Multiobjective Optim. results and approaches

Optimization of Complex Systems – March 10th 2026

Andrea Brilli

Sapienza University of Rome

Recap: where we are

In the previous class we introduced:

- The multiobjective problem $\min_{x\in\mathbb{R}^n} f(x) = (f_1(x), \ldots, f_p(x))^\top$
- Pareto dominance ($\preceq$) and strict dominance ($\prec$) between objective vectors
- The Pareto front Ω_P and the weak Pareto front Ω_D in objective space
- First-order optimality via the stationarity measure $\theta(x)$

Recap: where we are

In the previous class we introduced:

- The multiobjective problem $\min_{x \in \mathbb{R}^n} f(x) = (f_1(x), \ldots, f_p(x))^\top$
- Pareto dominance ($\preceq$) and strict dominance ($\prec$) between objective vectors
- The Pareto front Ω_P and the weak Pareto front Ω_D in objective space
- First-order optimality via the stationarity measure $\theta(x)$

Today we take a step back and study the geometry of the **objective space** more carefully:

- What is the structure of the feasible objective region $\mathcal{Z} = f(\mathbb{R}^n)$?
- Where, within $\mathcal{Z}$, can nondominated points lie?
- When is the Pareto front guaranteed to be non-empty?

The dominance cone

Pareto dominance has a natural geometric description in $\mathbb{R}^p$.

Definition (Dominance cone)

The **nonnegative orthant** *of* $\mathbb{R}^p$ *is*

$$\mathbb{R}^p_{\geq} = \{z \in \mathbb{R}^p : z_i \geq 0,\ i = 1, \ldots, p\}.$$

The dominance cone

Pareto dominance has a natural geometric description in $\mathbb{R}^p$.

Definition (Dominance cone)

The **nonnegative orthant** *of* $\mathbb{R}^p$ *is*

$$\mathbb{R}^p_{\geq} = \{z \in \mathbb{R}^p : z_i \geq 0,\ i = 1, \ldots, p\}.$$

Given two objective vectors $z^1, z^2 \in \mathbb{R}^p$, the dominance relations can be rewritten as:

$$z^1 \preceq z^2 \iff z^2 - z^1 \in \mathbb{R}^p_{\geq} \setminus \{0\} \qquad \text{(Pareto dominance)}$$

$$z^1 \prec z^2 \iff z^2 - z^1 \in \operatorname{int} \mathbb{R}^p_{\geq} \qquad \text{(strict dominance)}$$

The dominance cone

Pareto dominance has a natural geometric description in $\mathbb{R}^p$.

Definition (Dominance cone)

The **nonnegative orthant** *of* $\mathbb{R}^p$ *is*

$$\mathbb{R}^p_{\geq} = \{z \in \mathbb{R}^p : z_i \geq 0,\ i = 1, \ldots, p\}.$$

Given two objective vectors $z^1, z^2 \in \mathbb{R}^p$, the dominance relations can be rewritten as:

$$z^1 \preceq z^2 \iff z^2 - z^1 \in \mathbb{R}^p_{\geq} \setminus \{0\} \qquad \text{(Pareto dominance)}$$

$$z^1 \prec z^2 \iff z^2 - z^1 \in \operatorname{int} \mathbb{R}^p_{\geq} \qquad \text{(strict dominance)}$$

Question: what does the set of all points dominated by a given $z \in \mathbb{R}^p$ look like geometrically?

The dominance cone

Pareto dominance has a natural geometric description in $\mathbb{R}^p$.

Definition (Dominance cone)

The **nonnegative orthant** *of* $\mathbb{R}^p$ *is*

$$\mathbb{R}^p_{\geq} = \{z \in \mathbb{R}^p : z_i \geq 0, \;\; i = 1, \ldots, p\}.$$

Given two objective vectors $z^1, z^2 \in \mathbb{R}^p$, the dominance relations can be rewritten as:

$$z^1 \preceq z^2 \iff z^2 - z^1 \in \mathbb{R}^p_{\geq} \setminus \{0\} \qquad \text{(Pareto dominance)}$$

$$z^1 \prec z^2 \iff z^2 - z^1 \in \operatorname{int} \mathbb{R}^p_{\geq} \qquad \text{(strict dominance)}$$

Question: what does the set of all points dominated by a given $z \in \mathbb{R}^p$ look like geometrically?

It is the translated cone $z + \mathbb{R}^p_{\geq}$ — the "upper-right" quadrant rooted at z (in $p = 2$). A point z^1 dominates z if and only if $z \in z^1 + \mathbb{R}^p_{\geq}$.

The upward closure of $\mathcal{Z}$

Definition (Upward closure)

Given the feasible objective region $\mathcal{Z} \subset \mathbb{R}^p$, its **upward closure** *is*

$$\mathcal{Z} + \mathbb{R}^p_{\geq} = \{z + d : z \in \mathcal{Z},\ d \in \mathbb{R}^p_{\geq}\}.$$

Intuitively, $\mathcal{Z} + \mathbb{R}^p_{\geq}$ is the set of all points that are **dominated by or equal to** some point in $\mathcal{Z}$ — everything "above and to the right" of $\mathcal{Z}$.

The upward closure of $\mathcal{Z}$

Definition (Upward closure)

Given the feasible objective region $\mathcal{Z} \subset \mathbb{R}^p$, its **upward closure** *is*

$$\mathcal{Z} + \mathbb{R}^p_{\geq} = \{z + d : z \in \mathcal{Z},\ d \in \mathbb{R}^p_{\geq}\}.$$

Intuitively, $\mathcal{Z} + \mathbb{R}^p_{\geq}$ is the set of all points that are **dominated by or equal to** some point in $\mathcal{Z}$ — everything "above and to the right" of $\mathcal{Z}$.

Proposition (Prop. 2.3, Ehrgott 2005)

$$\mathcal{Z}_N = \left(\mathcal{Z} + \mathbb{R}^p_{\geq}\right)_N$$

i.e., the nondominated set of $\mathcal{Z}$ and the nondominated set of its upward closure coincide.

Consequence: using the upward closure of $\mathcal{Z}$ does not change the Pareto front. This justifies working with $\mathcal{Z} + \mathbb{R}^p_{\geq}$ instead of $\mathcal{Z}$ whenever it is more convenient — for instance, to obtain a **closed or convex** set.

Proof of Proposition 2.3

We need to show $\mathcal{Z}_N = (\mathcal{Z} + \mathbb{R}^p_{\geq})_N$.

Proof of Proposition 2.3

We need to show $\mathcal{Z}_N = (\mathcal{Z} + \mathbb{R}^p_{\geq})_N$.

Key observation: every point in $(\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$ is dominated.
Indeed, if $z \in (\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$, then by definition $z = z' + d$ for some $z' \in \mathcal{Z}$ and $0 \neq d \in \mathbb{R}^p_{\geq}$, so $z' \preceq z$.

Proof of Proposition 2.3

We need to show $\mathcal{Z}_N = (\mathcal{Z} + \mathbb{R}^p_{\geq})_N$.

Key observation: every point in $(\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$ is dominated. Indeed, if $z \in (\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$, then by definition $z = z' + d$ for some $z' \in \mathcal{Z}$ and $0 \neq d \in \mathbb{R}^p_{\geq}$, so $z' \preceq z$.

