ARTICLE IN PRESS YJSCO:1710

Journal of Symbolic Computation ••• (••••) •••–•••

Contents lists available at ScienceDirect

Journal of Symbolic Computation

www.elsevier.com/locate/jsc

Critical points via monodromy and local methods

Abraham Martín del Campo [a,1], Jose Israel Rodriguez [b,1,2]

[a] *CIMAT, Jalisco s/n, Col. Valenciana, 36023 Guanajuato, Gto., Mexico*
[b] *Department of Statistics, University of Chicago, 5747 S. Ellis Avenue, Chicago, IL 60637, USA*

A R T I C L E I N F O

Article history:
Received 1 July 2015
Accepted 19 August 2015
Available online xxxx

Keywords:
Critical points
Monodromy
Homotopy continuation
Gradient descent homotopy
Maximum likelihood degree
Euclidean distance
Trace test

A B S T R A C T

A fundamental problem in many areas of applied mathematics and statistics is to find the best representative of a model by optimizing an objective function. This can be done by determining critical points of the function restricted to the model.
We compile ideas arising from numerical algebraic geometry to compute these critical points. Our method consists of using numerical homotopy continuation and a monodromy action on the total critical space to compute all of the complex critical points. To illustrate the relevance of our method, we apply it to the Euclidean distance function to compute ED-degrees and the likelihood function to compute maximum likelihood degrees.

1. Introduction

In science and engineering, it is common to work with models that can be described as the solutions of a parametrized system of polynomial equations. For such algebraic models $X \subset \mathbb{R}^n$, we are interested in the following polynomial optimization problem: given $u \in \mathbb{R}^n$, find those points $x^* \in X$ that optimize an objective function $\Psi(x, u) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Common objective functions seen in applications include norms, distances, and other statistical functions. In this paper, we focus on the case when Ψ is a quadratic norm (such as the Euclidean distance) and when Ψ is the likelihood function.

E-mail addresses: abraham.mc@cimat.mx (A. Martín del Campo), JoIsRo@UChicago.edu (J.I. Rodriguez).
[1] Partly supported by 2014 NIMS Thematic Program on Applied Algebraic Geometry.
[2] Partially supported by the National Science Foundation under Award No. DMS-1402545.

http://dx.doi.org/10.1016/j.jsc.2016.07.019
0747-7171/

As a consequence of Fermat's theorem on stationary points, the method of Lagrange multipliers can be used to compute the set of critical points. Thus, our problem turns into finding the points $x \in X$ for which the gradient $\nabla\Psi(x, u)$ is orthogonal to the tangent space T_xX of X at the point x. We recast these constraints as the solutions to a system of polynomial equations. In the algebraic closure, the number d of critical points is an invariant that gives information about the algebraic complexity of the optimization problem. Finding the number d for different objective functions and models is a challenging problem. When Ψ is the Euclidean distance, the number d is called the Euclidean distance degree (ED-degree), and when Ψ is the likelihood function, the number d is the maximum likelihood degree (ML-degree).

Finding these degrees d is an active area of research. There are already some fundamental results about these algebraic degrees and the geometry behind them. A general degree theory for the ED-degree was introduced by Draisma et al. (2015) and for the ML-degree in Catanese et al. (2006), Hoşten et al. (2005), Huh and Sturmfels (2014). For some special cases, it is possible to find formulas for d using techniques from algebraic geometry (Agostini et al., 2014; Huh, 2013; Huh and Sturmfels, 2014; Lee, 2014; Nie and Ranestad, 2009; Ottaviani et al., 2014; Piene, 2015; Rodriguez, 2014; Rodriguez and Wang, 2015), some are summarized in Ranestad (2012). However, this is not always the case, and other methods are necessary.

There are algorithms proposed for identifying critical points based on local methods. Such methods involving Gröbner bases include (Faugère et al., 2012; Ottaviani et al., 2014; Spaenlehauer, 2014). From the numerical algebraic geometry perspective, those in (Wu and Reid, 2013; Bates et al., 2014) can be used, and involve SVD decomposition and regeneration (respectively) to find critical points for the computation of witness sets. For computing the real critical points, one method is to compute all complex solutions and determine the real solutions among them. This can be avoided by using local methods to determine local optima and using heuristics to decide if the local optima is in fact global. In these cases, one can study the expected number of real critical points when u is drawn from a given probability distribution. This has been studied for the ED-degree under the name of *average ED-degree* (e.g. Drusvyatskiy et al., 2015). Recently, a probabilistic method was proposed by Rodriguez and Tang (2015), where they classify the real critical points for the likelihood function.

In this paper, we propose a different numerical method to compute all of the complex critical points. Our method consists of two parts and complements those used in (Gross and Rodriguez, 2014; Hauenstein et al., 2014). The first part consists of exploiting a monodromy group action to randomly explore the variety of critical points. In this random exploration, we determine a subset of critical points. The second part consists of a trace test to determine, with probability one, if the computed subset contains *all* of the critical points.

This paper is organized as follows. In the next section, we give a geometric formulation of the critical equations along with a concrete formulation of the critical equations. In the following section, we review monodromy for a parameterized polynomial system. This includes a description of each part of our method, including a trace test. In Section 4, we summarize the steps of our procedure. We conclude by using these methods to reproduce some known ED-degrees and ML-degrees, and compare the times of those computations against our method. Throughout the paper, we include implementation subsections so the reader can use the methods in their own research. Supplementary materials can be found on the second author's website (Martín del Campo and Rodriguez, 2015).

We end this introduction with an illustrating example to the critical points problem.

Example 1. Let X be the ellipse defined as the algebraic variety of the points $x = (x_1, x_2) \in \mathbb{R}^2$ that satisfy the polynomial equation

$$1744x_1^2 - 2016x_1x_2 - 2800x_1 + 1156x_2^2 + 2100x_2 + 1125 = 0.$$

For a (general) choice of $u = (u_1, u_2) \in \mathbb{R}^2$, we are interested in the optimization problem

$$\min \Psi(x, u) = (x_1 - u_1)^2 + (x_2 - u_2)^2 \quad \text{subject to } x \in X. \tag{1}$$

Here, the objective function $\Psi(x, u)$ is the square of the Euclidean distance between the (fixed) point $u \in \mathbb{R}^2$ and the ellipse X. The critical points of (1) are those points $x^* \in X$ whose tangent is perpen-

Please cite this article in press as: Martín del Campo, A., Rodriguez, J.I. Critical points via monodromy and local methods. J. Symb. Comput. (2016), http://dx.doi.org/10.1016/j.jsc.2016.07.019