Computer Physics Communications 185 (2014) 572–577

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Tracking local optimality for cost parameterized optimization problems

Yueh-Cheng Kuo [a], Tsung-Lin Lee [b,*]

[a] Department of Applied Mathematics, National University of Kaohsiung, Kaohsiung 811, Taiwan
[b] Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung 804, Taiwan

ARTICLE INFO

Article history:
Received 19 February 2013
Received in revised form
10 August 2013
Accepted 23 October 2013
Available online 29 October 2013

Keywords:
Continuation methods
Optimization problems
Bifurcation

ABSTRACT

In this paper, a procedure for computing local optimal solution curves of the cost parameterized optimization problem is presented. We recast the problem to a parameterized nonlinear equation derived from its Lagrange function and show that the point where the positive definiteness of the projected Hessian matrix vanishes must be a bifurcation point on the solution curve of the equation. Based on this formulation, the local optimal curves can be traced by the continuation method, coupled with the testing of singularity of the Jacobian matrix. Using the proposed procedure, we successfully compute the energy diagram of rotating Bose–Einstein condensates.

© 2013 Published by Elsevier B.V.

1. Introduction

Cost parameterized optimization problems widely arise in many fields of science and engineering. Mathematically the problem can be described as

$$\begin{aligned} &\text{Minimize (or Maximize) } f(\mathbf{x}, p) \\ &\text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}, \end{aligned} \tag{1.1}$$

where $f(\mathbf{x}, p) \in C^2(\mathbb{R}^n \times [a, b])$ is a parameterized cost function with parameter $p \in [a, b]$ and $\mathbf{g}(\mathbf{x}) \equiv (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T$ with $g_j(\mathbf{x}) \in C^2(\mathbb{R}^n)$ for $j = 1, \ldots, m$ are constraints. The local optimal solution set within a certain parameter interval is desired. In this paper, we propose an efficient procedure to track local optimal curves within a prescribed parameter interval.

When the parameter p is given, various optimization algorithms, e.g. conjugate gradient methods, interior-point methods, gradient projection methods, etc., have been well developed to find a local optimal solution [1,2]. For computing the local optimal curves, applying these methods for a large number of given parameters is difficult to determine the end points of the maximal parameterized interval. Even worse, this exhausted approach may miss some critical optimal solutions especially when multi-stable regions exist. Fig. 1 illustrates a bi-stable region. A local optimal curve is parameterized over (p_1^ℓ, p_1^r), and the other is parameterized over (p_2^ℓ, p_2^r), where end points $p_1^\ell < p_2^\ell < p_1^r < p_2^r$. The region (p_2^ℓ, p_1^r) has two distinct local optimal solutions.

The continuation method is a powerful tool for tracking the solution curve of an equation with a parameter. The advantage of employing continuation method to problem (1.1) is its capability to trace the continuity property of optimal solution. We consider a parameterized nonlinear equation derived from the Lagrange function associated with (1.1). For a given parameter, the solutions of this equation are the candidates of local optimality. From the sufficient optimality conditions, the positive definiteness of its corresponding projected Hessian matrix guarantees the solution to be locally optimal. However, the additional computations for determining the positive definiteness of matrices are quite expensive. We give an invariant theorem (see Theorem 2.5) of inertia between the Jacobian matrix and projected Hessian matrix, and show that a bifurcation branch occurs at the point where the positive definiteness of the projected Hessian matrix vanishes. Therefore, we can track the solution curve by continuation method until a singular point is met. It is clear that checking the singularity of Jacobian is much cheaper than detecting the positive definiteness of a matrix, so the resulting algorithm is more efficient.

In the following subsection, we give a brief outline of the continuation method. In Section 2, we discuss the related theorems and propose a procedure for tracking the local optimal curves over a prescribed interval. The numerical results of computing the energy diagram of rotating Bose–Einstein condensates by the procedure will be presented in Section 3.

* Corresponding author.
E-mail addresses: yckuo@nuk.edu.tw (Y.-C. Kuo), leetsung@math.nsysu.edu.tw (T.-L. Lee).

0010-4655/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.cpc.2013.10.021

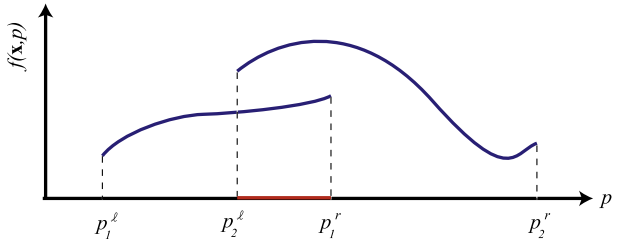

Fig. 1. The cost value of two local optimal curves. One is parameterized over (p_1^ℓ, p_1^r), and the other is parameterized over (p_2^ℓ, p_2^r), where end points $p_1^\ell < p_2^\ell < p_1^r < p_2^r$. The bi-stable region is (p_2^ℓ, p_1^r).

