Indicator Based Search in Variable Orderings: Theory and Algorithms

Pradyumn K. Shukla, Marlon A. Braun

Institute AIFB, Karlsruhe Institute of Technology

Outline

Outline

Multi-objective Optimization Problem

Let $F_1, \ldots, F_m : \mathbb{R}^n \to \mathbb{R}$ and $X \subseteq \mathbb{R}^n$ be given.

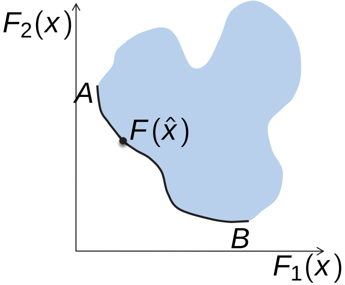
$$\begin{aligned} \text{min} \qquad & F(x) := (F_1(x), F_2(x), \ldots, F_m(x)) \\ \text{subject to} \qquad & x \in X \end{aligned}$$

- 'min' depends on a partial order of $\mathbb{R}^m$
- A set (or a cone) $D \subset \mathbb{R}^m$ is used to define such an order
- u 'D-**dominates**' v, if $v \in \{u\} + D \setminus \{\mathbf{0}\}$

An Optimality Notion

Definition

*A point $\hat{x} \in X$ is D-**optimal** if no feasible point 'D-dominates' $F(\hat{x})$. If $D = \mathbb{R}^m_+$, then $\hat{x}$ is Pareto-optimal and $F(\hat{x})$ is efficient.*



Efficient points of a bicritetia problem

Variable Ordering

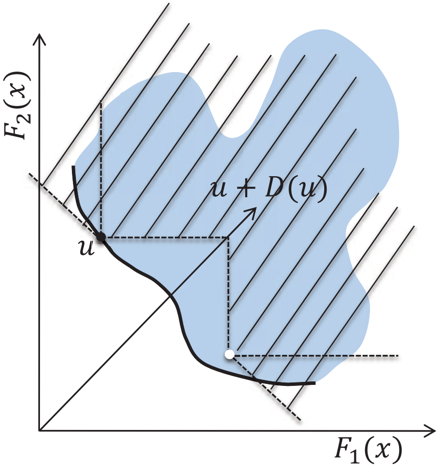
Definition

A partial ordering induced by sets $\mathcal{D}(\cdot)$ that depend on the point $u \in \mathbb{R}^m$ is called as variable ordering.

Applications

- Medical image registration
- Multicriteria game theory
- Problems with equitable efficiency

An Example From Multi-objective Resource Allocation



The shaded area $\mathcal{D}(u)$ is a non-convex set and is not a cone.

An Example From Medical Engineering

A weight is assigned to every point $u \in \mathbb{R}^m$. Corresponding to such a weight a set of preferred directions is defined by

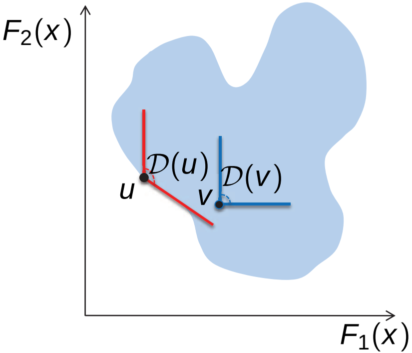
$$\mathcal{D}(u) := \left\{ d \in \mathbb{R}^m | \sum_{i=1}^{m} \text{sgn}(d_i) w_i(u) \geq 0 \right\}.$$

- Non-convex set in three (or more) dimensions
- Applications in image registration

Variable Domination

Definition

*u '$\leq_1$-**dominates**' v if $v \in \{u\} + \mathcal{D}(v) \setminus \{\mathbf{0}\}$. Similarly, we say that u '$\leq_2$-**dominates**' v if $v \in \{u\} + \mathcal{D}(u) \setminus \{\mathbf{0}\}$.*



u '$\leq_2$-*dominates*' v while u does not '$\leq_1$-*dominate*' v

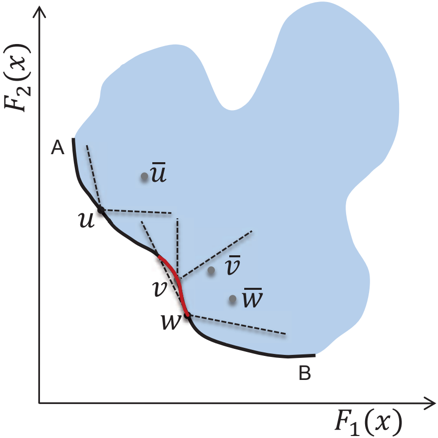
Variable Optimality

Definition

A point $\hat{u} \in F(X)$ is called a ***minimal point*** *of $F(X)$ if there is no feasible point which '$\leq_1$-dominates' $\hat{u}$.*

Similarly, a point $\hat{u} \in F(X)$ is called a ***nondominated point*** *of $F(X)$ if there is no feasible point which '$\leq_2$-dominates' $\hat{u}$.*

An Example



The red curve is not non-dominated but it may be minimal.

