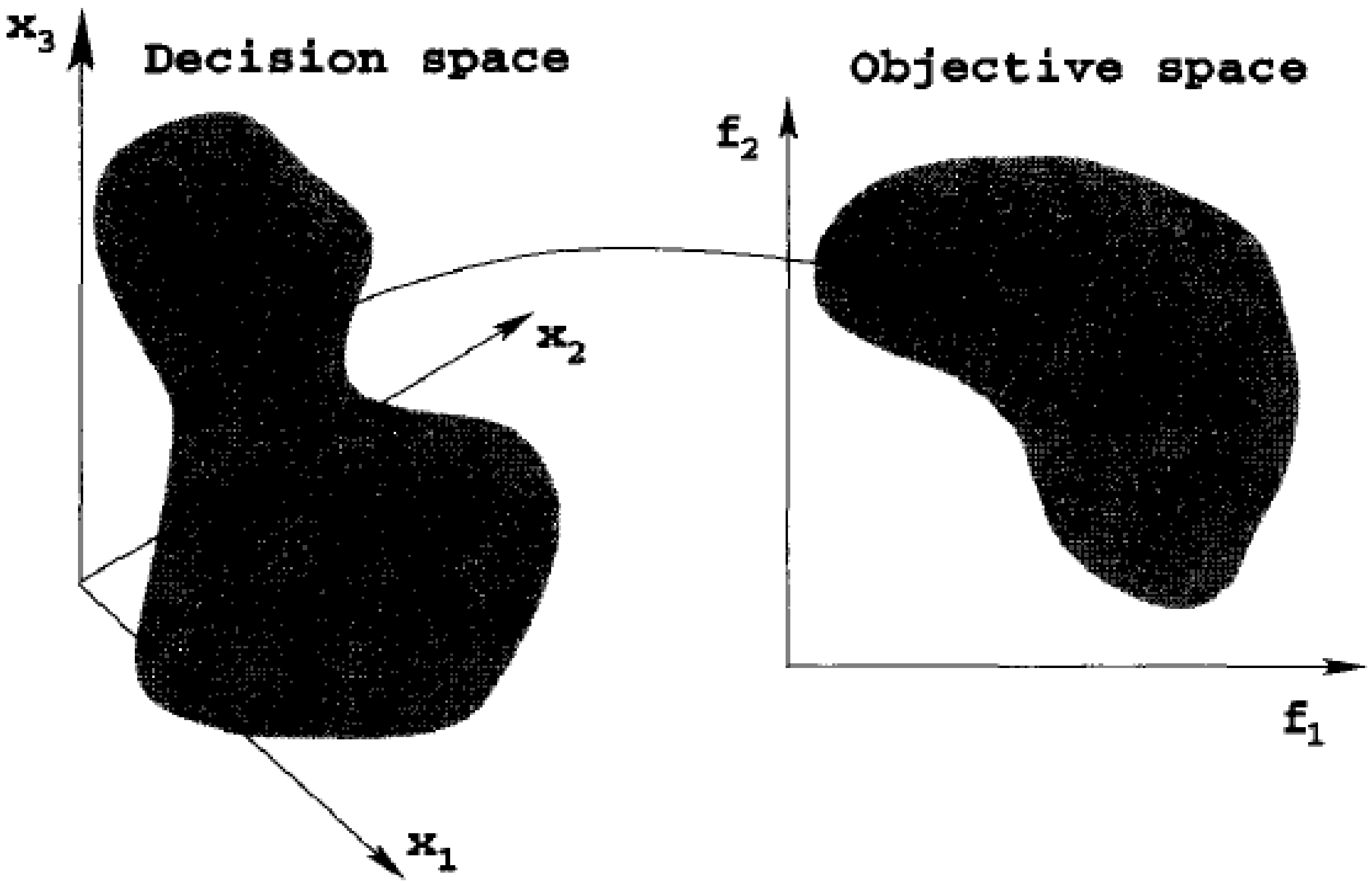
# Multi objective optimization problem

- There are M objective functions f(x) = (f, (x), f2(X), ... , fM(x))T considered in the above formulation.
- Each objective function can be either minimized or maximized.
- The duality principle (Deb, 1995; Rao, 1984; Reklaitis et al., 1983), in the context of optimization, suggests that we can convert a maximization problem into a minimization one by multiplying the objective function by -1.
- The duality principle has made the task of handling mixed type of objectives much easier.
- Many optimization algorithms are developed to solve only one type of optimization problems, such as e.g. minimization problems.
- When an objective is required to be maximized by using such an algorithm, the duality principle can be used to transform the original objective for maximization into an objective for minimization.