Consequence: $(\mathcal{Z} + \mathbb{R}^p_{\geq})_N \subseteq \mathcal{Z}$. That is, nondominated points of the upward closure must already lie in $\mathcal{Z}$.

Proof of Proposition 2.3

We need to show $\mathcal{Z}_N = (\mathcal{Z} + \mathbb{R}^p_{\geq})_N$.

Key observation: every point in $(\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$ is dominated. Indeed, if $z \in (\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$, then by definition $z = z' + d$ for some $z' \in \mathcal{Z}$ and $0 \neq d \in \mathbb{R}^p_{\geq}$, so $z' \preceq z$.

Consequence: $(\mathcal{Z} + \mathbb{R}^p_{\geq})_N \subseteq \mathcal{Z}$. That is, nondominated points of the upward closure must already lie in $\mathcal{Z}$.

Conclusion: since the nondominated points of $\mathcal{Z} + \mathbb{R}^p_{\geq}$ all lie in $\mathcal{Z}$, and $\mathcal{Z} \subseteq \mathcal{Z} + \mathbb{R}^p_{\geq}$, taking the nondominated set of either gives the same result $(\mathcal{Z} + \mathbb{R}^p_{\geq})_N = \mathcal{Z}_N$. □

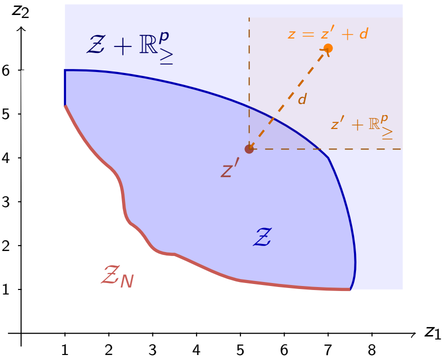
Proof of Proposition 2.3

We need to show $\mathcal{Z}_N = (\mathcal{Z} + \mathbb{R}^p_{\geq})_N$.

Key observation: every point in $(\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$ is dominated. Indeed, if $z \in (\mathcal{Z} + \mathbb{R}^p_{\geq}) \setminus \mathcal{Z}$, then by definition $z = z' + d$ for some $z' \in \mathcal{Z}$ and $0 \neq d \in \mathbb{R}^p_{\geq}$, so $z' \preceq z$.

Consequence: $(\mathcal{Z} + \mathbb{R}^p_{\geq})_N \subseteq \mathcal{Z}$. That is, nondominated points of the upward closure must already lie in $\mathcal{Z}$.

Conclusion: since the nondominated points of $\mathcal{Z} + \mathbb{R}^p_{\geq}$ all lie in $\mathcal{Z}$, and $\mathcal{Z} \subseteq \mathcal{Z} + \mathbb{R}^p_{\geq}$, taking the nondominated set of either gives the same result $(\mathcal{Z} + \mathbb{R}^p_{\geq})_N = \mathcal{Z}_N$. □



Nondominated points lie on the boundary

Proposition (Prop. 2.4, Ehrgott 2005)

$\mathcal{Z}_N \subset \mathrm{bd}(\mathcal{Z})$.

Nondominated points lie on the boundary

Proposition (Prop. 2.4, Ehrgott 2005)

$\mathcal{Z}_N \subset \mathrm{bd}(\mathcal{Z})$.

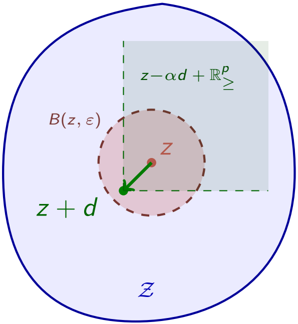
Proof: Let $z \in \mathcal{Z}_N$ and suppose $z \notin \mathrm{bd}(\mathcal{Z})$, i.e., $z \in \mathrm{int}(\mathcal{Z})$.
Then there exists $\varepsilon > 0$ such that $B(z, \varepsilon) \subset \mathcal{Z}$.
Let $d \neq 0$, $d \in \mathbb{R}^p_{\geq}$. Choose $\alpha \in \mathbb{R}$ with $0 < \alpha < \varepsilon / \|d\|$, so that $\alpha d \in B(0, \varepsilon)$.
Then $z - \alpha d \in B(z, \varepsilon) \subset \mathcal{Z}$, and $z - \alpha d \preceq z$ since $\alpha d \in \mathbb{R}^p_{\geq} \setminus \{0\}$.
This contradicts $z \in \mathcal{Z}_N$.

Nondominated points lie on the boundary

Proposition (Prop. 2.4, Ehrgott 2005)

$\mathcal{Z}_N \subset \mathrm{bd}(\mathcal{Z})$.

Proof: Let $z \in \mathcal{Z}_N$ and suppose $z \notin \mathrm{bd}(\mathcal{Z})$, i.e., $z \in \mathrm{int}(\mathcal{Z})$.
Then there exists $\varepsilon > 0$ such that $B(z, \varepsilon) \subset \mathcal{Z}$.
Let $d \neq 0$, $d \in \mathbb{R}^p_{\geq}$. Choose $\alpha \in \mathbb{R}$ with $0 < \alpha < \varepsilon / \|d\|$, so that $\alpha d \in B(0, \varepsilon)$.
Then $z - \alpha d \in B(z, \varepsilon) \subset \mathcal{Z}$, and $z - \alpha d \preceq z$ since $\alpha d \in \mathbb{R}^p_{\geq} \setminus \{0\}$.
This contradicts $z \in \mathcal{Z}_N$.



$z + d \prec z$

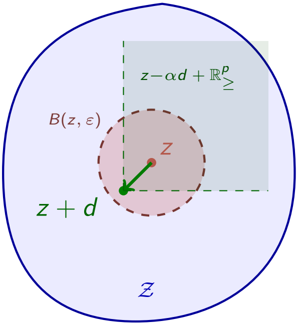
Nondominated points lie on the boundary

Proposition (Prop. 2.4, Ehrgott 2005)

$\mathcal{Z}_N \subset \mathrm{bd}(\mathcal{Z})$.

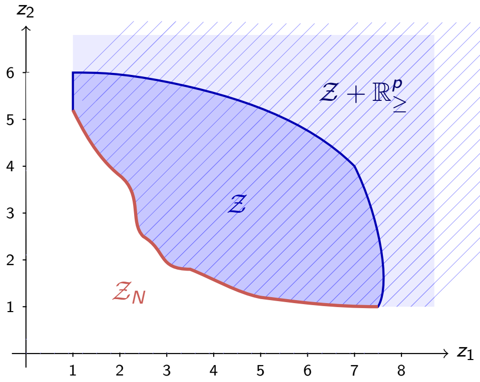
Proof: Let $z \in \mathcal{Z}_N$ and suppose $z \notin \mathrm{bd}(\mathcal{Z})$, i.e., $z \in \mathrm{int}(\mathcal{Z})$.
Then there exists $\varepsilon > 0$ such that $B(z, \varepsilon) \subset \mathcal{Z}$.
Let $d \neq 0$, $d \in \mathbb{R}^p_{\geq}$. Choose $\alpha \in \mathbb{R}$ with $0 < \alpha < \varepsilon / \|d\|$, so that $\alpha d \in B(0, \varepsilon)$.
Then $z - \alpha d \in B(z, \varepsilon) \subset \mathcal{Z}$, and $z - \alpha d \preceq z$ since $\alpha d \in \mathbb{R}^p_{\geq} \setminus \{0\}$.
This contradicts $z \in \mathcal{Z}_N$.

