Multi-criteria Optimization

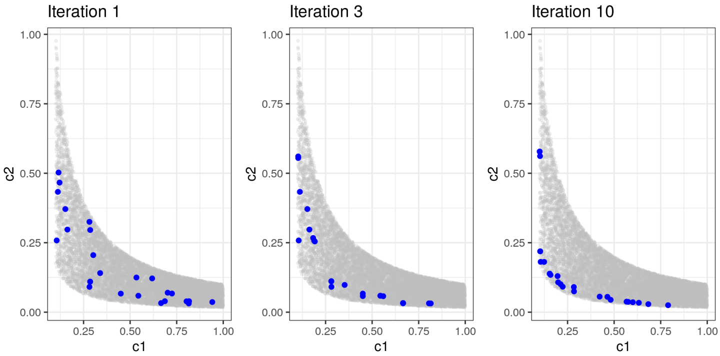
Evolutionary Approaches

Bernd Bischl   Frank Hutter   Lars Kotthoff
Marius Lindauer   Joaquin Vanschoren

A-posteriori methods and evolutionary algorithms I

Evolutionary multi-objective algorithms (EMOAs) evolve a diverse population over time to approximate the Pareto front.

Initialize population
↓
Eval population
↓
Parent selection
↓
Variation
↓
no　Eval offspring
↓
Survival selection
↓
Stop
↓ yes



A-posteriori methods and evolutionary algorithms II

Algorithm 1 Basic EA template loop

Init and eval population $\mathcal{P}_0 \subset \mathbf{\Lambda}$ with $|\mathcal{P}| = \mu$
$t \leftarrow 0$
repeat
    Select parents and generate offspring $\mathcal{Q}_t$ with $|\mathcal{Q}_t| = \lambda$
    Select μ survivors $\mathcal{P}_{t+1}$
    $t \leftarrow t + 1$
until Stop criterion fulfilled

- Note that (as in the EA lecture unit) we are using somewhat non-standard notation here.
- Nearly all steps in the above template work also for EMOAs but both parent and survival selection are now less obvious. How do we rank under multiple objectives?

NSGA-II

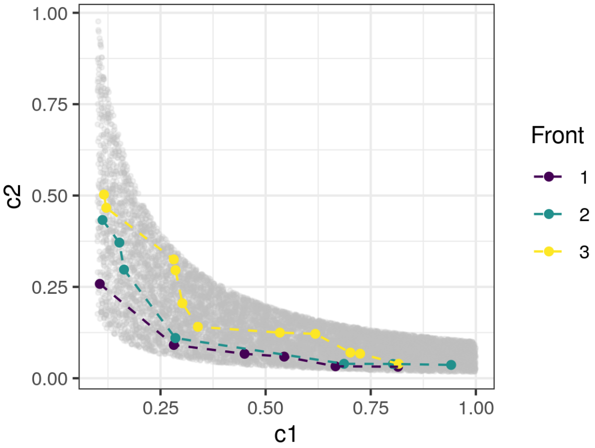
The **non-dominated sorting genetic algorithm (NSGA-II)** was published by [K. Dep et al. 2002].

- Follows a $(\mu + \lambda)$ strategy.
- All previously discussed variation strategies can be used; the original paper uses tournament selection, polynomial mutation and simulated binary crossover.
- Parent and survival selection rank candidates by
  1. **Non-dominated sorting** as main criterion
  2. **Crowding distance assignment** as tie breaker

NSGA-II: non-dominated sorting I

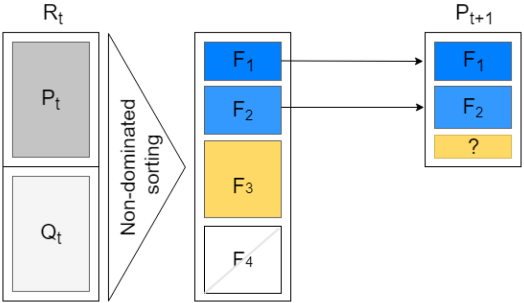
NDS partitions an objective space set into fronts $\mathcal{F}_1 \prec \mathcal{F}_2 \prec \mathcal{F}_3 \prec \dots$.

- $\mathcal{F}_1$ is non-dominated, each $\boldsymbol{\lambda} \in \mathcal{F}_2$ is dominated, but only by points in $\mathcal{F}_1$, each $\boldsymbol{\lambda} \in \mathcal{F}_3$ is dominated, but only by points in $\mathcal{F}_1$ and $\mathcal{F}_2$, and so on.
- We can easily compute the partitioning by computing all non-dominated points $\mathcal{F}_1$, removing them, then computing the next layer of non-dominated points $\mathcal{F}_2$, and so on.



NSGA-II: non-dominated sorting II

How does survival selection now work? We fill μ *places* one by one with $\mathcal{F}_1, \mathcal{F}_2, \ldots$ until a front can no longer **fully** survive (here: $\mathcal{F}_3$).

