Automatica 71 (2016) 135–142

Contents lists available at ScienceDirect

Automatica

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

Brief paper

Necessary and sufficient Karush–Kuhn–Tucker conditions for multiobjective Markov chains optimality☆

Julio B. Clempner [1]

Centro de Investigaciones Economicas, Administrativas y Sociales, Instituto Politecnico Nacional, Lauro Aguirre 120, col. Agricultura, Del. Miguel Hidalgo, 11360, Mexico City, Mexico

ARTICLE INFO

Article history:
Received 25 April 2015
Received in revised form
14 April 2016
Accepted 16 April 2016
Available online 31 May 2016

Keywords:
Multiobjective
Karush–Kuhn–Tucker
Tikhonov
Pareto
Markov chains

ABSTRACT

The solution concepts proposed in this paper follow the Karush–Kuhn–Tucker (KKT) conditions for a Pareto optimal solution in finite-time, ergodic and controllable Markov chains multi-objective programming problems. In order to solve the problem we introduce the Tikhonov's regularizator for ensuring the objective function is strict-convex. Then, we consider the c-variable method for introducing equality constraints that guarantee the result belongs to the simplex and satisfies ergodicity constraints. Lastly, we restrict the cost-functions allowing points in the Pareto front to have a small distance from one another. The computed image points give a continuous approximation of the whole Pareto surface. The constraints imposed by the c-variable method make the problem computationally tractable and, the restriction imposed by the small distance change ensures the continuation of the Pareto front. We transform the multi-objective nonlinear problem into an equivalent nonlinear programming problem by introducing the Lagrange function multipliers. As a result we obtain that the objective function is strict-convex, the inequality constraints are continuously differentiable and the equality constraint is an affine function. Under these settings, the KKT optimality necessary and sufficient conditions are elicited naturally. A numerical example is solved for providing the basic techniques to compute the Pareto optimal solutions by resorting to KKT conditions.

1. Introduction

1.1. Brief review

The optimization problems in real life are frequently multi-objective where several competing objective functions have to be minimized at the same time. Classical application areas are engineering, economics, others. Such optimization problems have a very large solution set. The main focus is on finding a minimal solution by applying objective numerical calculations involving subjective decisions made by a decision maker. The goal is to represent the whole efficient set to provide the decision maker with a practical understanding of the problem structure. There has been a great deal of effort by the researchers in the area for developing methods to generate an approximation of the Pareto front, see e.g. Clempner and Poznyak (2016), Dutta and Kaya (2011), Engau and Wiecek (2007), Haimes and Chankong (1979), Mueller-Gritschneder, Graeb, and Schlichtmann (2009) and Zitzler, Knowles, and Thiele (2008). The efficiency of the solution set depends significantly in the application approach. A Multi-objective problem (MOP) can be transformed into a tractability problem applying the KKT conditions. However, several problems arise. The resulting utility functions are in general nonconvex. In addition, the formulation of the problem after applying the KKT condition is in general non-linear in the case of Pareto optimality formulation. Moreover, the solution allows global and local Pareto optimal points. As a result, the formulation we obtained by using the Karush–Kuhn–Tucker conditions is only necessary for our original problem. Gatti, Rocco, and Sandholm (2013) prove that the KKT conditions lead to another set of necessary conditions that are not sufficient. The main reason of obtaining a sufficient formulation for KKT condition into the Pareto optimality formulation is to achieve a unique solution for every Pareto point.

1.2. Related work

Different methods for necessary and sufficient conditions for KKT optimality have been proposed in the literature. They are based on different types of suppositions as to the properties and

☆ The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Hyeong Soo Chang under the direction of Editor Ian R. Petersen.

E-mail address: julio@clempner.name.

[1] Tel.: +52 55 57296000x105; fax: +52 55 57296000x63105.

http://dx.doi.org/10.1016/j.automatica.2016.04.044
0005-1098/

form of the problem. Several are based on a different techniques of the original problem and different conditions are imposed to both, the original functions and the scalarization parameters.

For the interested reader we list some important works (excellent surveys related to multi-objective Markov decision process can be found in Chatterjee, Majumdar, and Henzinger (2006) and Roijers, Vamplew, Whiteson, and Dazeley (2013) and for multi-objective in Zitzler et al. (2008)). Durinovic, Lee, Katehakis, and Filar (1986) considered a multi-objective Markovian decision process based on the average reward function to characterize the complete sets of efficient policies, efficient deterministic policies and efficient points using linear programming. Barrett and Narayanan (2008) described an algorithm that can learn optimal policies for all linear preference assignments over the multiple reward criteria at once. Vamplew, Dazeley, Barker, and Kelarev (2009) discussed the advantages gained from applying stochastic policies to multi-objective tasks and examined a particular form of stochastic policy known as a mixture policy proposing two different methods. Roijers et al. (2014) proposed two algorithms for computing the relative importance of the objectives: (a) considering an anytime method, approximate optimistic linear support, that incrementally builds up a complete set of epsilon-optimal plans, exploiting the piecewise linear and convex shape of the value function and, (b) an approximate anytime method, scalarized sample incremental improve, that employs weight sampling to focus on the most interesting regions in weight space, as suggested by a prior over preferences. Pirotta, Parisi, and Restelli (2015) presented an idea to exploit a gradient-based approach to optimize the parameters of a function that defines a manifold in the policy parameter space so that the corresponding image in the objective space gets as close as possible to the Pareto frontier. In mathematical programming problems when are made use of functions, the focus is placed into substituting convex functions with new class of generalized convex functions. The goal is obtaining a solution through an optimality condition.