# Multi objective optimization problem

- One of the striking differences between single-objective and multiobjective optimization is that in multi-objective optimization the objective functions constitute a multi-dimensional space, in addition to the usual decision variable space.
- This additional space is called the *objective space, Z.*
- For each solution x in the decision variable space, there exists a point in the objective space, denoted by f(x) = z = (Z1,Z2, ... ,ZM)T.
- The mapping takes place between an n-dimensional solution vector and an M-dimensional objective vector.

# MULTI-OBJECTIVE OPTIMIZATION



# *Linear and Nonlinear MOOP*

- If all objective functions and constraint functions are linear, the resulting MOOP is called a multi-objective linear program (MOLP).
- Like the linear programming problems, MOLPs also have many theoretical properties.
- However, if any of the objective or constraint functions are nonlinear, the resulting problem is called a nonlinear multi-objective problem.
- Unfortunately, for nonlinear problems the solution techniques often do not have convergence proofs.
- Since most real-world multi-objective optimization problems are nonlinear in nature, we do not assume any particular structure of the objective and constraint functions here.

# *Convex and Nonconvex MOOP*

**Definition 2.1.** *A function* $f : \mathbb{R}^n \rightarrow \mathbb{R}$ *is a convex function if for any two pair of solutions* $\mathbf{x}^{(1)}, \mathbf{x}^{(2)} \in \mathbb{R}^n$, *the following condition is true:*

$$f\left(\lambda \mathbf{x}^{(1)} + (1-\lambda)\mathbf{x}^{(2)}\right) \leq \lambda f(\mathbf{x}^{(1)}) + (1-\lambda) f(\mathbf{x}^{(2)}),$$

*for all* $0 \leq \lambda \leq 1$.

The above definition gives rise to the following properties of a convex function:

1. The linear approximation of $f(\mathbf{x})$ at any point in the interval $[\mathbf{x}^{(1)}, \mathbf{x}^{(2)}]$ always *underestimates* the actual function value.
2. The Hessian matrix of $f(\mathbf{x})$ is positive definite for all $\mathbf{x}$.
3. For a convex function, a local minimum is always a global minimum.[1]

In the context of single-objective minimization problems, a solution having the smallest function value in its neighbourhood is called a local minimum solution, while a solution having the smallest function value in the feasible search space is called a global minimum solution.