Corollary 2.5. If $\mathcal{Z}$ is open, or if $\mathcal{Z} + \mathbb{R}^p_{\geq}$ is open, then $\mathcal{Z}_N = \emptyset$.



$z + d \prec z$

Illustration: $\mathcal{Z}$ and its upward closure



The nondominated set $\mathcal{Z}_N$ (red) lies on the **lower-left boundary** of $\mathcal{Z}$ (Prop. 2.4), and is unchanged when passing to the upward closure $\mathcal{Z} + \mathbb{R}^p_{\geq}$ (Prop. 2.3).

Problem setting

We consider the **unconstrained multiobjective problem**

$$\min_{x \in \mathbb{R}^n} f(x) = \big(f_1(x), \dots, f_p(x)\big)^\top$$

where each $f_i : \mathbb{R}^n \to \mathbb{R}$ is **convex** and **continuously differentiable**.

Proposition

For a convex multiobjective problem, every local Pareto optimum is also a global Pareto optimum.

Local implies global

Let x^* be a local Pareto optimum. Then $\exists \, \epsilon > 0$ such that

$$\nexists \, x \in B_\epsilon(x^*) : f(x) \preceq f(x^*).$$

Suppose by contradiction that x^* is **not** a global Pareto optimum.

Local implies global

Let x^* be a local Pareto optimum. Then $\exists\, \epsilon > 0$ such that

$$\nexists\, x \in B_\epsilon(x^*) : f(x) \preceq f(x^*).$$

Suppose by contradiction that x^* is **not** a global Pareto optimum. Then $\exists\, x^\circ \in \mathbb{R}^n$ with $f(x^\circ) \preceq f(x^*)$, i.e., $f_i(x^\circ) \leq f_i(x^*)$ for all i, and for at least one index j: $f_j(x^\circ) < f_j(x^*)$.

Consider the convex combination $x_\beta = \beta x^\circ + (1-\beta)x^*$, $\beta \in [0,1]$.
By definition, we can choose $\bar{\beta}$ small enough so that $\hat{x} := x_{\bar{\beta}} \in B_\epsilon(x^*)$.

By convexity of each f_i:

$$f_i(\hat{x}) \leq \bar{\beta}\, f_i(x^\circ) + (1-\bar{\beta})\, f_i(x^*) \leq f_i(x^*), \quad \text{and} \quad f_i(\hat{x}) < f_i(x^*),$$

which implies $f(\hat{x}) \preceq f(x^*)$. This contradicts local optimality of x^* since $\hat{x} \in B_\epsilon(x^*)$.

Sufficient condition for Pareto optimality

Proposition

A sufficient condition for $\bar{x} \in \mathbb{R}^n$ *to be a Pareto optimum is that there exist multipliers* $\bar{\lambda} > 0$ *(all strictly positive) such that*

$$\sum_{i=1}^{p} \bar{\lambda}_i \, \nabla f_i(\bar{x}) = 0.$$

Sufficient condition for Pareto optimality

Proposition

A sufficient condition for $\bar{x} \in \mathbb{R}^n$ *to be a Pareto optimum is that there exist multipliers* $\bar{\lambda} > 0$ *(all strictly positive) such that*

$$\sum_{i=1}^{p} \bar{\lambda}_i \nabla f_i(\bar{x}) = 0.$$

Construct the scalarized objective $\varphi(x) = \sum_{i=1}^{p} \bar{\lambda}_i f_i(x)$. Since each f_i is convex and $\bar{\lambda}_i > 0$, the function φ is convex.

The condition $\sum_i \bar{\lambda}_i \nabla f_i(\bar{x}) = 0$ is equivalent to $\nabla\varphi(\bar{x}) = 0$, which for a convex function is **sufficient for global optimality**. Hence $\bar{x}$ is a global minimizer of φ.

Sufficient condition for Pareto optimality

Proposition

A sufficient condition for $\bar{x} \in \mathbb{R}^n$ to be a Pareto optimum is that there exist multipliers $\bar{\lambda} > 0$ (all strictly positive) such that

$$\sum_{i=1}^{p} \bar{\lambda}_i \nabla f_i(\bar{x}) = 0.$$

Construct the scalarized objective $\varphi(x) = \sum_{i=1}^{p} \bar{\lambda}_i f_i(x)$. Since each f_i is convex and $\bar{\lambda}_i > 0$, the function φ is convex.

The condition $\sum_i \bar{\lambda}_i \nabla f_i(\bar{x}) = 0$ is equivalent to $\nabla\varphi(\bar{x}) = 0$, which for a convex function is **sufficient for global optimality**. Hence $\bar{x}$ is a global minimizer of φ.

Suppose by contradiction $\bar{x}$ is not a Pareto optimum. Then $\exists\, \tilde{x}$ with $f(\tilde{x}) \preceq f(\bar{x})$, i.e., $f_i(\tilde{x}) \leq f_i(\bar{x})$ for all i, strict for some j. Multiplying by $\bar{\lambda}_i > 0$ and summing:

$$\varphi(\tilde{x}) = \sum_{i=1}^{p} \bar{\lambda}_i f_i(\tilde{x}) < \sum_{i=1}^{p} \bar{\lambda}_i f_i(\bar{x}) = \varphi(\bar{x}),$$

contradicting the global optimality of $\bar{x}$ for φ.

Why $\bar{\lambda} > 0$?

Key remark: the strict positivity $\bar{\lambda} > 0$ is **essential**.

The dominance $f(\tilde{x}) \preceq f(\bar{x})$ guarantees:

- $f_i(\tilde{x}) \leq f_i(\bar{x})$ for **all** i
- $f_j(\tilde{x}) < f_j(\bar{x})$ for **at least one** j

When we multiply by $\bar{\lambda}_i$ and sum:

$$\varphi(\tilde{x}) - \varphi(\bar{x}) = \sum_i \bar{\lambda}_i \big(f_i(\tilde{x}) - f_i(\bar{x})\big)$$

- Each term $\bar{\lambda}_i(f_i(\tilde{x}) - f_i(\bar{x})) \leq 0$
- The j-th term is strictly negative: $\bar{\lambda}_j(f_j(\tilde{x}) - f_j(\bar{x})) < 0$ **only if** $\bar{\lambda}_j > 0$

Conclusion: without $\bar{\lambda} > 0$, the single strict inequality might be "killed" by a zero multiplier and the contradiction would fail.

Sufficient condition for weak Pareto optimality

Proposition

A sufficient condition for $\bar{x} \in \mathbb{R}^n$ to be a **weakly** *Pareto optimal point is that there exist multipliers $\bar{\lambda} \geq 0$, $\bar{\lambda} \neq 0$, such that*

$$\sum_{i=1}^{p} \bar{\lambda}_i \nabla f_i(\bar{x}) = 0.$$

Construct $\varphi(x) = \sum_i \bar{\lambda}_i f_i(x)$ as before. Again φ is convex and $\bar{x}$ is its global minimizer.