Outline

Relations Between Optimal Solutions

Let $\mathcal{E}$, $\mathcal{E}_{\mathcal{N}}$ and $\mathcal{E}_{\mathcal{M}}$ be the set of efficient, nondominated and minimal points. Let X_p, $X_{\mathcal{N}}$ and $X_{\mathcal{M}}$ be their pre-images.

We assume throughout that all the sets are closed and that $\mathcal{D}(u) + \mathbb{R}^m_+ \subseteq \mathcal{D}(u)$ for every $u \in F(X)$.

Lemma

$X_{\mathcal{M}} \subseteq X_p$, $X_{\mathcal{N}} \subseteq X_p$, $\mathcal{E}_{\mathcal{M}} \subseteq \mathcal{E}$ and $\mathcal{E}_{\mathcal{N}} \subseteq \mathcal{E}$.

Characterization of Minimal Points

Lemma

$\hat{u} \in F(X)$ is a minimal point of $F(X)$ if and only if $\hat{u} \in \mathcal{E}$ and $\hat{u}$ is a minimal point of $\mathcal{E}$.

In order to check if a point is minimal or not it is sufficient to check the $\leq_1$-domination w.r.t. the efficient points only.

Characterization of Nondominated Points

Assumption

If u Pareto-dominates v, then $\mathcal{D}(v) \subseteq \mathcal{D}(u)$.

Lemma

Let the above assumption hold. Then, $\hat{u} \in F(X)$ is a nondominated point of $F(X)$ if and only if $\hat{u} \in \mathcal{E}$ and $\hat{u}$ is a nondominated point of $\mathcal{E}$.

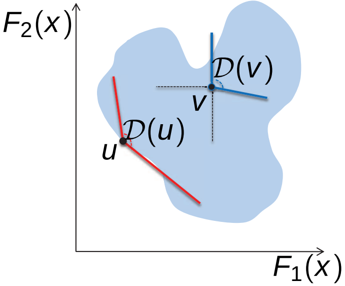
In order to check if a point is nondominated or not it is sufficient to check the $\leq_2$-domination w.r.t. efficient points only.

Weaker Assumptions

Assumption

For every $v \in F(X)$ *there exists a* $u \in \left(F(v) - \mathbb{R}^m_+\right) \cap \mathcal{E}$ *so that* $\mathcal{D}(v) \subseteq \mathcal{D}(u)$.

- Holds for equitable (and other) variable orderings
- Related to the transitivity of the $\leq_2$-domination



Algorithmic Implications

- Almost every population based algorithm finds/ uses Pareto non-dominated solutions
- The characterization reduces the additional burden of finding minimal/ nondominated points
- Jahn-Graef-Younes sorting technique to reduce pairwise comparisons

Minimal Variable Ordering Hypervolume

Definition

Let $\mathcal{S} \subset \mathbb{R}^m$, and let $\mathbf{r} \in \mathbb{R}^m$ indicate the reference point. The *minimal hypervolume* is defined by

$$\mathcal{H}^m(\mathcal{S}, \mathbf{r}) := \mathrm{Vol}\left(\{\mathbf{w} \in \mathbb{R}^m | \exists \mathbf{v} \in \mathcal{E}_{\mathcal{M}}(\mathcal{S}) : \mathbf{v} \leq \mathbf{w} \leq \mathbf{r}\}\right).$$

Minimal Variable Ordering Hypervolume

Definition

Let $\mathcal{S} \subset \mathbb{R}^m$, and let $\mathbf{r} \in \mathbb{R}^m$ indicate the reference point. The *minimal hypervolume* is defined by