1.3. Main results

This paper focuses on the conventional Markov chains multi-objective programming problems in which are considered the Karush–Kuhn–Tucker (KKT) conditions for a Pareto optimal solution. Our aim is to study and provide the conditions on the involved objective functions such that these are sufficient and necessary in order for a feasible point satisfying KKT type conditions to be a strong Pareto policy for the MOP. The main results are as follows. We provide necessary and sufficient of KKT conditions for efficiency to multi-objective programs introducing the Tikhonov's regularizator for ensuring that the objective function is strict-convex. Because the optimization problem is strict convex for any fixed value, then our results assert that efficient solutions can be found by strict convex optimization, something that does not necessary hold for efficient solutions in general. We introduce a more restrictive concept of efficiency by ensuring strong Pareto policies using the Tikhonov's regularizator. Then, all efficient solutions can be found by minimization of strictly convex functions. We consider the c-variable method for introducing the equality constraints that ensure the result belongs to the simplex and it satisfies ergodicity constraints. The constraints imposed by the c-variable method make the problem computationally tractable and, the restriction imposed by the small distance change ensures the continuation of the Pareto front. We restrict the cost-functions allowing points in the Pareto front to have a small distance from one another. We transform the multi-objective nonlinear problem into an equivalent nonlinear programming problem by introducing the Lagrange function multipliers. We obtain that the objective function is strict-convex, the inequality constraints are continuously differentiable and the equality constraint is an affine function. Under these settings, the KKT optimality necessary and sufficient conditions are elicited naturally. In addition, we present the convergence conditions and compute the estimate rate of convergence of variables μ and δ corresponding to the step size parameter of the gradient method and the Tikhonov's regularization respectively.

1.4. Organization of the paper

In the remainder we proceed as follows: in Section 2 we recall the basic concepts in Markov chains and multi-objective optimization for Markov chains. In Section 3, we discuss the sufficiency and necessity for optimal conditions. In Section 4 we present a solution concept for multi-objective optimization problem and the KKT optimality conditions for the problem are derived. Also a numerical example is solved for providing the basic techniques to compute the Pareto optimal solutions by resorting to KKT conditions in Section 5. We conclude with some remarks on the presented approach in Section 6.

2. Basic notations and concepts

2.1. Controllable Markov process

A *controllable Markov chain* is a 5-tuple $MC = \{S, A, A(s), \Pi\}$ where S is a finite set of states, $S \subset \mathbb{N}$, endowed with discrete topology; A is the set of actions, which is a metric space. For each $s \in S$, $A(s) \subset A$ is the non-empty set of admissible actions at state $s \in S$. Without loss of generality we may take $A = \cup_{s\in S} A(s)$; $\mathbb{K} = \{(s, a) | s \in S, a \in A(s)\}$ is the set of admissible state-action pairs, which is a measurable subset of $S \times A$; $\Pi(k) = \left[\pi_{(ij|k)}\right]$ is a stationary transition controlled matrix, where

$$\pi_{(ij|k)} \equiv P\left(s(n+1) = s_{(j)} | s(n) = s_{(i)}, a(n) = a_{(k)}\right)$$

representing the probability associated with the transition from state $s_{(i)}$ to state $s_{(j)}$ under an action $a_{(k)} \in A\left(s_{(i)}\right)$ $(k = 1, \ldots, M)$ at time $n \in \mathbb{N}$: 0, 1, ... The relations $\pi_{(i,j|k)} \geq 0$ and $\sum_{j=1}^{N} \pi_{(i,j|k)} = 1$ are satisfied. A *Markov process* is a tuple $MP = \{MC, J\}$ where MC is a controllable Markov chain and $J : S \times \mathbb{K} \to \mathbb{R}^n$ is a cost function.

We will restrict attention to stationary policies. A policy d is a (measurable) rule for choosing actions which, at each time $n \in \mathbb{N}$, may depend on the current state and on the record of previous states and actions. The class of all policies is denoted by D and, given the initial state $s \in S$ and the policy d being used for choosing actions, the distribution of the state-action process $\{(s_n, a_n)\}$ is uniquely determined. Following, we will denote by P and E respectively the probability measure and the expectation operator induced by the policy d. Next, define $\mathbb{F} := \prod_{s\in S} A(s)$ and notice that $\mathbb{F}$ is a compact metric space in the product topology which consists of all functions $f : S \to A$ such that $f(s) \in A(s)$ for each $s \in S$. A policy d is *stationary* iff there exists $f \in \mathbb{F}$ such that the equality $A_n = f(s_n)$ is always valid under d, i.e. $d_{(k|i)}(n) = d_{(k|i)}$. Also, under the action of any stationary policy $d_{(k|i)}(n) = d_{(k|i)}$ the state process is a Markov chain with stationary transition mechanism. For each strategy $d_{(k|i)}$ the associated transition matrix is defined as:

$$\Pi(d) := [\pi_{(ij|k)}(d)] = \sum_{k=1}^{M} \pi_{(ij|k)} d_{(k|i)}$$

such that on a stationary state distribution for all $d_{(k|i)}$ and $n \geq 0$.