1.1. A framework of continuation method

Continuation methods are numerical schemes aiming to compute approximate solutions of a system of parameterized nonlinear equations. Here, we give a brief account of the main ideas. The detailed description of the methods can be found, for example, in [3,4].

In general, a parameterized nonlinear equation system can be denoted as

$$\mathbf{G}(\mathbf{z}, p) = \mathbf{0}, \tag{1.2}$$

where $\mathbf{G} : \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}^N$ is a continuously differentiable function, $\mathbf{z} \in \mathbb{R}^N$, $p \in \mathbb{R}$ is a natural parameter. Parameterizing $\mathbf{z}$ by p, a solution curve of (1.2) can be defined as

$$\mathcal{C}_I = \left\{(\mathbf{z}(p), p) \in \mathbb{R}^{N+1} |\ \mathbf{G}(\mathbf{z}(p), p) = \mathbf{0},\ p \in I\right\}, \tag{1.3}$$

where I is an interval. To follow the solution curve (1.3), a continuation method usually takes a prediction–correction approach. In particular, starting from the point $\mathbf{z}_i = \mathbf{z}(p_i) \in \mathbb{R}^N$ which is a solution point lying (approximately) on a solution curve $\mathcal{C}_I$, the prediction and correction steps are described as follows:

- in the prediction step, the Euler predictor

  $$\mathbf{z}_{i+1,1} = \mathbf{z}_i + h_i \dot{\mathbf{z}}_i,$$

  is computed. Here, $\dot{\mathbf{z}}_i = \dot{\mathbf{z}}(p_i)$ is the tangent vector of the solution curve at $\mathbf{z}_i$ and $h_i > 0$ is a suitable step length. Set $p_{i+1} = p_i + h_i$. To compute $\dot{\mathbf{z}}_i$, we solve the linear system

  $$\mathbf{G}_{\mathbf{z}}(\mathbf{z}_i, p_i)\dot{\mathbf{z}}_i = -\mathbf{G}_p(\mathbf{z}_i, p_i),$$

  which is obtained by differentiating (1.2) with respect to p;
- in the correction step, Newton's method is usually used to compute the approximate solution of $\mathbf{G}(\mathbf{z}, p_{i+1}) = \mathbf{0}$ with initial value $\mathbf{z}_{i+1,1}$. That is, for the correction vector δ_l, the iteration $\mathbf{z}_{i+1,l+1} = \mathbf{z}_{i+1,l} + \delta_l$ is computed for $l = 1, 2, \ldots$ until a convergence criterion is satisfied for $l = l_\infty$. Finally, $\mathbf{z}_{i+1} = \mathbf{z}_{i+1,l_\infty}$ is taken as a new approximate solution on the solution curve $\mathcal{C}_I$.

2. The tracking procedure

In the following, we only discuss the cost parameterized minimization problem:

$$\begin{aligned} &\text{Minimize } f(\mathbf{x}, p) \\ &\text{subject to } \mathbf{g}(\mathbf{x}) = \mathbf{0}, \end{aligned} \tag{2.1}$$

where $f(\mathbf{x}, p) \in C^2(\mathbb{R}^n \times [a, b])$ is a parameterized cost function with parameter p and $\mathbf{g}(\mathbf{x}) \equiv (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))^T$ with $g_j(\mathbf{x}) \in C^2(\mathbb{R}^n)$ for $j = 1, \ldots, m$. The theories and the algorithms for the maximization model can be derived in a similar way. Throughout this paper, we use the bold face letter or symbol to denote the matrix or vector. For a matrix $\mathbf{A}$, $\mathbf{A}^H$ and $\mathbf{A}^T$ denote the conjugate transpose and transpose of $\mathbf{A}$, respectively. For a Hermitian matrix $\mathbf{A}$, we use $\mathbf{A} > 0$ ($\mathbf{A} \geqslant 0$) to denote that $\mathbf{A}$ is positive definite (positive semi-definite). For $\mathbf{A} \in \mathbb{R}^{n\times m}$, $\mathcal{N}(\mathbf{A})$ and $\mathbf{A}^\dagger \in \mathbb{R}^{m\times n}$ denote the null space of $\mathbf{A}$ and the pseudoinverse of $\mathbf{A}$, respectively.