# *Convex and Nonconvex MOOP*

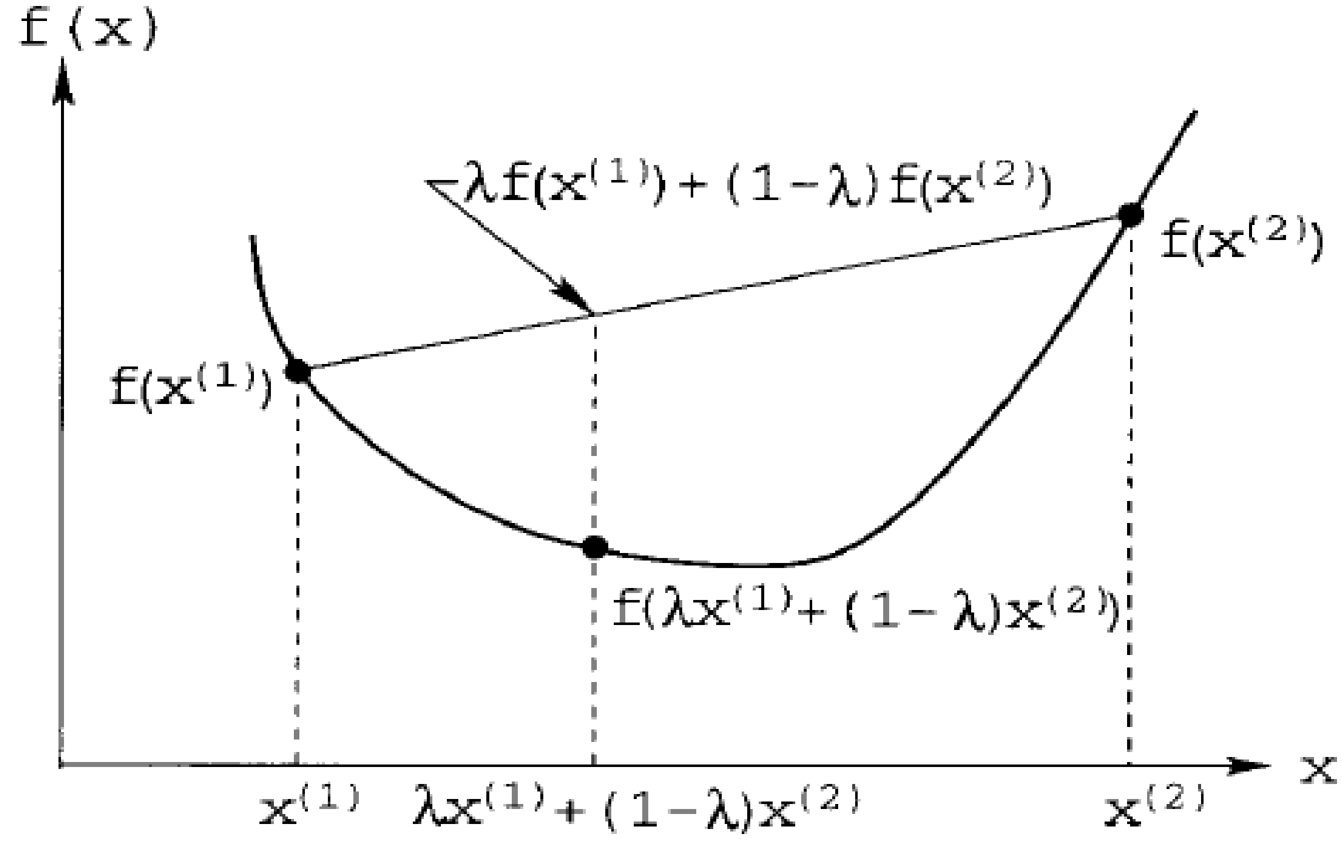



**Figure 6** A convex function is illustrated. A line joining function values at two points $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ always estimates a large value of the true convex function.

# *Convex and Nonconvex MOOP*

equation (2.2) with a '$>$' sign instead of a '$\leq$' sign is called a nonconvex function. To test if a function is convex within an interval, the Hessian matrix $\nabla^2 f$ is calculated and checked for its positive-definiteness at all points in the interval. One of the ways to check the positive-definiteness of a matrix is to compute the eigenvalues of the matrix and check to see if all eigenvalues are positive. To test if a function $f$ is nonconvex in an interval, the Hessian matrix $-\nabla^2 f$ is checked for its positive-definiteness. If it is positive-definite, the function $f$ is nonconvex.

It is interesting to realize that if a function $g(\mathbf{x})$ is nonconvex, the set of solutions satisfying $g(\mathbf{x}) \geq 0$ represents a convex set. Thus, a feasible search space formed with nonconvex constraint functions will enclose a convex region. Now, we are ready to define a convex MOOP.

# *Convex and Nonconvex MOOP*

**Definition 2.2.** *A multi-objective optimization problem is convex if all objective functions are convex and the feasible region is convex (or all inequality constraints are nonconvex and equality constraints are linear).*

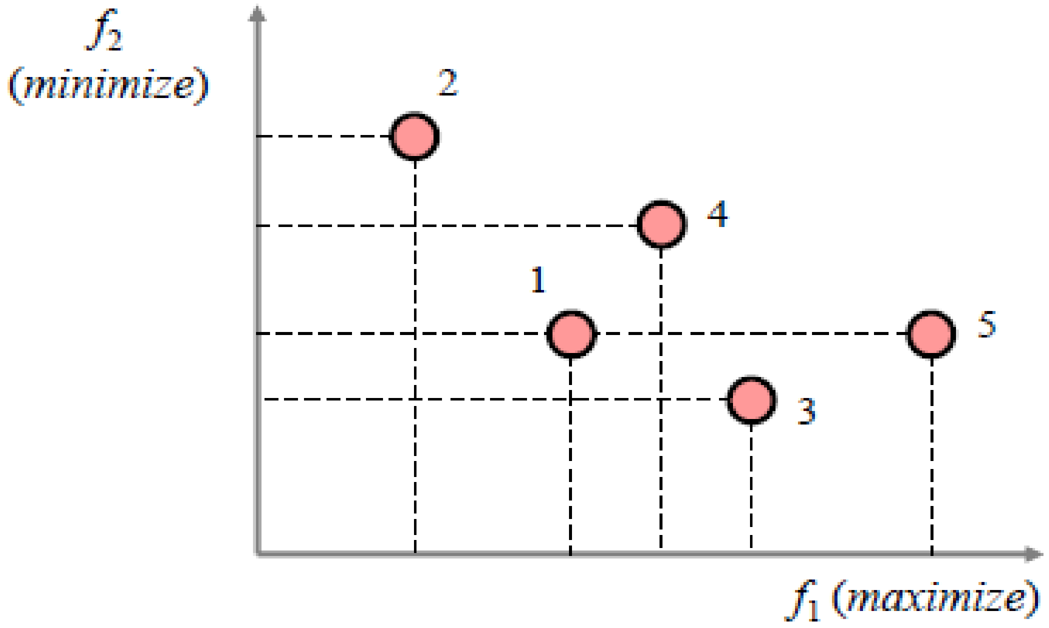
According to this definition, an MOLP is a convex problem. The convexity of an MOOP is an important matter, which we shall see in subsequent chapters. There exist many algorithms which can handle convex MOOPs well, but face difficulty in solving nonconvex MOOPs. Since an MOOP has two spaces, the convexity in each space (objective and decision variable space) is important to a multi-objective optimization algorithm. Moreover, although the search space can be nonconvex, the Pareto-optimal front may be convex.

