



HAL
open science

Rank-based Linear-Quadratic Surrogate Assisted CMA-ES

Mohamed Gharafi, Nikolaus Hansen, Rodolphe Le Riche, Dimo Brockhoff

► **To cite this version:**

Mohamed Gharafi, Nikolaus Hansen, Rodolphe Le Riche, Dimo Brockhoff. Rank-based Linear-Quadratic Surrogate Assisted CMA-ES. GECCO '25 Companion - Proceedings of the Genetic and Evolutionary Computation Conference Companion, Jul 2025, Malaga, Spain. . <hal-05035436>

HAL Id: hal-05035436

<https://hal.science/hal-05035436v1>

Submitted on 15 Apr 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Rank-based Linear-Quadratic Surrogate Assisted CMA-ES

Mohamed Gharafi
mohamed.gharafi@inria.fr

Inria, CMAP, Ecole Polytechnique, IP Paris
Palaiseau, France

Rodolphe Le Riche
CNRS LIMOS
St-Etienne, France

Nikolaus Hansen

Inria, CMAP, Ecole Polytechnique, IP Paris
Palaiseau, France

Dimo Brockhoff

Inria, CMAP, Ecole Polytechnique, IP Paris
Palaiseau, France

Abstract

In this poster, we introduce a rank-based surrogate-assisted variant of CMA-ES. Unlike previous methods that employ rank information as constraints to train an SVM classifier, our approach employs a linear-quadratic regression on the ranks. We investigate the method's invariance empirically. While this first algorithm outperforms CMA-ES with a few exceptions, it falls short to entirely meet the lq-CMA-ES performance levels. To address this, we propose an enhanced variant that handles together two alternative surrogates, one based on the ranks and one based on the original function values. Although this variant sacrifices strict invariance, it gains in robustness and achieves performance comparable to, or even exceeding, lq-CMA-ES on transformed problems. This last algorithm shows how simply incorporating new transformations of rank values could improve any surrogate-based CMA-ES variant.

CCS Concepts

• **Computing methodologies** → **Machine learning approaches**; *Cross-validation*; • **Mathematics of computing** → **Stochastic control and optimization**.

Keywords

Surrogate-assisted optimization, Surrogate models, CMA-ES, Invariance

ACM Reference Format:

Mohamed Gharafi, Nikolaus Hansen, Rodolphe Le Riche, and Dimo Brockhoff. 2025. Rank-based Linear-Quadratic Surrogate Assisted CMA-ES. In *Genetic and Evolutionary Computation Conference (GECCO '25 Companion)*, July 14–18, 2025, Malaga, Spain. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3712255.3726606>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
GECCO '25 Companion, Malaga, Spain
© 2025 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-1464-1/2025/07
<https://doi.org/10.1145/3712255.3726606>

1 Introduction

The Covariance Matrix Adaptation Evolution Strategy (CMA-ES) is a well-established algorithm for single-objective black-box optimization, known for its efficiency [8, 13]. It evolves a normal distribution iteratively, relying solely on solution rankings, ensuring invariance to strictly increasing transformations—a valuable property in optimization [5, 9]. Reducing function evaluations is critical, especially in expensive optimization, where surrogate models can help by approximating objective values for ranking or pre-screening candidates [1].

Surrogate assistance in CMA-ES often compromises its invariance, as the surrogate model may not preserve invariance to strictly increasing transformations of the objective function. This issue was observed in lq-CMA-ES [6], which employs successive linear and quadratic models. Our goal is to design surrogates that preserve said invariance property. We first review invariant surrogate-assisted CMA-ES methods, then introduce our rank-based variant of lq-CMA-ES. Finally, we analyze its invariance and compare its performance to CMA-ES, lq-CMA-ES, and S^* ACM-ES.

2 Existing work related to Surrogate assisted CMA-ES

Existing literature dealing with the addition of surrogates to CMA-ES can be found in [3, 4, 6, 10, 12]. We narrow our focus to review rank-based surrogate-assisted variants of CMA-ES. To the best of our knowledge, the most notable ones include RankSVM coupled with CMA-ES [14] and S^* ACM-ES [12].

Rank Support Vector Machine (RankSVM). The Rank Support Vector Machine (RankSVM) approach, as detailed in [14], employs a Support Vector Machine (SVM) classifier to learn a utility function. This utility function replaces the original function calls to rank candidate solutions. This approach is inherently invariant to strictly increasing transformation of the objective function since the model building depends only on the pairwise comparison of solutions. The gains in terms of number of function evaluations depend on the choice of the kernel, as is the case for any kernel method. These gains however are reported to rapidly decrease with the dimension d of the problem in [14].