2.1. Preliminaries

We shall introduce some definitions and preliminary theorems in this subsection. The feasible domain of (2.1) is the set of points $\mathbf{x}$ that satisfy the constraints; namely,

$$\mathcal{F} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{g}(\mathbf{x}) = \mathbf{0}\}.$$

A point $\mathbf{x}_* \in \mathcal{F}$ is a (strict) local solution of the optimization problem (2.1) with a parameter p if there is a neighborhood $\mathcal{N}$ of $\mathbf{x}_*$ such that $f(\mathbf{x}, p) \geqslant f(\mathbf{x}_*, p)$ ($f(\mathbf{x}, p) > f(\mathbf{x}_*, p)$) for all $\mathbf{x} \in \mathcal{N} \cap \mathcal{F}$ with $\mathbf{x} \neq \mathbf{x}_*$. A curve $\Gamma = \{\mathbf{x}(p) : p \in I\} \subset \mathcal{F}$ is called a local minimal curve of problem (2.1) if $I \subseteq [a, b]$ is an interval and $\mathbf{x}(p)$ is a local solution of problem (2.1) with each $p \in I$. We say that the **linear independence constraint qualification** (**LICQ**) holds at $\mathbf{x} \in \mathcal{F}$ if

$$\nabla \mathbf{g}(\mathbf{x}) \equiv [\nabla g_1(\mathbf{x}), \ldots, \nabla g_m(\mathbf{x})] \in \mathbb{R}^{n\times m}$$

is of full column rank. The corresponding **Lagrange function** to the optimization problem (2.1) is

$$\mathcal{L}(\mathbf{x}, \mathbf{v}, p) = f(\mathbf{x}, p) + \sum_{j=1}^{m} v_j g_j(\mathbf{x}),$$

where $\mathbf{v} = (v_1, \ldots, v_m)^T \in \mathbb{R}^m$ with the Lagrange multipliers v_i's.

For a given parameter $p \in [a, b]$, we say **Karush–Kuhn–Tucker** (**KKT**) **conditions** hold at $\mathbf{x}$, if there exists a vector $\mathbf{v} \in \mathbb{R}^m$ such that

$$\begin{aligned} &\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mathbf{v}, p) = \nabla_{\mathbf{x}} f(\mathbf{x}, p) + \sum_{j=1}^{m} v_j \nabla g_j(\mathbf{x}) = \mathbf{0} \quad \text{and} \\ &\mathbf{g}(\mathbf{x}) = \mathbf{0}. \end{aligned} \tag{2.2}$$

In this case, $\mathbf{x} \in \mathbb{R}^n$ is called a **KKT point** for the optimization problem (2.1) with a parameter p.

Let $\mathbf{P}(\mathbf{x}) \in \mathbb{R}^{n\times(n-m)}$ be a matrix whose columns form an orthonormal basis of $\mathcal{N}(\nabla \mathbf{g}(\mathbf{x})^T)$. Then the **projected Hessian matrix** is defined as

$$\mathbf{H}(\mathbf{x}, \mathbf{v}, p) = \mathbf{P}(\mathbf{x})^T \nabla^2_{\mathbf{xx}} \mathcal{L}(\mathbf{x}, \mathbf{v}, p)\mathbf{P}(\mathbf{x}) \in \mathbb{R}^{(n-m)\times(n-m)}. \tag{2.3}$$

Theorems 2.1 and 2.2 are the direct results of the standard necessary and sufficient conditions theorem for (non-parameterized) optimization problem, respectively.

Theorem 2.1 (*Necessary Conditions*). *Suppose that $\mathbf{x}_*$ is a local solution of optimization problem (2.1) with a given parameter p and that the LICQ holds at $\mathbf{x}_*$. Then there is a Lagrange multiplier vector $\mathbf{v}_* \in \mathbb{R}^m$ such that KKT conditions (2.2) are satisfied and the projected Hessian matrix $\mathbf{H}(\mathbf{x}_*, \mathbf{v}_*, p)$ is positive semi-definite.*

Theorem 2.2 (*Sufficient Conditions*). *Suppose that the parameter $p \in [a, b]$ is given and $\mathbf{x}_* \in \mathbb{R}^n$ is a KKT point for optimization problem (2.1) with Lagrange multiplier vector $\mathbf{v}_* \in \mathbb{R}^m$. If the projected Hessian matrix $\mathbf{H}(\mathbf{x}_*, \mathbf{v}_*, p)$ is positive definite, then $\mathbf{x}_*$ is a strict local solution of problem (2.1).*

Let $\mathcal{C}_I$ in (1.3) be a solution curve of (1.2). A point $(\mathbf{z}_0, p_0) \in \mathcal{C}_I$ with p_0 being an interior point of I is said to be a **bifurcation point** of (1.2) if every ball $B_\rho(\mathbf{z}_0, p_0)$ of radius $\rho > 0$ contains solutions of (1.2) which are not on $\mathcal{C}_I$. The following bifurcation test theorem can be found in [4].

Theorem 2.3 (*[4, Bifurcation Test]*). *Let $\mathcal{C}_I$ be a solution curve of (1.2) parameterized by $p \in I$. If $\det(\mathbf{G}_{\mathbf{z}}(\mathbf{z}(p), p))$ changes sign at some interior point p_0 of I, then $(\mathbf{z}(p_0), p_0)$ is a bifurcation point of (1.2).*

Download English Version:

https://daneshyari.com/en/article/501676

Download Persian Version:

https://daneshyari.com/article/501676

Daneshyari.com