Suppose by contradiction $\bar{x}$ is not weakly Pareto optimal. Then $\exists\, \tilde{x}$ with $f(\tilde{x}) \prec f(\bar{x})$, i.e., $f_i(\tilde{x}) < f_i(\bar{x})$ for **all** i. Multiplying by $\bar{\lambda}_i \geq 0$ (not all zero) and summing:

$$\varphi(\tilde{x}) < \varphi(\bar{x}),$$

contradicting the global optimality of $\bar{x}$.

Note: why does $\bar{\lambda} \geq 0$ suffice here? Because $\prec$ gives strict inequality in **all** components — even a single nonzero $\bar{\lambda}_i$ guarantees a strict decrease in φ.

Comparing the two sufficient conditions

The only difference between Propositions 2 and 3 is the strength of the dominance assumed in the contradiction:

Optimality	Contradiction uses	Multipliers needed
Pareto	$f(\tilde{x}) \preceq f(\bar{x})$, $\exists j$ strict	$\bar{\lambda} > 0$
Weak Pareto	$f(\tilde{x}) \prec f(\bar{x})$, all strict	$\bar{\lambda} \geq 0$, $\bar{\lambda} \neq 0$

Geometric intuition:

- For Pareto, the dominating point only needs to improve **at least one** objective — a zero multiplier on that component would miss it
- For weak Pareto, the dominating point improves **all** objectives — so any nonzero multiplier is enough to detect the improvement in φ

Necessary and sufficient condition

Proposition

For the convex unconstrained multiobjective problem, x^* is a **weakly Pareto optimal** *point if and only if there exist multipliers $\lambda^* \geq 0$, $\lambda^* \neq 0$, such that*

$$\sum_{i=1}^{p} \lambda_i^* \nabla f_i(x^*) = 0.$$

- **($\Leftarrow$) Sufficiency:** Proposition 3 above.
- **($\Rightarrow$) Necessity:** Recall from the general (non-convex) theory that a necessary condition for weak Pareto optimality is

$$\theta(x^*) = 0, \qquad \text{where } \theta(x) = \min_{d \in \mathbb{R}^n} \max_{i=1,\ldots,p} \nabla f_i(x)^\top d.$$

  The condition $\theta(x^*) = 0$ is equivalent to $0 \in \mathrm{conv}\{\nabla f_1(x^*), \ldots, \nabla f_p(x^*)\}$, i.e., $\exists\, \lambda^* \geq 0$, $\lambda^* \neq 0$, $\sum_i \lambda_i^* = 1$ with $\sum_i \lambda_i^* \nabla f_i(x^*) = 0$.

Key point: convexity makes this **also sufficient** (via Proposition 3), closing the gap between necessary and sufficient conditions.

Summary

Optimality	Condition	Multipliers
Pareto optimum (suff.)	$\sum_i \bar{\lambda}_i \nabla f_i(\bar{x}) = 0$	$\bar{\lambda} > 0$
Weak Pareto optimum (nec. & suff.)	$\sum_i \bar{\lambda}_i \nabla f_i(\bar{x}) = 0$	$\bar{\lambda} \geq 0,\ \bar{\lambda} \neq 0$

Take-home message:

- In the convex case, scalarization with $\bar{\lambda} > 0$ is a **complete certificate** for Pareto optimality
- For weak optimality, convexity gives a clean **iff** characterization via the stationary condition
- The distinction $\bar{\lambda} > 0$ vs. $\bar{\lambda} \geq 0$ tracks the distinction $\preceq$ vs. $\prec$ in dominance

Multiobjective portfolio optimization

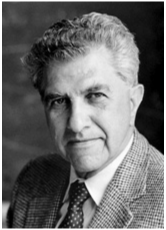
The portfolio selection problem formulated by Markowitz (mean-variance model) is a special case of a **two-objective optimization problem**:

$$
\begin{aligned}
\max \quad & \left(\mu^\top x,\ -x^\top Q x\right) \\
s.t. \quad & \sum_{i=1}^{n} x_i = 1 \\
& 0 \leq x_i \leq 1, \quad i = 1, \ldots, n.
\end{aligned}
$$

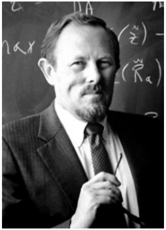
A Nobel-worthy idea

In 1990, the Prize in Memory of A. Nobel was jointly awarded to:

Harry M. Markowitz



Merton H. Miller



William F. Sharpe

Portfolio selection

Markowitz observed that the typical behavior of a financial investor is to *diversify* their portfolio of assets.

"The hypothesis that the investor does maximize discounted return must be rejected [...] the foregoing rule never implies that there is a diversified portfolio which is preferable to all non-diversified portfolios."

"There is a rule which implies both that the investor should diversify and that he should maximize expected return. [...] This rule is a special case of the expected returns–variance of returns rule."

Portfolio selection

Consider n assets on the financial market, and let R_i be the random variable representing the return of the i-th asset.

Let μ_i be the expected return of R_i, and σ_{ij} the covariance between R_i and R_j (so $\sigma_{ii} = \sigma_i^2$ is the variance of R_i).

Let x_i denote the fraction of capital invested in asset i, and assume all capital is invested, so

$$\sum_{i=1}^{n} x_i = 1.$$

N.B. the x_i are quantities chosen by the investor — they are **not** random variables.

Given the x_i, the portfolio return **is a random variable**, a combination of the R_i with coefficients x_i:

$$R = \sum_{i=1}^{n} x_i R_i.$$

Portfolio selection

Consider n assets on the financial market, and let R_i be the random variable representing the return of the i-th asset.

Let μ_i be the expected return of R_i, and σ_{ij} the covariance between R_i and R_j (so $\sigma_{ii} = \sigma_i^2$ is the variance of R_i).

Let x_i denote the fraction of capital invested in asset i, and assume all capital is invested, so

$$\sum_{i=1}^{n} x_i = 1.$$

N.B. the x_i are quantities chosen by the investor — they are **not** random variables.

Given the x_i, the portfolio return **is a random variable**, a combination of the R_i with coefficients x_i:

$$R = \sum_{i=1}^{n} x_i R_i.$$

Portfolio selection

Consider n assets on the financial market, and let R_i be the random variable representing the return of the i-th asset.

Let μ_i be the expected return of R_i, and σ_{ij} the covariance between R_i and R_j (so $\sigma_{ii} = \sigma_i^2$ is the variance of R_i).

Let x_i denote the fraction of capital invested in asset i, and assume all capital is invested, so

$$\sum_{i=1}^{n} x_i = 1.$$

N.B. the x_i are quantities chosen by the investor — they are **not** random variables.

Given the x_i, the portfolio return **is a random variable**, a combination of the R_i with coefficients x_i:

$$R = \sum_{i=1}^{n} x_i R_i.$$

Portfolio selection

Consider n assets on the financial market, and let R_i be the random variable representing the return of the i-th asset.

Let μ_i be the expected return of R_i, and σ_{ij} the covariance between R_i and R_j (so $\sigma_{ii} = \sigma_i^2$ is the variance of R_i).

Let x_i denote the fraction of capital invested in asset i, and assume all capital is invested, so

$$\sum_{i=1}^{n} x_i = 1.$$

N.B. the x_i are quantities chosen by the investor — they are **not** random variables.