Self-Adaptive ACM-ES. Using the RankSVM method, the self-adaptive ACM-ES (S^* ACM-ES) uses the utility function values to pre-screen offsprings from a sample larger than

the population size. Similarly, this approach is invariant to strictly increasing transformations of the objective function. Additionally, S^* ACM-ES remains invariant to rotations of the search space, which is achieved by using the covariance matrix adapted by CMA-ES as an SVM kernel [11].

Both RankSVM and S^* ACM-ES rely on SVM classifiers, which are known to be sensitive to the kernel choice. In the following, we introduce a method that uses linear-quadratic regression on ranks instead of function values.

3 lq-CMA-ES: A Linear-Quadratic Surrogate-Assisted CMA-ES

The lq-CMA-ES is a surrogate-assisted variant of CMA-ES, originally introduced in [6]. In this method, a linear, first, then quadratic model acts as the surrogate, hence its name. The switch from linear to quadratic surrogate occurs when sufficiently many observations of the true function have been gathered. In lq-CMA-ES, the training dataset comprises the solutions sampled by CMA-ES evaluated on the true function. The decision to utilize or refine the surrogate depends on a ranking quality criterion: the Kendall's tau (τ) correlation coefficient between the rankings derived from the actual function values and those inferred from the surrogate values at a subset of assessed solutions [6]. The linear-quadratic model is defined as the scalar product $\mathbf{m}^\top z(\mathbf{x}_i)$ where z is the full quadratic features mapping of the solution \mathbf{x}_i with \mathbf{m} the vector of the model coefficients, learned by minimizing the following weighted error for positive weights w_i :

$$\sum_i w_i^2 (\mathbf{m}^\top z(\mathbf{x}_i) - f(\mathbf{x}_i))^2 \quad (1)$$

The performance of the algorithm depends on the function values. As discussed in [6], the lq-CMA-ES algorithm demonstrates strong performance on several functions such as the sphere function and the ellipsoid function. However, the application of a strictly increasing transformation strongly affects the performance of lq-CMA-ES as shown in [6]. This is because the model building relies on the function values, as evident in the expression of the weighted error Equation (1). Next, we present a simple way to achieve invariance with respect to any increasing transformation of f in the context of optimization with surrogates.

4 lq-CMA-ES on a dataset of ranks

In our proposed method, Rank-lq-CMA-ES, we apply lq-CMA-ES to a dataset of rank values computed from the original function values. Given candidate solutions $\mathbf{x}_{i:n}$, where i represents the i^{th} best solution in the model's archive, we transform them using the mapping $\mathbf{x}_{i:n} \mapsto i$ to replace function values with rank values. The new weighted error, derived from Equation (1), is expressed as: $\sum_i w_{i:n}^2 (\mathbf{m}^\top z(\mathbf{x}_{i:n}) - i)^2$.

Any strictly increasing transformation of the original fitness values results in the same rank data, ensuring invariance of the model building and therefore the resulting algorithm to strictly increasing fitness transformations. Nevertheless, since the algorithm employs a linear-quadratic model, the use of

rank-transformed data inherently affects the quality of the model either positively or negatively.

5 Rank-lq-CMA-ES on transformed ranks

Invariant surrogate assisted CMA-ES with transformed ranks. In general, ranks can be substituted with any strictly increasing transformation of ranks $(t_i)_{0 \leq i < n}$. Since the t_i 's only depend on the ordering of the function values, the algorithm remains invariant. We propose an extended variant of Rank-lq-CMA-ES using multiple transformations of rank values to build the linear quadratic model. Algorithm 1 details the steps of this new variant.

Core Concept. The extended Rank-lq-CMA-ES's main idea is to dynamically select, during runtime, the most suitable transformation from a set of rank-value transformations for constructing the linear-quadratic model. The chosen transformation maximizes the quality of the resulting model.

Algorithm 1 Model selection based on transformations

Require: m transformations of rank values $\{T_1, \dots, T_m\}$ of size n with $T_j = \{t_1^j, \dots, t_n^j\}$, model archive $X = \{x_{1:n}, \dots, x_{n:n}\}$ sorted by true fitness, lq learning model, k , the number of cross-validation sampling steps.

