On Invariances in Evolutionary Algorithms

Anne Auger[1], Sylvain Gelly[1], Sylvie Ruette[2], Olivier Teytaud[1]

1 - TAO Team, INRIA Futurs, LRI, UMR 8623 (CNRS - Univ. Paris-Sud),
Bat. 490 Paris-Sud University 91405 Orsay, France
2 - Laboratoire de Mathematiques, CNRS UMR 8628,
Bat. 425, Paris-Sud University, 91405 Orsay cedex, France

We summarize current research on the pros and cons of invariance properties of optimization algorithms:

1. Robustness optimality of comparison-based algorithms for the worst case among increasing transformations of the fitness function. Informally, we show that for the worst case on increasing transformations of the fitness function, comparison-based algorithms are in fact optimal. Precisely, we consider $x_1(o,f)$,...,$x_n(o,f)$ the points visited by an optimization algorithm o when working on a fitness function f, and we let $x^*(f)$ be the arg min of a fitness function f. We show that (under mild technical assumptions) for all random function f, for all optimization algorithm o, there exists an optimization algorithm o' which only depends on comparisons such that
$$E_f \sup_{g\in G} ||x_n(o, g\circ f) - x^*(g\circ f)||^2 = E_f \sup_{g\in G} ||x_n(o', g\circ f) - x^*(g\circ f)||^2$$
where G is the set of increasing mappings from $\mathbb{R}$ to $\mathbb{R}$ [2].
2. Lower-bounds on the convergence rates in continuous optimization for comparison-based algorithms. Precisely, we show that all algorithms o for which $\forall f, \forall g \in G, \forall n, x_n(o,f) = x_n(o, g\circ f)$ have a convergence rate at most linear, with a constant scaling as $O(1/d)$ where d is the dimension ([5], [3]), in the sense that
$$-\log(||x_n(o,f) - x^*(f)||) = O\left(k(n)/d\right) = O\left(n\log(n)/d\right)$$
where $k(n)$ is the number of comparisons used before visiting $x_n(o,f)$.
3. Lower-bounds on the convergence rates in multi-objective optimization for comparison-based multi-objective algorithms [4]. These results show lower complexity bounds $u(\epsilon)$ on the number of iterations points necessary for ensuring a precision ϵ for the Hausdorff metric and $u(\epsilon)$ is close to random search's number of iterations $rs(\epsilon)$ when the number of conflicting objectives increase. Namely, the ratio $log\left(u(\epsilon)\right)/log\left(rs(\epsilon)\right)$ converges to $d/(d-1)$ as ϵ goes to 0, i.e. $u(\epsilon) = O\left(rs(\epsilon)^{(d-1)/d}\right)$ if there is no more assumption than the C^1 nature of Pareto-fronts.
4. Homogeneity of the Markov Chain ([1]) resulting from linear invariance properties of the algorithm, e.g. when the output of the algorithm when inputs are rotated/translated is similarly rotated/translated). From this homogeneity plus from the stability of the resulting Markov chains follow linear convergence. The use of drift conditions for proving the stability of the homogeneous Markov chain will be discussed.

References

1. A. Auger. Convergence results for (1,λ)-SA-ES using the theory of φ-irreducible markov chains. *Theoretical Computer Science*, 334:35–69, 2005.
2. S. Gelly, S. Ruette, and O. Teytaud. Comparison-based algorithms: worst-case optimality, optimality w.r.t a bayesian prior, the intraclass-variance minimization in eda, and implementations with billiards. In *PPSN-BTP workshop*, 2006.
3. J. Jägersküpper. *Probabilitic Analysis of Evolution Strategies Using Isotropic Mutations*. PhD thesis, Dortmund Unversity, 2006.
4. O. Teytaud. How entropy theorems can show that approximating high dimensional pareto-fronts is too hard. In *PPSN-BTP workshop*, 2006.
5. O. Teytaud and S. Gelly. General lower bounds for evolutionary algorithms. In 10^{th} *International Conference on Parallel Problem Solving from Nature (PPSN 2006)*, 2006.