# Principles of Multi-Objective Optimization

# Definition of Dominance

- **Dominance Test**
  - $x_1$ dominates $x_2$, if
    - Solution $x_1$ is no worse than $x_2$ in all objectives
    - Solution $x_1$ is strictly better than $x_2$ in at least one objective
  - $x_1$ dominates $x_2$ $\Longleftrightarrow$ $x_2$ is dominated by $x_1$

# Example Dominance Test

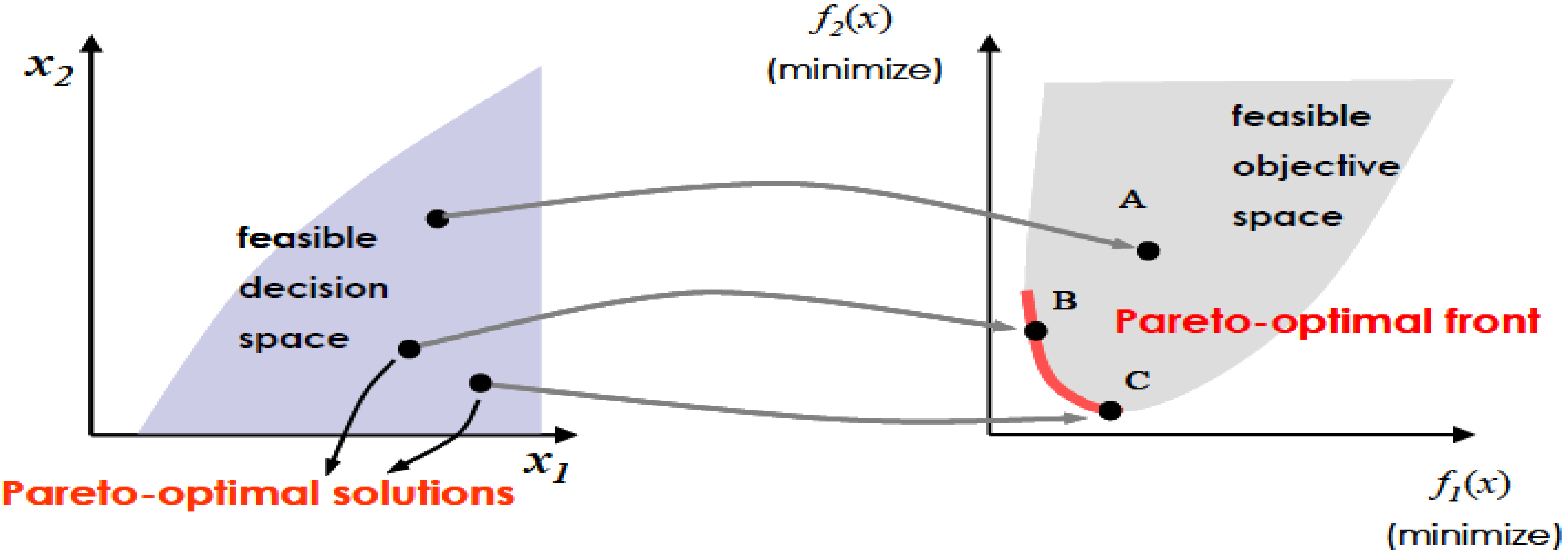



- 1 Vs 2: 1 dominates 2
- 1 Vs 5: 5 dominates 1
- 1 Vs 4: Neither solution dominates