Ensure: $\forall j \in \{1, \dots, m\} \quad t_1^j < t_2^j < \dots < t_n^j$

- 1: $\tau \leftarrow \{\}$
- 2: **for** $j \leftarrow 1$ to m **do**
- 3: $S_j \leftarrow \{(x_{1:n}, t_1^j), \dots, (x_{n:n}, t_n^j)\}$
- 4: $\tau_j \leftarrow \mathbf{k}\text{-Fold}\text{-Cross}\text{-Validation}(S_j, k, \text{lq})$
- 5: Append τ_j to τ
- 6: **end for**
- 7: Identify T_{best} with the best kendall- τ
- 8: Train lq on S_{best} to obtain m_{best}
- 9: **return** m_{best}

Transformation selection. We perform a cross-validation process to evaluate the ranking quality of each transformation/model. Specifically, we use an implementation of the k -fold cross-validation method from the library `scikit-learn`. The quality of the transformation/model is measured by the Kendall's τ .

Independent transformations. When transformations are applied without incorporating any information from the original function values, aside from their ranking, the resulting algorithm remains invariant to strictly increasing transformations of the objective function. This foundational structure can then be combined with any learning method to develop an invariant surrogate-assisted CMA-ES variant.

As a proof of concept, we examine two algorithms: the basic Rank-lq-CMA-ES and an extended variant, where additionally the original function values are used. Remark that the function values *are* an increasing transformation of the ranks. Hence we use two datasets, defined as $\{t_i^1 = i, t_i^2 = f(x_{i:n})\}$. Since the transformation set includes the original function values, the extended variant is not invariant to strictly increasing transformations of the objective function. The following

sections detail the experiments conducted on both variants of Rank-lq-CMA-ES, encompassing both the experimental setup and results.

6 COCO benchmarking

COCO Experimental Setup. We evaluate Rank-lq-CMA-ES against CMA-ES, lq-CMA-ES, and S^* ACM-ES using the `bbob` testbed in the COCO platform [7]. Rank-lq-CMA-ES in both variants uses an implementation¹ based on the code from [6]. The rest of the performance data is taken from COCO archive. Following [6], algorithms run until reaching `maxfevals` = $10^4 \times$ dimension function evaluations. IPOPOP restarts are used, doubling the population size before each restart [2]. The initial solution per restart is set by the `initial_solution_proposal` method from `cocoex`, with an all-zeros solution evaluated at the start. The initial step size is 2, and no parameter tuning was performed.

COCO plots. The runtime distributions (in function evaluations) to reach specific function value targets are shown in Figure 1 for dimensions 3, 5, 10, and 20, aggregated over all test functions in the testbed. The results include also data from COCO’s data archive for CMA-ES (CMA-ES-2019), lq-CMA-ES (lq-CMA-ES), and S^* ACM-ES (IPOPsaACM), along with results of Rank-lq-CMA-ES (`ranklq-CMA-ES`) and the extended variant (`cvranklq-CMA-ES`). We use `lq-xcma`, a re-run of lq-CMA-ES from [6], as the baseline instead of lq-CMA-ES to account for changes observed between COCO’s dataset and our implementation.

The `cvranklq-CMA-ES` performs similarly to `lq-CMA-ES`, while `ranklq-CMA-ES` experiences a performance decline, which worsens in higher dimensions, aligning with CMA-ES at dimension 20. Compared to S^* ACM-ES, Rank-lq-CMA-ES shows advantages in lower dimensions (up to 5), but S^* ACM-ES scales better and outperforms Rank-lq-CMA-ES in 20-D.

7 Invariance Experiments

Setup. We empirically investigate the invariance behavior of the extended Rank-lq-CMA-ES algorithm with regard to strictly increasing transformations of the objective function. We follow the experimental setup outlined in [6]. We apply the CMA-ES, lq-CMA-ES, and Rank-lq-CMA-ES with both variants to the transformation $f \mapsto (f - f^{\text{opt}})^\alpha$ across various problems sourced from the `bbob` testbed in COCO. We take α in a range a little wider than $[0.1, 10]$. Initial solutions are sampled from a normal distribution $x_0 \sim \mathcal{N}(x^{\text{opt}} - 2, 0.1)$, with an initial step size set to 1. To account for stochastic effects, each experiment was repeated $1 + \lfloor 20/\text{dimension} \rfloor$ times.

Figure 2 shows the average runtime in terms of number of function evaluations to reach a target $(f - f^{\text{opt}})^\alpha \leq 10^{-\alpha}$ against α (in x axis). We can see that CMA-ES is invariant w.r.t α .

The Rank-lq-CMA-ES implementation is also invariant. It shows a little more variance than CMA-ES. This variance can be attributed to the sensitivity of the algorithm to the initial conditions in particular the initial training set. We confirmed that by running the same experiments and removing the stochastic effects by fixing the initial conditions for all α -values (not shown), we got perfect horizontal lines² for both CMA-ES and Rank-lq-CMA-ES.