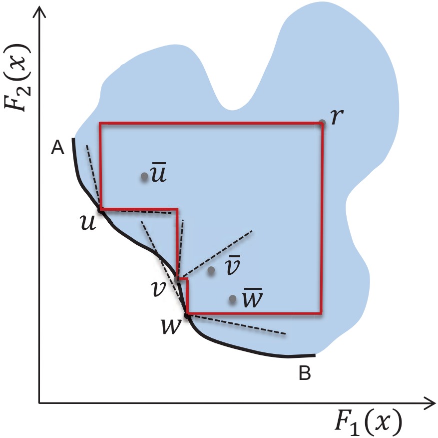
$$\mathcal{H}^m(\mathcal{S}, \mathbf{r}) := \mathrm{Vol}\left(\{\mathbf{w} \in \mathbb{R}^m | \exists \mathbf{v} \in \mathcal{E}_{\mathcal{M}}(\mathcal{S}) : \mathbf{v} \leq \mathbf{w} \leq \mathbf{r}\}\right).$$

The *minimal set hypervolume* of a set $\mathcal{A} \subseteq \mathcal{S}$ is defined by

$$\mathcal{H}^m(\mathcal{A}, \mathcal{S}, \mathbf{r}) \quad := \quad \mathrm{Vol}\left(\{\mathbf{w} \in \mathbb{R}^m | \exists \mathbf{v} \in \mathcal{E}_{\mathcal{M}}(\mathcal{S}) \cap \mathcal{E}_{\mathcal{M}}(\mathcal{A}) : \mathbf{v} \leq \mathbf{w} \leq \mathbf{r}\}\right).$$

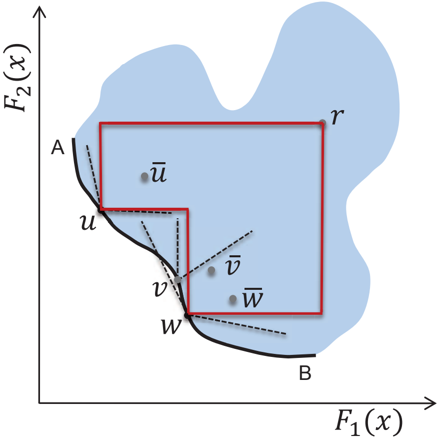
Nondominated notions are defined in a similar way.

An Example



The volume enclosed by the red lines is the *minimal hypervolume* of the set $\{u, v, w\}$.

An Example



The volume enclosed by the red lines is the *nondominated hypervolume* of the set $\{u, v, w\}$.

Compatibility and Completeness

Let $\mathcal{A}, \mathcal{B} \subset \mathbb{R}^m$ be two finite sets.

Theorem

($\not\leq_1$-Compatibility)

$$\mathcal{B} \not\leq_1 \mathcal{A} \Leftarrow \exists \mathbf{r} \in \mathbb{R}^m : \mathcal{H}^m(\mathcal{A}, \mathcal{A} \cup \mathcal{B}, \mathbf{r}) > \mathcal{H}^m(\mathcal{B}, \mathcal{A} \cup \mathcal{B}, \mathbf{r}).$$

($\leq_1$-Completeness)

$$\mathcal{A} \leq_1 \mathcal{B}\,, \mathcal{B} \not\leq_1 \mathcal{A} \;\Rightarrow\; \mathcal{H}^m(\mathcal{A}, \mathcal{A} \cup \mathcal{B}, \mathbf{r}) > \mathcal{H}^m(\mathcal{B}, \mathcal{A} \cup \mathcal{B}, \mathbf{r})$$

for all $\mathbf{r}$ *such that* $\mathrm{nad}(\mathcal{A} \cup \mathcal{B}) < \mathbf{r}$.

Algorithmic Implications

- Computing minimal hypervolume is *almost* the same as computing classical hypervolume
- Minimal hypervolume computes the volume in the original objective space
- A direct extension of the classical hypervolume to variable orderings is theoretically (and computationally) intractable

Outline

Algorithms

Three versions of SMS-EMOA were implemented.

- LF-SMS-EMOA
  - Splits the last non-dominated (using Pareto ordering) front
- FF-SMS-EMOA
  - Splits the first non-dominated (using Pareto ordering) front
- CF-SMS-EMOA
  - Sorts the population using the variable domination structure

Bishop-Phelps Cones

Bishop-Phelps cones are described by two parameters:

- A scalar γ controlling the angle of the cone
- A reference (ideal) vector $p \in \mathbb{R}^m$

Based on this, variable domination cone $\mathcal{C}(u)$ is defined by

$$\mathcal{C}(u) := \{d \mid \langle d, u - p \rangle \geq \gamma \cdot \|d\| \cdot [u - p]_{\min}\},$$

where $[u - p]_{\min}$ is the minimal component of the vector $u - p$.