# Pareto Optimal Solution

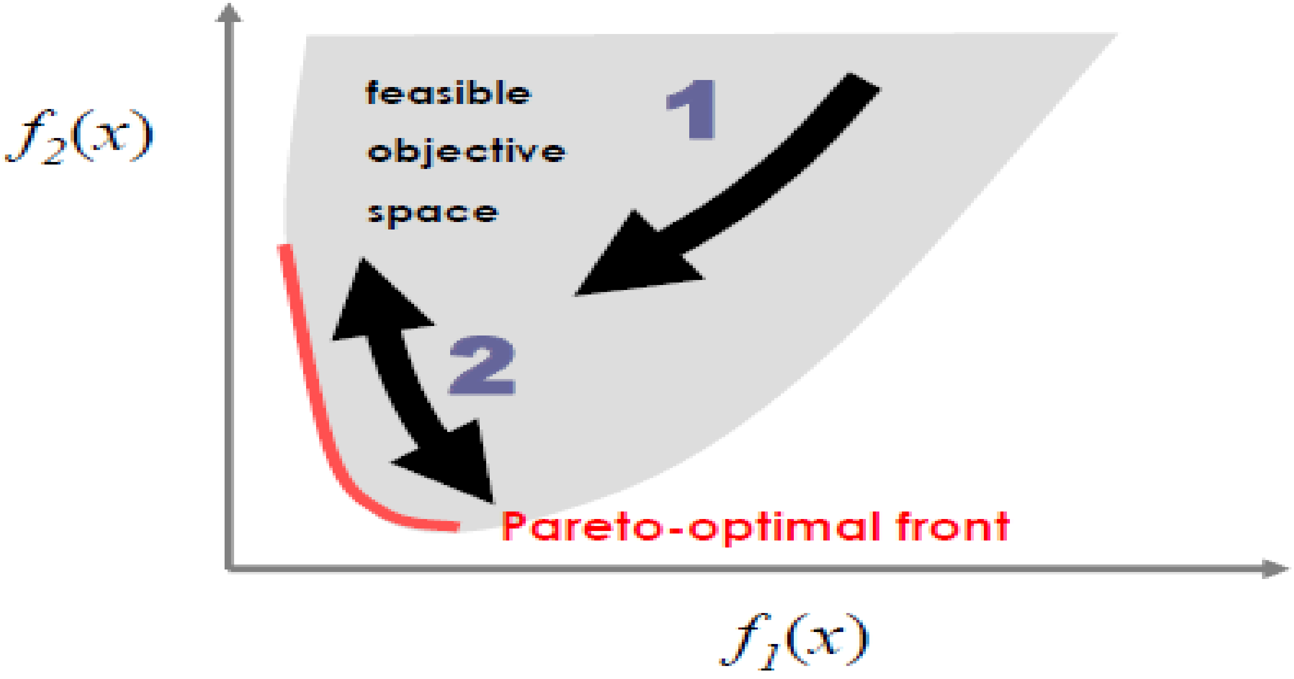
- **Non-dominated solution set**
  - Given a set of solutions, the non-dominated solution set is a set of all the solutions that are not dominated by any member of the solution set
- The non-dominated set of the entire feasible decision space is called the **Pareto-optimal set**
- The boundary defined by the set of all point mapped from the Pareto optimal set is called the **Pareto-optimal front**

# Graphical Depiction of Pareto Optimal Solution



# Goals in MOO

- Find set of solutions as close as possible to Pareto-optimal front
- To find a set of solutions as diverse as possible



# Illustrating Pareto-Optimal Solutions

**More Examples**

A cheaper but inconvenient flight

A convenient but expensive flight

# *Illustrating Pareto-Optimal Solutions*

- Airline routing often involves balancing cost and passenger convenience, exemplified by the trade-off between direct flights and stopovers.
- Airlines typically operate modular networks with main hubs, which help in reducing operational costs by concentrating services and staff at key locations.
- Passengers may prefer direct routes for convenience, but these often come at a higher cost due to increased operational expenses.
- The network design of airlines can vary from hub-centric models, like the one shown in Figure 7, to fully connected networks, as depicted in Figure 8, with direct flights between all airports.
- There is no single optimal solution to this trade-off between cost and convenience; instead, a spectrum of compromise solutions exists.
- Innovative airlines continuously seek compromises that provide passengers with more convenience while maintaining operational efficiency, resulting in evolving route networks.

# *Objectives in Multi-Objective Optimization*

- In multi-objective optimization, the search space can be divided into optimal and non-optimal regions, with the optimal region containing Pareto-optimal solutions.
- Even with more than two objectives, this division holds true, and conflicting objectives often result in multiple Pareto-optimal solutions.
- Without higher-level information, all Pareto-optimal solutions are equally important, emphasizing the need to find as many as possible.
- The primary goals in multi-objective optimization are to converge towards the Pareto-optimal front and to achieve diversity among solutions.
- Converging close to the true optimal solutions ensures near-optimality properties, similar to single-objective optimization.
- Diversity among solutions is crucial in multi-objective optimization to ensure a good set of trade-off solutions among objectives, which can be defined in both decision variable and objective spaces.

# *Non-Conflicting Objectives*

- Multiple Pareto-optimal solutions exist in a problem only when the objectives are conflicting.
- If the objectives are not conflicting, the Pareto-optimal set contains only one solution.
- For example, in a cantilever design problem where both end-deflection and maximum stress need to be minimized, the Pareto-optimal solution may reduce to a single solution.
- The example illustrates that when objectives are not conflicting, the minimum solution for one objective coincides with the minimum solution for another.
- In cases where it's not obvious whether objectives conflict, the resulting Pareto-optimal set may still contain only one optimal solution.
- Understanding the nature of objective conflicts is crucial for determining the cardinality of the Pareto-optimal set in a multi-objective optimization problem.

# Dominance and Pareto-Optimality

# Dominance and Pareto-Optimality

- Most multi-objective optimization algorithms use the concept of dominance in their search.
- Here, we define the concept of dominance and related terms and present a number of techniques for identifying dominated solutions in a finite population of solutions.
- We first define some special solutions which are often used In multi-objective optimization algorithms.