The Extended Rank-lq-CMA-ES: Striking a Balance. The extended variant of Rank-lq-CMA-ES strikes a balance between invariance and performance, as illustrated by the magenta curve labeled `cvranklq-CMA-ES`. This extended variant achieves performance comparable to lq-CMA-ES on raw problems ($\alpha = 1$) while maintaining robustness on transformed problems, closely aligning with the performance of the strictly invariant algorithm `ranklq-CMA-ES`. However, there are still some rare cases where the extended variant performs slightly worse than both Rank-lq-CMA-ES and lq-CMA-ES.

8 Summary and Conclusion

We introduce a surrogate-assisted CMA-ES using linear-quadratic regression on ranks, ensuring invariance to strictly increasing transformations of the objective function. Unlike previous rank-based approaches, our method performs regression on rank data instead of using it as constraints for an SVM classifier. While it shows some advantages over CMA-ES, it generally underperforms compared to both lq-CMA-ES and S^* ACM-ES. To address this, we propose an extension that applies transformed rank values, selecting a strictly increasing transformation from a predefined set via cross-validation. An example demonstrates improved performance, highlighting the potential of this approach despite deviation from strict invariance. This opens avenues for research on optimizing transformations to enhance algorithm performance.

9 Acknowledgement

This research was conducted with the support of the consortium in Applied Mathematics CIROQUO, gathering partners in industry and academia, in particular Storengy, for the development of advanced methods for Computer Experiments.

References

- [1] Richard Allmendinger, Michael T. M. Emmerich, Jussi Hakanen, Yaochu Jin, and Enrico Rigoni. 2017. Surrogate-assisted multicriteria optimization: Complexities, prospective solutions, and business case. *Journal of Multi-Criteria Decision Analysis* 24, 1-2 (2017), 5–24. doi:10.1002/mcda.1605 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/mcda.1605
- [2] Anne Auger and Nikolaus Hansen. 2005. A restart CMA evolution strategy with increasing population size. In *2005 IEEE congress on evolutionary computation*, Vol. 2. IEEE, 1769–1776.
- [3] Anne Auger, Marc Schoenauer, and Nicolas Vanhaccke. 2004. LS-CMA-ES: A second-order algorithm for covariance matrix

¹Python implementation is available at <https://github.com/mgharafi/rank-lq-cma-es> as well as the full COCO performance data

²Disabling the optional parameter `f_offset` in lq-CMA-ES was required to have perfect horizontal lines. This adjustment was to address numerical issues, specifically flat fitness, caused by large numbers when f-values were raised to $\alpha \approx 5$.

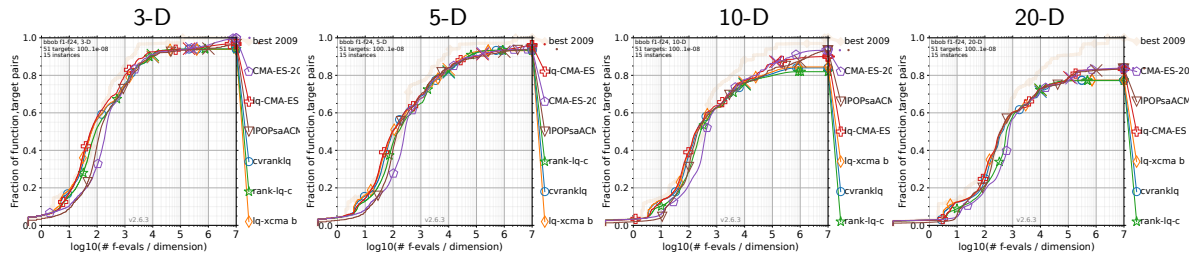


Figure 1: Bootstrapped empirical cumulative distribution of the number of f -evaluations divided by dimension (FEvals/DIM) for 51 targets with target precision in $10^{[-8..2]}$ for all functions and subgroups in 3, 5, 10 and 20-D. As reference algorithm, the best algorithm from BBOB 2009 is shown as light thick line with diamond markers.