Test Problems and Performance Metrics

Test problems

- Many CTP, DTLZ, CEC07, WFG, and ZDT instances
- Zero vector as the ideal point and $\gamma = 0.5$

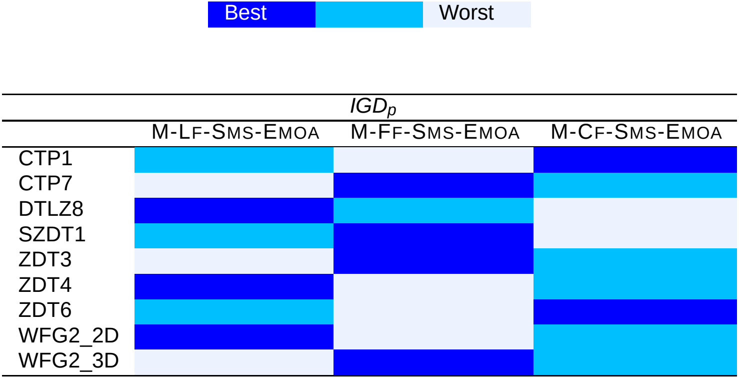
Performance metrics

- Power mean based IGD
  - First diverse points on the efficient front are generated
  - From these we calculate the minimal (or nondominated) points
- Minimal (or nondominated) hypervolume metric

Outline

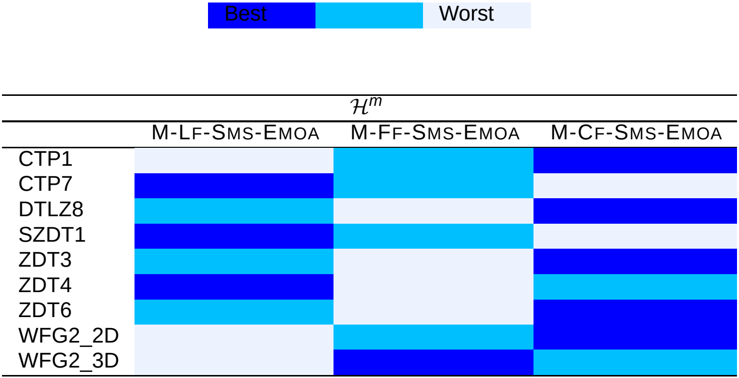
Minimal Study

Power mean based inverted generational distance metric



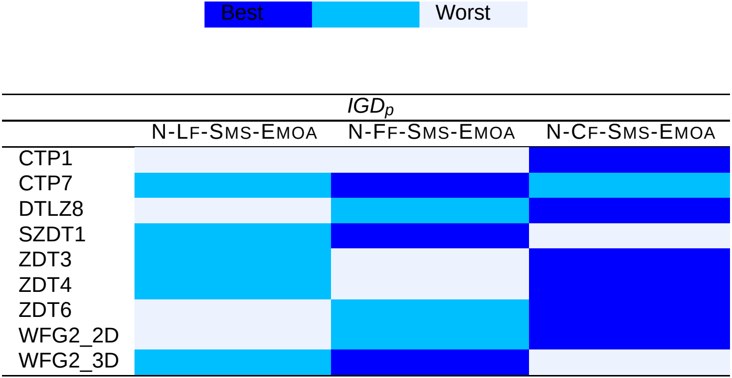
Minimal Study

Minimal hypervolume metric



Nondominated Study

Power mean based inverted generational distance metric



Non-dominated hypervolume metric

Best		Worst

		$\mathcal{H}^n$	
	N-LF-SMS-EMOA	N-FF-SMS-EMOA	N-CF-SMS-EMOA
CTP1			
CTP7			
DTLZ8			
SZDT1			
ZDT3			
ZDT4			
ZDT6			
WFG2_2D			
WFG2_3D			

Outline

Summary

- Analyzed minimal and nondominated points

Summary

- Analyzed minimal and nondominated points
- Presented new theoretical results

Summary

- Analyzed minimal and nondominated points
- Presented new theoretical results
- Proposed new hypervolume based indicators

Summary

- Analyzed minimal and nondominated points
- Presented new theoretical results
- Proposed new hypervolume based indicators
- Based on the the above three algorithms were developed

Summary

- Analyzed minimal and nondominated points
- Presented new theoretical results
- Proposed new hypervolume based indicators
- Based on the the above three algorithms were developed
- For nondominated variable orderings N-CF-SMS-EMOA performed the best