## Ideal Objective Vector

For each of the $M$ conflicting objectives, there exists one different optimal solution. An objective vector constructed with these individual optimal objective values constitutes the ideal objective vector.

**Definition 2.3.** *The* $m$*-th component of the ideal objective vector* $\mathbf{z}^*$ *is the constrained minimum solution of the following problem:*

$$\left.\begin{array}{ll} \textit{Minimize} & f_m(\mathbf{x}) \\ \textit{subject to} & \mathbf{x} \in \mathcal{S}. \end{array}\right\} \tag{2.5}$$

Thus, if the minimum solution for the $m$-th objective function is the decision vector $\mathbf{x}^{*(m)}$ with function value $f_m^*$, the ideal vector is as follows:



$$\mathbf{z}^* = \mathbf{f}^* = (f_1^*, f_2^*, \cdots, f_M^*)^T.$$

## Utopian Objective Vector

The ideal objective vector denotes an array of the lower bound of all objective functions. This means that for every objective function there exists at least one solution in the feasible search space sharing an identical value with the corresponding element in the ideal solution. Some algorithms may require a solution which has an objective value strictly better than (and not equal to) that of any solution in the search space. For this purpose, the utopian objective vector is defined as follows.

**Definition 2.4.** *A utopian objective vector* $z^{**}$ *has each of its components marginally smaller than that of the ideal objective vector, or* $z_i^{**} = z_i^{*} - \epsilon_i$ *with* $\epsilon_i > 0$ *for all* $i = 1, 2, \ldots, M$.

## Nadir Objective Vector

Unlike the ideal objective vector which represents the lower bound of each objective in the entire feasible search space, the nadir objective vector, $\mathbf{z}^{\mathrm{nad}}$, represents the upper bound of each objective in the entire Pareto-optimal set, and not in the entire search space. A nadir objective vector must not be confused with a vector of objectives (marked as 'W' in Figure 13) found by using the worst feasible function values, $f_i^{\max}$, in the entire search space.

Although the ideal objective vector is easy to compute (except in complex multi-modal objective problems), the nadir objective vector is difficult to compute in practice. However, for well-behaved problems (including linear MOOPs), the nadir objective vector can be derived from the ideal objective vector by using the *payoff table* method described in Miettinen (1999). For two objectives (Figure 13), if $\mathbf{z}^{*(1)} = \left(f_1(\mathbf{x}^{*(1)}), f_2(\mathbf{x}^{*(1)})\right)^T$ and $\mathbf{z}^{*(2)} = \left(f_1(\mathbf{x}^{*(2)}), f_2(\mathbf{x}^{*(2)})\right)^T$ are coordinates of the minimum solutions of $f_1$ and $f_2$, respectively, in the objective space, then the nadir objective vector can be estimated as $\mathbf{z}^{\mathrm{nad}} = \left(f_1(\mathbf{x}^{*(2)}), f_2(\mathbf{x}^{*(1)})\right)^T$.

The nadir objective vector may represent an existent or a non-existent solution, depending on the convexity and continuity of the Pareto-optimal set. In order to normalize each objective in the entire range of the Pareto-optimal region, the knowledge of nadir and ideal objective vectors can be used as follows:

$$f_i^{\mathrm{norm}} = \frac{f_i - z_i^*}{z_i^{\mathrm{nad}} - z_i^*} \qquad (2.6)$$