Given the x_i, the portfolio return **is a random variable**, a combination of the R_i with coefficients x_i:

$$R = \sum_{i=1}^{n} x_i R_i.$$

Portfolio return: mean and variance

One can verify that the **expected return** and **variance** of the portfolio are:

$$\mu_R(x) = \sum_{i=1}^{n} x_i \mu_{R_i} = \mu^\top x,$$

$$\sigma_R^2(x) = x^\top Q x,$$

where $Q \in \mathbb{R}^{n \times n}$ is the **covariance matrix**, with entries $q_{ij} = \sigma_{ij}$.

Question: given this structure, how should a rational investor choose x?

Multiobjective formulation

Markowitz showed that the observed behavior of investors can be explained by considering the following **multiobjective problem**:

$$\max\ \mu_R(x), \quad \min\ \sigma_R^2(x)$$

$$s.t. \quad \sum_{i=1}^{n} x_i = 1$$

$$x_i \geq 0, \quad i = 1, \ldots, n,$$

in which one wishes to *simultaneously* maximize expected return and minimize variance. For this reason, the model is also known as the *mean-variance* model.

Key insight: the two objectives are in conflict — higher expected return typically comes at the cost of higher risk.

The problem

Given:

- $\mu \in \mathbb{R}^n$: vector of expected returns
- $Q \in \mathbb{R}^{n \times n}$: covariance matrix (symmetric positive semidefinite)

Define the **mean-variance multiobjective problem**:

$$\max \ \mu^\top x, \quad \min \ \frac{1}{2} x^\top Q x$$
$$s.t. \qquad e^\top x = 1, \quad x \geq \mathbf{0}_n$$

where $e = (1,\ 1, \ldots, 1)^\top \in \mathbb{R}^n$ is the vector of all ones.

Numerical example

Setting:

- We consider $n = 50$ assets from a financial market
- Each asset i has an estimated expected return μ_i and pairwise covariances σ_{ij}

Questions:

- How do we compute the **ideal objective vector** z^{id}?
- How can we compute **Pareto optimal solutions**?

Computing the ideal vector z^{id}

Solve the two **single-objective** problems separately:

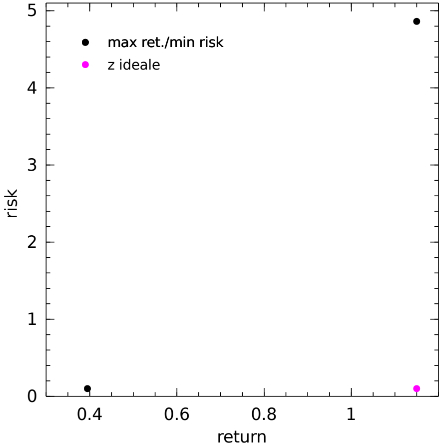
$$\begin{aligned} \max \quad & \mu^\top x \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned} \quad \Downarrow \quad x^*_\mu$$

$$\begin{aligned} \min \quad & \frac{1}{2} x^\top Q x \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned} \quad \Downarrow \quad x^*_\sigma$$

The ideal vector is then:

$$z^{id} = \left(\mu^\top x^*_\mu,\ \frac{1}{2} (x^*_\sigma)^\top Q x^*_\sigma \right)^\top$$

Ideal vector: graphical illustration



The “decision maker”

- Single objective: if x^* and $\tilde{x}$ are both global optima, then $f(x^*) = f(\tilde{x})$ and there is no (reasonable) ground for preferring one over the other
- Multiobjective: if x^* and $\tilde{x}$ are both (global) Pareto optima, then necessarily

$$f(x^*) \not\preceq f(\tilde{x})$$
$$f(\tilde{x}) \not\preceq f(x^*)$$

i.e., $\tilde{x}$ and x^* are mutually nondominated, **but** there may be (valid) reasons to prefer one over the other

In the multiobjective setting it is customary to distinguish two “figures”:

1) **optimizer**: whoever (or whatever) is able to determine one or more Pareto optimal solutions of the problem
2) **decision maker**: whoever (or whatever) is able to (arbitrarily) select **one** Pareto optimal solution from a set of nondominated solutions

The "decision maker"

- Single objective: if x^* and $\tilde{x}$ are both global optima, then $f(x^*) = f(\tilde{x})$ and there is no (reasonable) ground for preferring one over the other
- Multiobjective: if x^* and $\tilde{x}$ are both (global) Pareto optima, then necessarily

$$\begin{aligned} f(x^*) &\not\leq f(\tilde{x}) \\ f(\tilde{x}) &\not\leq f(x^*) \end{aligned}$$

  i.e., $\tilde{x}$ and x^* are mutually nondominated, **but** there may be (valid) reasons to prefer one over the other

In the multiobjective setting it is customary to distinguish two "figures":

1) **optimizer**: whoever (or whatever) is able to determine one or more Pareto optimal solutions of the problem
2) **decision maker**: whoever (or whatever) is able to (arbitrarily) select **one** Pareto optimal solution from a set of nondominated solutions

The “decision maker”

- Single objective: if x^* and $\tilde{x}$ are both global optima, then $f(x^*) = f(\tilde{x})$ and there is no (reasonable) ground for preferring one over the other
- Multiobjective: if x^* and $\tilde{x}$ are both (global) Pareto optima, then necessarily

$$\begin{aligned} f(x^*) &\not\leq f(\tilde{x}) \\ f(\tilde{x}) &\not\leq f(x^*) \end{aligned}$$

  i.e., $\tilde{x}$ and x^* are mutually nondominated, **but** there may be (valid) reasons to prefer one over the other

In the multiobjective setting it is customary to distinguish two “figures”:

1) **optimizer**: whoever (or whatever) is able to determine one or more Pareto optimal solutions of the problem
2) **decision maker**: whoever (or whatever) is able to (arbitrarily) select **one** Pareto optimal solution from a set of nondominated solutions

Classification of solution methods

Depending on the role played by the *decision maker* in the solution process and the moment at which they intervene, solution methods can be broadly classified as follows.

- *No-preference methods*: the decision maker plays no role; finding any Pareto optimum is considered satisfactory
- *A posteriori methods*: a (possibly complete) set of Pareto optima is generated first, and then presented to the decision maker for selection
- *A priori methods*: the decision maker specifies their preferences *before* the solution process begins; based on this information, the "best" Pareto optimal solution for the decision maker is directly computed

Classification of solution methods

Depending on the role played by the *decision maker* in the solution process and the moment at which they intervene, solution methods can be broadly classified as follows.

- *No-preference methods*: the decision maker plays no role; finding any Pareto optimum is considered satisfactory
- *A posteriori methods*: a (possibly complete) set of Pareto optima is generated first, and then presented to the decision maker for selection
- *A priori methods*: the decision maker specifies their preferences *before* the solution process begins; based on this information, the "best" Pareto optimal solution for the decision maker is directly computed

Classification of solution methods

Depending on the role played by the *decision maker* in the solution process and the moment at which they intervene, solution methods can be broadly classified as follows.

