UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Essays on empirical likelihood in economics

Gao, Z.

Publication date
2012

Link to publication

Citation for published version (APA):
Gao, Z. (2012). *Essays on empirical likelihood in economics*. Thela Thesis.

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:05 Dec 2021

1

INTRODUCTION

1.1 EL, MATHEMATICAL PROGRAMMING, AND COMPUTATION

This thesis intends to exploit the roots of Empirical likelihood (EL) and EL's related methods in mathematical programming and computation. The roots will be connected and the connections will induce new solutions for the problems of estimation, computation, and generalization of EL.

In economics, the study of resource allocation under scarcity often refers to optimization or mathematical programming. Identifying optima was firstly proposed by Fermat and Lagrange in their calculus-based formulas, while the way of solving this optimal problem was initiated by Newton and Gauss using their iterative computational methods. The modern optimization which utilizes the dual theory for attaining optima was developed by Kantorovich and was introduced to economics by von Neumann to solve the primal production maximization problem or the dual, cost minimization problem. The existence of the dual of a primal problem requires some regularity conditions which would construct a feasible solution set. In most cases, simplifications of these conditions appear as constraints in optimization problems. Such problems are known as constrained optimization problems. Constrained optimization appears to be *a crucial link* of connecting EL and its related methods to economic and econometric problems. The difference between EL (and its related methods) and many other extremum estimation methods is that EL uses constraints in its estimation.

Prior to EL, in econometrics and statistics, the essential likelihood based topic that relates to mathematical programming and computation is Maximum Likelihood Estimation (MLE). The connection between mathematical programming, MLE and computational algorithms can be traced back to Wald (1943) and Kiefer and Wolfowitz (1952). Recently Owen (1988, 2001) introduced a nonparametric likelihood-based method, Empirical Likelihood, which pushes the implementations of estimation methodology from the classical optimization to the constrained optimization. While EL itself is a statistical problem, applications of EL to economics require the integration of the advanced techniques of mathematical programming and related computational methods.

EL has attracted a lot of attention in econometrics since Qin and Lawless (1994) incorporated estimating equations into EL. The estimating equations in this modified EL play the same role as moment conditions in the Generalized Methods of Moments (GMM) which is currently one of the most popular estimation approaches in econometrics. Moreover, the estimators in both EL (Qin and Lawless, 1994) and GMM share many similar statistical features. Therefore, EL estimation in the framework of Qin and Lawless (1994) has been recognized as a moment-based estimation method in econometrics. For a sample of n observations, the moment-based EL is as follows:

$$\max_{\theta, p_1, \ldots, p_n} \left\{ \prod_{i=1}^{n} n p_i \,\middle|\, \sum_{i=1}^{n} p_i m(X_i, \theta) = 0,\ p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1 \right\}.$$

The likelihood ratio function $\prod_{i=1}^{n} n p_i$ is called, variously, an objective function, cost function, energy function, or energy functional in different applications of mathematical programming. The estimating function $m(X_i, \theta)$ is of main interest in all moment-based estimation methods.

A remarkable connection between the moment-based estimation method and the mathematical programming problem is established in Kitamura and Stutzer (1997) via the parameters in the dual problem. The use of dual parameters also appears in Owen (1990) and Qin and Lawless (1994) for Lagrangian type testing. Kitamura and Stutzer (1997) use the dual parameter to derive an alternative representation (or objective function) of moment-based estimation methods. Their method which they refer to as Exponential Tilting is based on optimizing the Kullback-Leibler divergence, instead of the log-likelihood ratio, subjected to moment constraints. The dual problem shows an alternative way of incorporating moment constraints. The moment constraints no longer appear directly in the objective functions as they do in GMM or other minimum distance methods. The moment constraints are controlled dually by the Lagrangian multiplier instead and then appear indirectly in the modified objective functions.

General speaking, "duality theory" in optimization means the simultaneous study of a *pair* of optimization problems, the

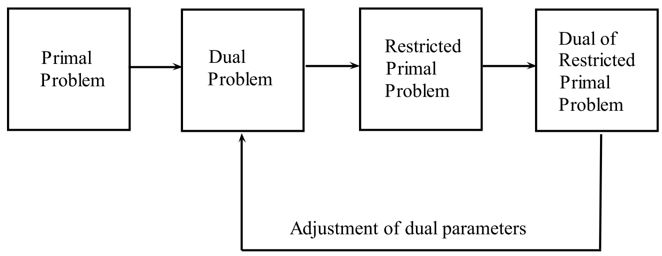

Figure 1: An outline of the primal-dual method.

initial problem which is also called the "primal problem" and the dual problem. The aim of dual problem is to obtain more information about the primal problem. For EL and its related methods, the information of constraints and the information of optimal "weights" w of these constraints are presented in a single objective function by the duality theory. Using the auxiliary dual parameters, Smith (1997) and Newey and Smith (2004) show that a class of estimators including Exponential Tilting, continuous updating GMM and EL, will have better statistical properties than the original GMM whose weights are not necessarily optimal.