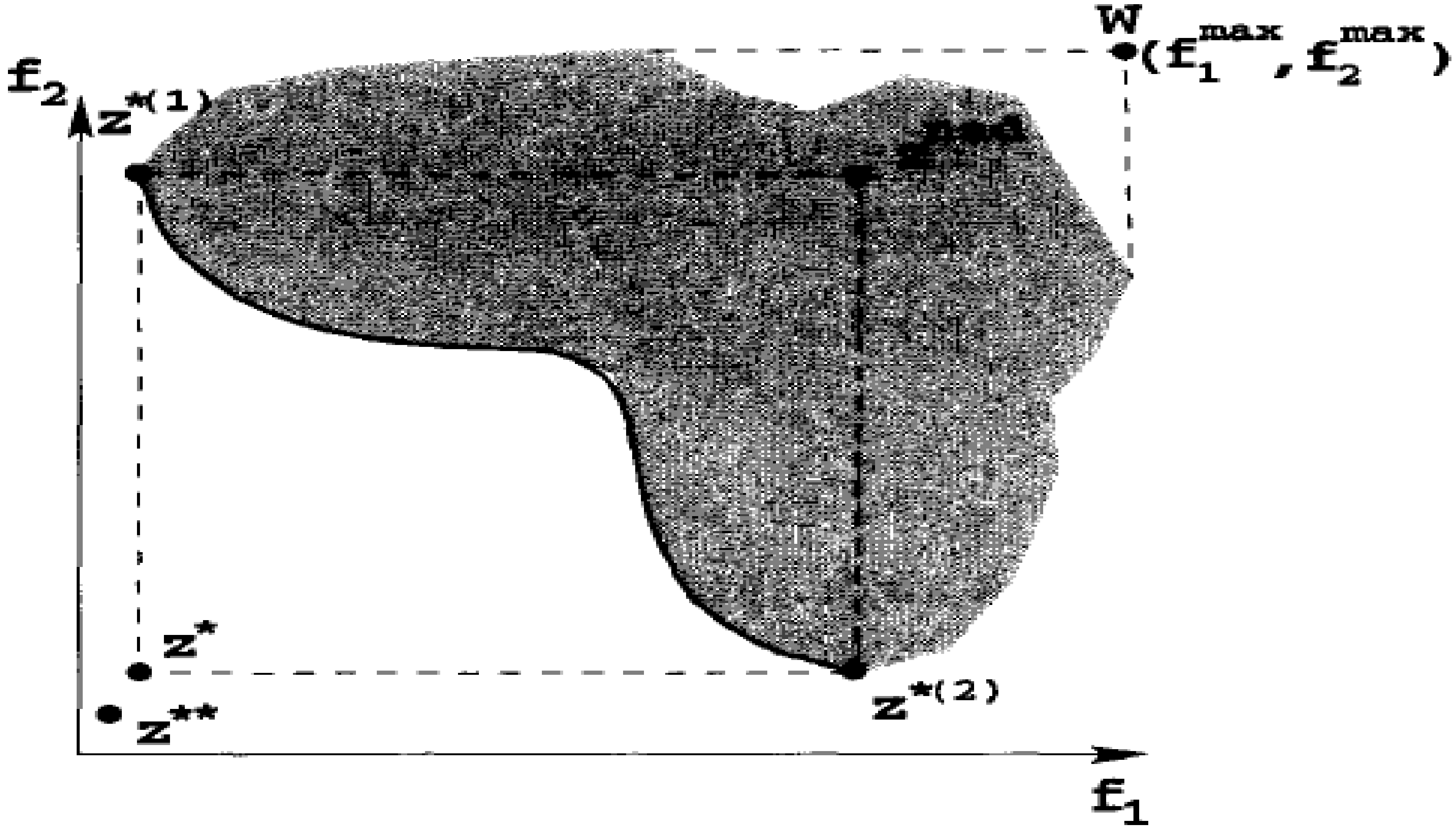



**Figure 13** The ideal, utopian, and nadir objective vectors.

# *Properties of Dominance Relation*

**Reflexive:** The dominance relation is *not reflexive*, since any solution p does not dominate itself according to Definition 2.5. The second condition of dominance relation in Definition 2.5 does not allow this property to be satisfied.

**Symmetric:** The dominance relation is also *not symmetric*, because $p \preceq q$ does not imply $q \preceq p$. In fact, the opposite is true. That is, if p dominates q, then q does not dominate p. Thus, the dominance relation is *asymmetric*.

**Antisymmetric:** Since the dominance relation is not symmetric, it cannot be antisymmetric as well.

**Transitive:** The dominance relation is *transitive*. This is because if $p \preceq q$ and $q \preceq r$, then $p \preceq r$.

There is another interesting property that the dominance relation possesses. If solution p does not dominate solution q, this does not imply that q dominates p.

In order for a binary relation to qualify as an ordering relation, it must be at least transitive (Chankong and Haimes, 1983). Thus, the dominance relation qualifies as an ordering relation. Since the dominance relation is not reflexive, it is a *strict partial order*. In general, if a relation is reflexive, antisymmetric, and transitive, it is loosely called a *partial order* and a set on which a partial order is defined is called a *partially ordered set*. However, it is important to note that the dominance relation is not reflexive and is not antisymmetric. Thus, the dominance relation is not a partial order relation in its general sense. The dominance relation is only a strict partial order relation.

# *Pareto-Optimality*

**Definition 2.7 (Globally Pareto-optimal set).** *The non-dominated set of the entire feasible search space $S$ is the globally Pareto-optimal set.*

On many occasions, the globally Pareto-optimal set is simply referred to as the Pareto-optimal set. Since solutions of this set are not dominated by any feasible member of the search space, they are optimal solutions of the MOOP. We define a locally Pareto-optimal set as follows (Deb, 1999c; Miettinen, 1999).

**Definition 2.8.** *If for every member* $\mathbf{x}$ *in a set* $\underline{\mathrm{P}}$ *there exists no solution* $\mathbf{y}$ *(in the neighborhood of* $\mathbf{x}$ *such that* $\|\mathbf{y} - \mathbf{x}\|_\infty \leq \epsilon$*, where* $\epsilon$ *is a small positive number) dominating any member of the set* $\underline{\mathrm{P}}$*, then solutions belonging to the set* $\underline{\mathrm{P}}$ *constitute a locally Pareto-optimal set.*

### 2.4.5 Strong Dominance and Weak Pareto-Optimality

The dominance relationship between the two solutions defined in Definition 2.5 is sometimes referred to as a *weak* dominance relation. This definition can be modified and a strong dominance relation can be defined as follows.