- *No-preference methods*: the decision maker plays no role; finding any Pareto optimum is considered satisfactory
- *A posteriori methods*: a (possibly complete) set of Pareto optima is generated first, and then presented to the decision maker for selection
- *A priori methods*: the decision maker specifies their preferences *before* the solution process begins; based on this information, the "best" Pareto optimal solution for the decision maker is directly computed

No-preference methods — goal programming

After computing the ideal objective vector z^{id}, define the problem

$$\begin{array}{ll} \min_x & \|z^{id} - f(x)\|_p \\ s.t. & x \in \mathcal{F} \end{array}$$

where $\|\cdot\|_p$ is the p-norm of a vector ($1 \leq p \leq \infty$). Recall that for $v \in \mathbb{R}^k$:

- $1 \leq p < \infty$: $\|v\|_p = \left(\sum_{i=1}^{k} |v_i|^p \right)^{1/p}$
- $p = \infty$: $\|v\|_\infty = \max\{|v_1|, \ldots, |v_k|\}$

Goal programming — linear case

The norms $p = 1$ and $p = \infty$ are particularly interesting because they allow one to obtain *linear* problems starting from a linear multiobjective problem. Consider:

$$\begin{array}{ll} \min_x & (c_1^\top x,\ c_2^\top x,\ \ldots,\ c_k^\top x) \\ s.t. & Ax \leq b \end{array}$$

- when $p = 1$:

$$\begin{array}{ll} \min_x & \sum_{i=1}^{k} (c_i^\top x - z_i^{id}) \\ s.t. & Ax \leq b \end{array}$$

- when $p = \infty$:

$$\begin{array}{ll} \min_x & \max_{i=1,\ldots,k} \{c_i^\top x - z_i^{id}\} \\ s.t. & Ax \leq b \end{array}$$

N.B. $c_i^\top x \geq z_i^{id}$ for all i (the ideal is unattainable in general).

Goal programming — LP reformulations

Both cases admit an equivalent **linear programming** reformulation:

- $p = 1$: introduce auxiliary variables α_i

$$\begin{aligned} \min_{x,\,\alpha_i} \quad & \sum_{i=1}^{k} \alpha_i \\ s.t. \quad & c_i^\top x - z_i^{id} \leq \alpha_i, \quad i = 1, \dots, k \\ & Ax \leq b \end{aligned}$$

- $p = \infty$: introduce a single auxiliary variable α

$$\begin{aligned} \min_{x,\,\alpha} \quad & \alpha \\ s.t. \quad & c_i^\top x - z_i^{id} \leq \alpha, \quad i = 1, \dots, k \\ & Ax \leq b \end{aligned}$$

Goal programming ($\|\cdot\|_\infty$) — portfolio

Find a nondominated solution by solving:

$$\begin{array}{ll} \min & \left\| \mu^\top x - z_1^{id},\ \frac{1}{2}x^\top Qx - z_2^{id} \right\|_\infty \\ s.t. & e^\top x = 1,\ x \geq \mathbf{0}_n \end{array}$$

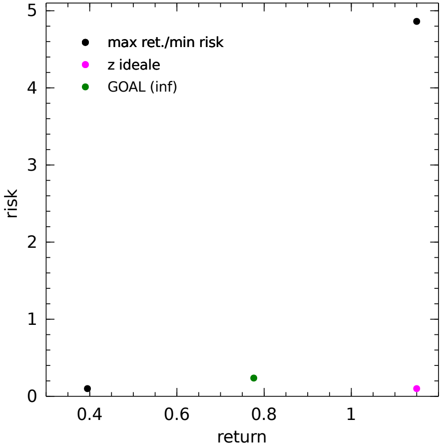
i.e., equivalently:

$$\begin{array}{ll} \min_{x,\,\alpha} & \alpha \\ s.t. & z_1^{id} - \mu^\top x \leq \alpha \\ & \frac{1}{2}x^\top Qx - z_2^{id} \leq \alpha \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{array}$$

$$\Downarrow$$

$$x_\infty^*$$

Goal programming ($\|\cdot\|_\infty$) — portfolio



Goal programming ($\|\cdot\|_1$) — portfolio

Find a nondominated solution by solving:

$$\begin{aligned} \min \quad & \left\| \mu^\top x - z_1^{id},\ \tfrac{1}{2}x^\top Q x - z_2^{id} \right\|_1 \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

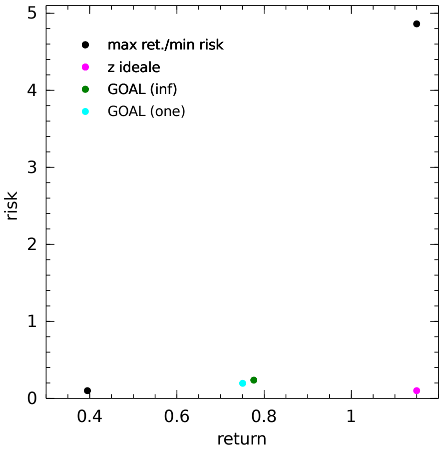
i.e., equivalently:

$$\begin{aligned} \min \quad & -\mu^\top x + \tfrac{1}{2}x^\top Q x \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

$$\Downarrow$$

$$x_1^*$$

Goal programming ($\|\cdot\|_1$) — portfolio



Weighted sum method

0) Set $X = \emptyset$
1) Choose weights w_i, $i = 1, \ldots, k$, such that

$$\sum_{i=1}^{k} w_i = 1, \quad w \geq 0$$

2) Compute

$$x^* \in \begin{array}{c} \text{argmin} \sum_{i=1}^{k} w_i f_i(x) \\ s.t.\ x \in \mathcal{F} \end{array}$$

3) Set $X = X \cup \{x^*\}$, go to step 1)

Weighted sum method — portfolio

Choose a weight vector $w \in \mathbb{R}^2$, $\mathbf{0}_2 \leq w$, and find a nondominated solution by solving:

$$\begin{aligned} \min \quad & w_1(-\mu^\top x) + w_2 \tfrac{1}{2} x^\top Q x \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

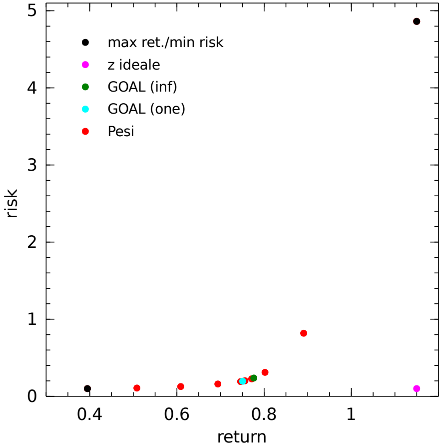
Setting $w_1 = \beta$, $w_2 = (1-\beta)$, $\beta \in [0,1]$, this becomes:

$$\begin{aligned} \min \quad & -\beta\,\mu^\top x + (1-\beta)\,\tfrac{1}{2} x^\top Q x \\ s.t. \quad & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

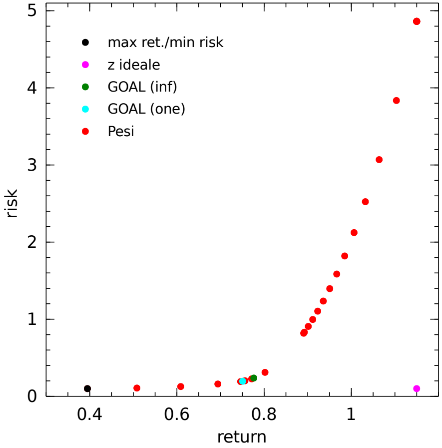
$$\Downarrow$$

$$x^*_\beta$$

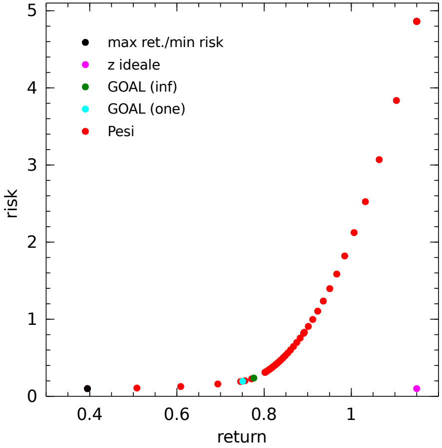
Weighted sum method ($\beta = \text{linspace}(0, 1, 10)$)