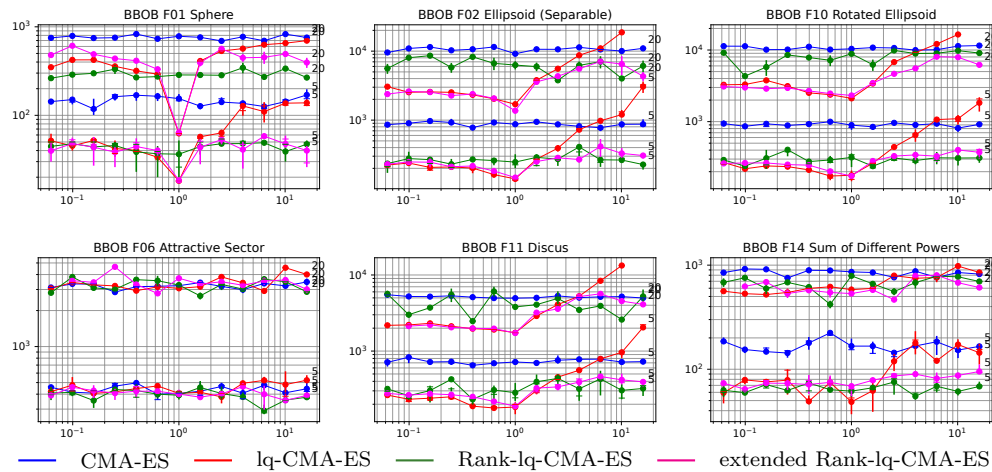


Figure 2: Average number of f -evaluations from $1 + \lfloor 20/n \rfloor$ runs to reach $f(x) < f^{opt} + 10^{-1}$ on $x \mapsto (f(x) - f^{opt})^\alpha$ versus α , in dimension $n = 5, 20$ with initial point coordinate-wise i.i.d. $(x_i^{opt} - 2, 0.12)$ -normally distributed. Bars indicate the interquartile range, the horizontal thick line is the median. The first row shows results for the convex-quadratic Sphere function and two Ellipsoid functions from the bboB implementation in COCO. The second row show results on three other unimodal but not convex quadratic functions from the bboB suite.

adaptation, Vol. 3242. 182–191. doi:10.1007/978-3-540-30217-9_19

[4] Lukáš Bajer, Zbyněk Pitra, Jakub Repický, and Martin Holeňa. 2019. Gaussian process surrogate models for the CMA evolution strategy. *Evolutionary computation* 27, 4 (2019), 665–697.

[5] Sylvain Gelly, Sylvie Ruetten, and Olivier Teytaud. 2007. Comparison-based algorithms are robust and randomized algorithms are anytime. *Evol. Comput.* 15, 4 (dec 2007), 411–434. doi:10.1162/evco.2007.15.4.411

[6] Nikolaus Hansen. 2019. A global surrogate assisted CMA-ES. In *Genetic and Evolutionary Computation Conference (GECCO 2019)*. ACM-balance, New York, NY, USA, 664–672.

[7] Nikolaus Hansen, Anne Auger, Raymond Ros, Olaf Mersmann, Tea Tušar, and Dimo Brockhoff. 2021. COCO: A platform for comparing continuous optimizers in a black-box setting. *Optimization Methods and Software* 36, 1 (2021), 114–144.

[8] Nikolaus Hansen and Andreas Ostermeier. 2001. Completely randomized self-adaptation in evolution strategies. *Evolutionary computation* 9, 2 (2001), 159–195.

[9] Nikolaus Hansen, Raymond Ros, Nicolas Mauny, Marc Schoenauer, and Anne Auger. 2011. Impacts of invariance in search: When CMA-ES and PSO face ill-conditioned and non-separable problems. *Applied Soft Computing* 11, 8 (2011), 5755–5769. doi:10.1016/j.asoc.2011.03.001

[10] Stefan Kern, Nikolaus Hansen, and Petros Koumoutsakos. 2006. Local meta-models for optimization using evolution strategies.

In *International Conference on Parallel Problem Solving from Nature*. Springer, 939–948.

[11] Ilya Loshchilov, Marc Schoenauer, and Michèle Sebag. 2010. Comparison-Based Optimizers Need Comparison-Based Surrogates. In *Parallel Problem Solving from Nature, PPSN XI*, Robert Schaefer, Carlos Cotta, Joanna Kolodziej, and Günter Rudolph (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 364–373. doi:10.1007/978-3-642-15844-5_37

[12] Ilya Loshchilov, Marc Schoenauer, and Michele Sebag. 2012. Self-adaptive surrogate-assisted covariance matrix adaptation evolution strategy. In *Proceedings of the 14th annual conference on Genetic and evolutionary computation*. 321–328.

[13] Ilya Loshchilov, Marc Schoenauer, and Michele Sebag. 2013. Bi-population CMA-ES algorithms with surrogate models and line searches. In *Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation (Amsterdam, The Netherlands) (GECCO '13 Companion)*. Association for Computing Machinery, New York, NY, USA, 1177–1184. doi:10.1145/2464576.2482696

[14] Thomas Runarsson. 2006. Ordinal Regression in Evolutionary Computation. 1048–1057. doi:10.1007/11844297_106

accepted 19 March 2025