**Definition 2.9.** *A solution* $\mathbf{x}^{(1)}$ strongly *dominates a solution* $\mathbf{x}^{(2)}$ *(or* $\mathbf{x}^{(1)} \prec \mathbf{x}^{(2)}$*), if solution* $\mathbf{x}^{(1)}$ *is strictly better than solution* $\mathbf{x}^{(2)}$ *in all* $M$ *objectives.*

Referring to Figure 14, we now observe that solution 5 does not strongly dominate solution 1, although we have seen earlier that solution 5 weakly dominates solution 1. However, solution 3 strongly dominates solution 1, since solution 3 is better than solution 1 in both objectives. Thus, if a solution $\mathbf{x}^{(1)}$ strongly dominates a solution $\mathbf{x}^{(2)}$, the solution $\mathbf{x}^{(1)}$ also weakly dominates solution $\mathbf{x}^{(2)}$, but not vice versa. The strong dominance operator has the same properties as that described in Section 2.4.3 for the weak dominance operator.

The above definition of strong dominance can be used to define a *weakly non-dominated set.*

**Definition 2.10 (Weakly non-dominated set).** *Among a set of solutions* $P$*, the weakly non-dominated set of solutions* $P'$ *are those that are not strongly dominated by any other member of the set* $P$.

# Optimality Conditions

**Theorem 2.5.1.** *(Fritz–John necessary condition). A necessary condition for* $\mathbf{x}^*$ *to be Pareto-optimal is that there exist vectors* $\boldsymbol{\lambda} \geq \mathbf{0}$ *and* $\mathbf{u} \geq \mathbf{0}$ *(where* $\boldsymbol{\lambda} \in \mathbb{R}^M$, $\mathbf{u} \in \mathbb{R}^J$ *and* $\boldsymbol{\lambda}, \mathbf{u} \neq \mathbf{0}$*) such that the following conditions are true:*

1. $\sum_{m=1}^{M} \lambda_m \nabla f_m(\mathbf{x}^*) - \sum_{j=1}^{J} u_j \nabla g_j(\mathbf{x}^*) = \mathbf{0}$, *and*
2. $u_j g_j(\mathbf{x}^*) = 0$ *for all* $j = 1, 2, \ldots, J$.

For a proof, readers may refer to Cunha and Polak (1967). Miettinen (1999) argues that the above theorem is also valid as the necessary condition for a solution to be weakly Pareto-optimal. Those readers familiar with the Kuhn–Tucker necessary conditions for single-objective optimization will immediately recognize the similarity between the above conditions and that of the single-objective optimization. The difference is in the inclusion of a $\lambda$-vector with the gradient vector of the objectives.

For an unconstrained MOOP, the above theorem requires the following condition:

$$\sum_{m=1}^{M} \lambda_m \nabla f_m(\mathbf{x}^*) = \mathbf{0}$$

to be necessary for a solution to be Pareto-optimal. Writing the above vector equation in matrix form, we have the following necessary condition for an $M$-objective and $n$-variable unconstrained MOOP:

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_M}{\partial x_1} \\ \frac{\partial f_1}{\partial x_2} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_M}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_1}{\partial x_n} & \frac{\partial f_2}{\partial x_n} & \cdots & \frac{\partial f_M}{\partial x_n} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_M \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.7}$$

**Theorem 2.5.2.** *(Karush–Kuhn–Tucker sufficient condition for Pareto-optimality). Let the objective functions be convex and the constraint functions of the problem shown in equation (2.1) be nonconvex. Let the objective and constraint functions be continuously differentiable at a feasible solution* $\mathbf{x}^*$. *A sufficient condition for* $\mathbf{x}^*$ *to be Pareto-optimal is that there exist vectors* $\boldsymbol{\lambda} > \mathbf{0}$ *and* $\mathbf{u} \geq \mathbf{0}$ *(where* $\boldsymbol{\lambda} \in \mathbb{R}^M$ *and* $\mathbf{u} \in \mathbb{R}^J$*) such that the following equations are true:*

1. $\sum_{m=1}^{M} \lambda_m \nabla f_i(\mathbf{x}^*) - \sum_{j=1}^{J} u_j \nabla g_j(\mathbf{x}^*) = \mathbf{0}$, *and*
2. $u_j g_j(\mathbf{x}^*) = 0$ *for all* $j = 1, 2, \ldots, J$.

For a proof, see Miettinen (1999). If the objective functions and constraints are not convex, the above theorem does not hold. However, for pseudo-convex and non-differentiable problems, different necessary and sufficient conditions do exist (Bhatia and Aggarwal, 1992).