Weighted sum method ($\beta = \text{linspace}(0.89, 1, 20)$)



Weighted sum method ($\beta = \text{linspace}(0.778, 0.889, 20)$)



ϵ-constraint method

0) Set $X = \emptyset$
1) Choose an index $\ell \in \{1, \ldots, k\}$ and thresholds ϵ_i, $i = 1, \ldots, k$, $i \neq \ell$
2) Compute

$$x^* \in \operatorname{argmin} \ f_\ell(x)$$
$$s.t. \ f_i(x) \leq \epsilon_i, \ i = 1, \ldots, k, \ i \neq \ell$$
$$x \in \mathcal{F}$$

3) Set $X = X \cup \{x^*\}$, go to step 1)

ϵ-constraint method — portfolio

Choose:

- one objective to minimize, e.g. $\frac{1}{2}x^\top Qx$
- a threshold ϵ for the remaining objective(s), e.g. for $\mu^\top x$

Solve the problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^\top Qx \\ s.t. \quad & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

Warning! We are modifying the feasible region of the problem. In our case:

- if ϵ is too **large**, the problem may become **infeasible**
- if ϵ is too **small**, the problem may return **again** the solution x^*_σ

ϵ-constraint method — portfolio

Choose:

- one objective to minimize, e.g. $\frac{1}{2}x^\top Qx$
- a threshold ϵ for the remaining objective(s), e.g. for $\mu^\top x$

Solve the problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^\top Qx \\ s.t. \quad & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

Warning! We are modifying the feasible region of the problem. In our case:

- if ϵ is too **large**, the problem may become **infeasible**
- if ϵ is too **small**, the problem may return **again** the solution x^*_σ

ϵ-constraint method — portfolio

Choose:

- one objective to minimize, e.g. $\frac{1}{2}x^\top Qx$
- a threshold ϵ for the remaining objective(s), e.g. for $\mu^\top x$

Solve the problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^\top Qx \\ s.t. \quad & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

Warning! We are modifying the feasible region of the problem. In our case:

- if ϵ is too **large**, the problem may become **infeasible**
- if ϵ is too **small**, the problem may return **again** the solution x^*_σ

ϵ-constraint method — portfolio

Choose:

- one objective to minimize, e.g. $\frac{1}{2}x^\top Qx$
- a threshold ϵ for the remaining objective(s), e.g. for $\mu^\top x$

Solve the problem

$$\begin{aligned} \min \quad & \tfrac{1}{2}x^\top Qx \\ s.t. \quad & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

Warning! We are modifying the feasible region of the problem. In our case:

- if ϵ is too **large**, the problem may become **infeasible**
- if ϵ is too **small**, the problem may return **again** the solution x^*_σ

ϵ-constraint method — choice of ϵ

$$\begin{aligned} \min \quad & \tfrac{1}{2} x^\top Q x \\ s.t. \quad & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{aligned}$$

Question: what is a "reasonable" range of values for the parameter ϵ?

Answer: the constraint $\mu^\top x \geq \epsilon$ is feasible and non-trivial precisely when

$$\mu^\top x^*_\sigma \leq \epsilon \leq \mu^\top x^*_\mu,$$

i.e., ϵ must lie between the minimum and maximum attainable expected return on the feasible set.

ϵ-constraint method — choice of ϵ

$$\begin{array}{ll} \min & \frac{1}{2}x^\top Qx \\ s.t. & \mu^\top x \geq \epsilon \\ & e^\top x = 1,\ x \geq \mathbf{0}_n \end{array}$$

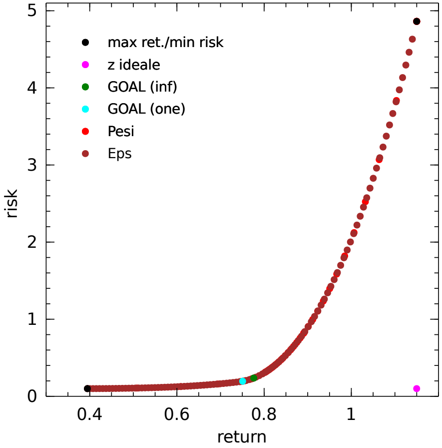
Question: what is a "reasonable" range of values for the parameter ϵ?

Answer: the constraint $\mu^\top x \geq \epsilon$ is feasible and non-trivial precisely when

$$\mu^\top x_\sigma^* \leq \epsilon \leq \mu^\top x_\mu^*,$$

i.e., ϵ must lie between the minimum and maximum attainable expected return on the feasible set.

ϵ-constraint method ($\epsilon = \mathrm{linspace}(0.4, 1.14, 100)$)



Lexicographic ordering method

The decision maker specifies a **ranking of the objective functions**. Let

$$f_1(x), \ldots, f_k(x)$$

be the objectives ordered by **decreasing importance** (from most to least important).

The idea: optimize objectives one at a time in order of priority, adding the previously attained optimal value as a constraint at each step — so that more important objectives are never sacrificed for less important ones.

Lexicographic ordering method — algorithm

- Step 1

$$x^{*,1} \in \operatorname*{argmin}_{x \in \mathcal{F}} f_1(x)$$

- Step 2

$$x^{*,2} \in \operatorname*{argmin}_{\substack{x \in \mathcal{F} \\ f_1(x) \leq f_1(x^{*,1})}} f_2(x)$$

- Step 3

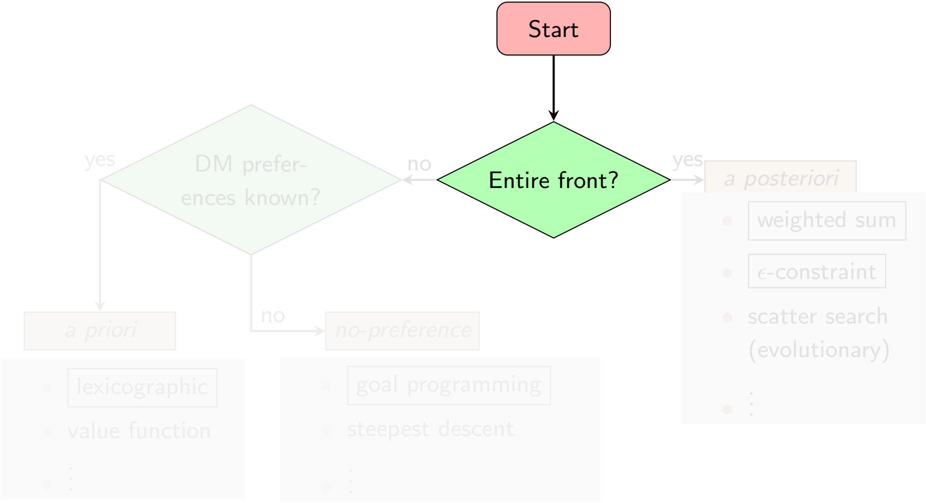
$$x^{*,3} \in \operatorname*{argmin}_{\substack{x \in \mathcal{F} \\ f_1(x) \leq f_1(x^{*,2}) \\ f_2(x) \leq f_2(x^{*,2})}} f_3(x)$$