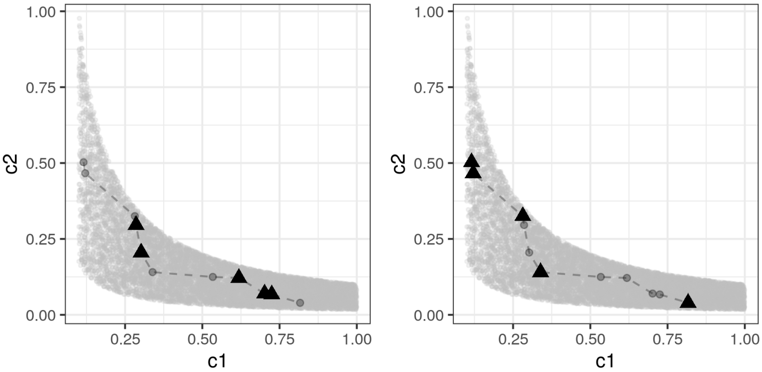


Which individuals survive from $\mathcal{F}_3$? $\rightarrow$ **crowding sort**

NB: the same principle to rank individuals is applied in tournament selection in parent selection.

NSGA-II: crowding distance I

Idea: Add *good* representatives of front $\mathcal{F}_3$, define this as points of "low density" in c-space.



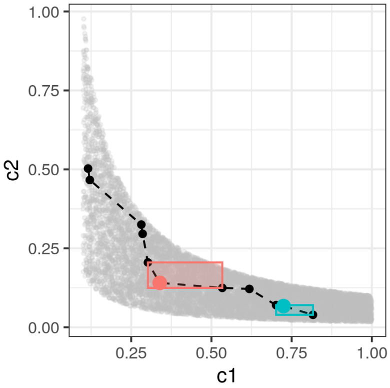
Left: Not good, points very close together. Right: better.

NSGA-II: crowding distance II

For each objective c_j

- Sort points by c_j
- Normalize scores to [0,1]
- Assign border points (which have score 0 or 1) a CD of ∞ (they should always be selected, if possible)
- Each point gets a distance score, which is the distance between its 2 next-neighbors w.r.t. the sorting of c_j

For each point, all of its m distance scores are summed up (or averaged) and points are ranked w.r.t. to this overall score.



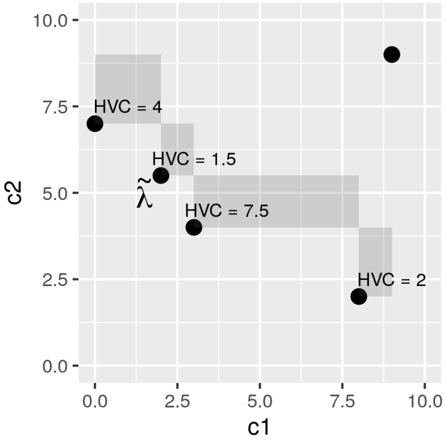
Red: Point with high CD. Blue: Low CD.

Selection criteria: contribution to the hypervolume I

SMS-EMOA
(S-Metric-Selection-EMOA) [Beume et al. 2007]
is a $(\mu + 1)$ EMOA and evaluates fitness of an individual $\boldsymbol{\lambda} \in \mathcal{P} \subset \boldsymbol{\Lambda}$ based on its contribution to the dominated HV:

$$\Delta s(\boldsymbol{\lambda}, \mathcal{P}) = S(\mathcal{P}, R) - S(\mathcal{P} \setminus \{\boldsymbol{\lambda}\}, R).$$

- Dark rectangles: HV contribution of dots.
- Grey point: reference point.
- The HVC contribution is the volume of space that is dominated only by $\boldsymbol{\lambda}$, and nothing else.
- $\tilde{\boldsymbol{\lambda}}$ has lowest S-metric contribution.



SMS-EMOA algorithm I

Algorithm 2 SMS-EMOA

1: Generate start population $\mathcal{P}_0$ of size μ
2: $t \leftarrow 0$
3: **repeat**
4: Generate **one** individual $\mathbf{q}$ by recombination and mutation of $\mathcal{P}_t$
5: $\{\mathcal{F}_1, ..., \mathcal{F}_k\} \leftarrow \mathsf{NDS}(\mathcal{P}_t \cup \{\mathbf{q}\})$
6: $\tilde{\boldsymbol{\lambda}} \leftarrow \mathrm{argmin}_{\boldsymbol{\lambda} \in \mathcal{F}_k} \Delta s(\boldsymbol{\lambda}, \mathcal{F}_k)$
7: $\mathcal{P}_{t+1} \leftarrow (\mathcal{P}_t \cup \{\mathbf{q}\}) \setminus \{\tilde{\boldsymbol{\lambda}}\}$
8: $t \leftarrow t + 1$
9: **until** Termination criterion fulfilled

- L5: the set of temporary $(\mu + 1)$ individuals is partitioned by NDS into k fronts $\mathcal{F}_1, ..., \mathcal{F}_k$.
- L6-7: In last front, find $\tilde{\boldsymbol{\lambda}} \in \mathcal{F}_k$ with smallest HV contribution - and kill it.
- Fitness of an individual is mainly the rank of its front and HV contribution as tie-breaker.