Stochastic and deterministic algorithms for continuous black-box optimization

Wang, H.

Citation

Wang, H. (2018, November 1). *Stochastic and deterministic algorithms for continuous black-box optimization*. Retrieved from https://hdl.handle.net/1887/66671

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/66671

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

Leiden University
Repository

The handle http://hdl.handle.net/1887/66671 holds various files of this Leiden University dissertation.

Author: Wang, H.

Title: Stochastic and deterministic algorithms for continuous black-box optimization

Issue Date: 2018-11-01

Stochastic and Deterministic Algorithms for Continuous Black-Box Optimization

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 1 november 2018
klokke 10.00 uur

door

Hao Wang

geboren te Baoji, China
in 1989

Promotiecommissie

Promotor:	Prof. Dr. T.H.W. Bäck	
Co-promotor:	Dr. M.T.M. Emmerich	
Overige leden:	Prof. Dr. A. Plaat	(voorzitter)
	Prof. Dr. H. Trautmann	(WWU Münster and LIACS)
	Dr. C. Dörr	(CNRS and Sorbonne University, FR)
	Prof. Dr. S. Manegold	(secretaris, CWI and LIACS, NL)
	Dr. W.A. Kosters	
	Prof. X. Liu	(Brunel University, UK)

Copyright © 2018 Hao Wang.

This research is financially supported by the Dutch funding agency NWO, under project number 650.002.001 (the PROMIMOOC project), in collaboration with Tata Steel IJmuiden, BMW Group Regensburg, Centrum voor Wiskunde en Informatica (CWI) and MonetDB.

Figures and diagrams are generated using GGPLOT2, PGF/TIKZ and MATPLOTLIB.

Abstract

Continuous optimization is never easy: the exact solution is always a luxury demand and the theory of it is not always analytical and elegant. Continuous optimization, in practice, is essentially about the efficiency: how to obtain the solution with same quality using as minimal resources (e.g., CPU time or memory usage) as possible? In this thesis, the number of function evaluations is considered as the most important resource to save. To achieve this goal, various efforts have been implemented and applied successfully. One research stream focuses on the so-called stochastic variation (mutation) operator, which conducts an (local) exploration of the search space. The efficiency of those operator has been investigated closely, which shows a good stochastic variation should be able to generate a good coverage of the local neighbourhood around the current search solution. The first part (Chapter 2) of this thesis contributes on this issue by formulating a novel stochastic variation that yields good space coverage.

Alternative research stream approaches the efficiency issue differently: we should keep record of the evaluated solutions and re-use them as they carry partially information about the objective function. This leads to studies on the so-called surrogate modeling. Here, the second part (Chapter 3) of this thesis dives into the one specific surrogate modeling technique, called Kriging/Gaussian Process Regression (GPR). In addition, we try to keep a precise and theoretical treatment on Kriging/GPR and several improvements over it. Closely related to the surrogate modeling, it is crucial to exploit the surrogate model properly. A common approach is to take the so-called infill criteria/acquisition function, which measures the potential gain we could obtain by evaluate one candidate solution on the model. Lastly, the efficiency issues can also be tackled by generalizing a well-performing algorithm from a specific domain to a broader class of problems. One prominent example is to extend the gradient-based optimization algorithms (that are devised

for single objective problems) to the multi-objective scenario. The last part of this thesis (Chapter 5) is on this topic, where both the first- and second-order methods are generalized for multi-objective optimization problems.

Contents

List of Symbols

$\mathbf{1}$	Vector of ones, whose dimension is implied in context
$\mathbf{0}$	Vector of zeros, whose dimension is implied in context
$\mathbf{I}$	Identity matrix, whose shape is implied in context
$\mathbb{N}_{>0}$	Positive natural numbers
$\mathbb{R}^d$	d-dimensional Euclidean space
S	Search space/decision space/domain of the objective function
$\mathcal{X}$	Pareto efficient set
$P_{\mathcal{X}}$	Pareto front
$\mathcal{H}$	Hilbert space of functions $f\colon \mathrm{S} \to \mathbb{R}$
$L^2(\mathrm{S})$	Space of square-integrable functions on S
$\lVert\cdot\rVert$	Euclidean norm
$\lVert\cdot\rVert_{\mathbf{\Sigma}}$	Mahalanobis norm with respect to a covariance matrix $\mathbf{\Sigma}$
$\lVert\cdot\rVert_{\mathcal{H}}$	Norm in Hilbert space $\mathcal{H}$
$\lVert\cdot\rVert_{\infty}$	supremum norm
$\langle\cdot,\cdot\rangle$	Dot product in Euclidean spaces.
$\langle\cdot,\cdot\rangle_{\mathcal{H}}$	Inner product in Hilbert space $\mathcal{H}$
$\mathcal{U}(0,1)$	Uniform distribution over $[0,1]$
$\mathcal{N}(m,\sigma^2)$	Gaussian random variable with mean m and variance σ^2
$\boldsymbol{\mathcal{N}}(\mathbf{m},\mathbf{K})$	Gaussian random vector with mean $\mathbf{m}$ and covariance matrix $\mathbf{K}$

$\mathcal{GP}(m(\cdot), k(\cdot,\cdot))$ Gaussian process with mean function $m(\cdot)$ and kernel $k(\cdot,\cdot)$

$\Pr(\cdot)$ Probability of an event

$p(\cdot)$ Probability density function (p.d.f.)

$P(\cdot)$ Cumulative distribution function (c.d.f.)

$\mathbb{P}$ Probability measure

ϕ Probability density function of $\mathcal{N}(0,1)$

ϕ_{m,σ^2} Probability density function of $\mathcal{N}(m,\sigma^2)$

Φ Cumulative distribution function of $\mathcal{N}(0,1)$

Φ_{m,σ^2} Cumulative distribution function of $\mathcal{N}(m,\sigma^2)$

$\Phi^n_{\mathbf{m},\mathbf{K}}$ Cumulative distribution function of a multivariate Gaussian $\boldsymbol{\mathcal{N}}(\mathbf{m},\mathbf{K})$

$\mathbb{E}$ Expectation

Var Variance

Cov Covariance

$\perp\!\!\!\perp$ Statistical independence

σ_n^2 Variance of the white noise process

$\mathscr{A}$ Infill criteria/acquisition function

det Determinant of square matrices

κ Condition number of matrices