$$\vdots$$

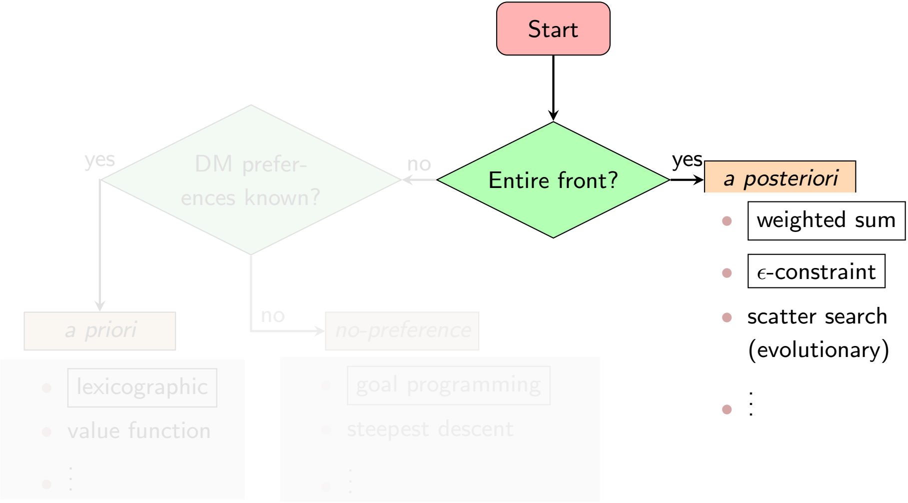
- Step h

$$x^{*,h} \in \operatorname*{argmin}_{\substack{x \in \mathcal{F} \\ f_i(x) \leq f_i(x^{*,h-1}) \\ i = 1, \dots, h-1}} f_h(x)$$

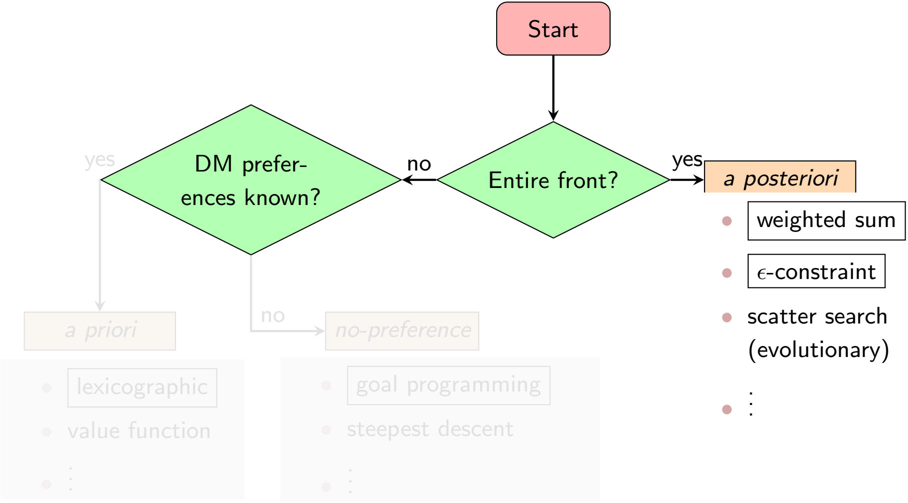
Summary


Start
Entire front?
yes
a posteriori
weighted sum
ϵ-constraint
scatter search (evolutionary)
⋮
no
DM preferences known?
yes
a priori
lexicographic
value function
⋮
no
no-preference
goal programming
steepest descent
⋮

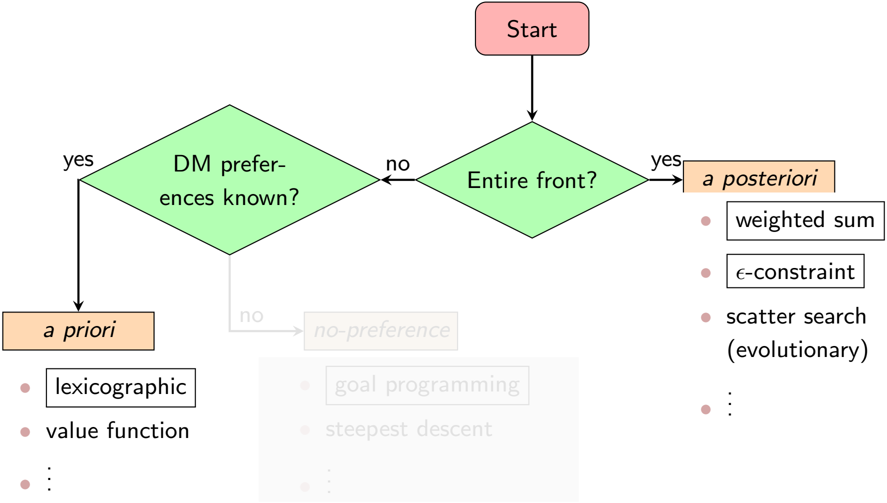
Summary


Start
Entire front?
yes
a posteriori
weighted sum
ϵ-constraint
scatter search (evolutionary)
⋮
no
DM preferences known?
yes
a priori
lexicographic
value function
⋮
no
no-preference
goal programming
steepest descent
⋮

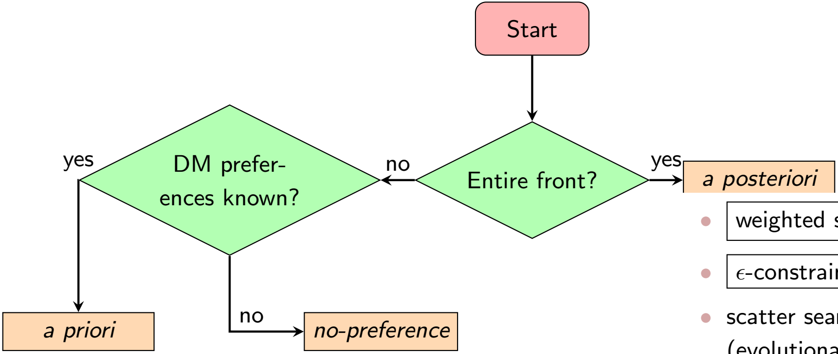
Summary


Start
Entire front?
yes
a posteriori
no
DM preferences known?
yes
a priori
no
no-preference
weighted sum
ϵ-constraint
scatter search (evolutionary)
⋮
lexicographic
value function
⋮
goal programming
steepest descent
⋮

Summary


Start
Entire front?
yes
a posteriori
weighted sum
ϵ-constraint
scatter search (evolutionary)
⋮
no
DM preferences known?
yes
a priori
lexicographic
value function
⋮
no
no-preference
goal programming
steepest descent
⋮

Summary



a priori:

- lexicographic
- value function
- ⋮

no-preference:

- goal programming
- steepest descent
- ⋮

a posteriori:

- weighted sum
- ϵ-constraint
- scatter search (evolutionary)
- ⋮