Although they are important improvements from the statistical point of view, EL and its related moment-based methods still face several practical difficulties when they are applied in economic models. One of these issues addressed in Chapter 2 and 3 is about non-linear constraints. Non-linearity represents a multitude of complex economic phenomena, and in turn, applications in structural econometrics ubiquitously give rise to problems formulated as nonlinear constrained optimization problems. The duals of these non-linear constrained optimization problems are generally large, complex and infinite dimensional. As a result, this class of optimization problems present significant computational challenges.

As we know, nonlinear equations in general have no closed form solutions but only numerical solutions. Optimization can be thought of as a way of finding approximating solutions to some equations. If the equations are nonlinear, optimization problems also become nonlinear. The computational mechanism, however, is developed from a linear or quasi-linear environment. Thus when one attempts to solve a nonlinear optimization problem, one should first think about transferring the problem to a linear or quasi-linear problem. Then firstly one needs to make a choice for linearizing the problem:

(a): should I linearize the nonlinear problem first and solve a linearized optimization problem?

(b): should I optimize the nonlinear problem first and obtain a set of solution equations to be linearized?

The approach (a) often relates to approximation and the second approach often relates to the optimization given the first order condition. In general, these two steps do not commute.

Chapter 2 suggests a localization method to solve the nonlinear optimization problem of EL, which belongs to category (a). This idea originally appeared in Newton's algorithmic scheme. Le Cam (1974) used Newton's scheme in statistics to solve the irregular likelihood (objective) function caused by the highly degree of nonlinearity. Unlike classic MLE, the moment-based EL estimator depends on its dual optimal value. Le Cam's localization technique might not be proper in the dual problem. As we will see, if the likelihood function in the primal problem of EL has (highly) nonlinear constraints, the dual problem will have a single objective function of a minimax problem with an infinite dimensional Lagrangian multiplier. I apply the primal-dual scheme together with Netwon-Le Cam's localization to solve this problem. From Figure 1 above, one can see the principle role of the dual problem: it helps to restrict the solution set of the primal problem. It implies that a tractable dual representation would be helpful for finding the solution. By this property and the primal likelihood problem, we will have a Kitamura-Stutzer type duality. We will apply this dual result to a restricted problem - a locally linearized (linear-quadratic) representation of the primal problem. Finally the dual of this restricted problem will give a signal how to adjust the multiplier and thereafter the value of the primal likelihood function. Like Newton's algorithm, this is an iterative scheme. Statistical properties of this scheme will depend only on the construction of the estimator rather than the number of iterations in the algorithm.

Chapter 3 considers a rather different problem: how to make an optimal decision when agents face various uncertainties in a dynamic economic system. We will see that this problem is again a mathematical programming problem. The solution of this problem seeks to optimize an objective defined over many points in time taking into account how the objective function changes if there is a small change in the choice path. Usually, economists apply dynamic programming to study this case and the optimization strategy is based on splitting the problem into smaller subproblems. The equation that describes one of these subproblems is called a Bellman equation. We will see that this Bellman type problem can be dually formulated as a mathematical programming with equilibrium constraints where the constraints are Bellman type equations. Then we can easily introduce robustness concerns of uncertainty to the new dual problem even when it is hard to introduce it to the original problem. The solution of the dual problem will be approximately represented by a linear function.

The connection between the contents of these two chapters is the estimation method. The estimation in Chapter 3 relies on the approximating solution of the robust decision problem and it belongs to the category (b). The robust decision problem in Chapter 3 induces non-differentiable components when we construct the constraints for estimation. For non-differentiable components, there is a risk of obtaining inconsistent gradients of the objective functionals. One may approximate these gradients, but in our specific setting (with robustness concerns) the approximate gradient obtained is not a true gradient of anything: neither of the continuous functional nor the discrete functional. To avoid the effects of non-differentiability, we transfer this problem to the one in the category (a). In other words, while the objective functions in the category (b) are continuously differentiable before linearization, we might obtain non-differentiable components after linearizing such problems.

In Chapter 3, one can see that using a usual localization procedure we show that the estimation problem can be trivially discretized first and then the discretized problem will be represented by a continuous local objective function. In a number of subfields in control theory, localization techniques are designed specifically for optimization in dynamic contexts.

The last chapter considers a generalization of EL dealing with the primal-dual concern. Duality, roughly speaking, is a fundamental concept that underlies many aspects of extremum principles in natural systems. In EL and its related methods, the optimal probabilistic weights of the moment constraints are convexifying these constraint functions. This is quite natural in the light of the results of constrained optimization, since best approximation by convex sets is a particular case of convex optimization and for convex systems, the mathematical theory of duality is well established due to the existence of a common symmetric framework (Hahn-Banach Theorem). Therefore, a naive guess is that a primal-dual relation must exist for a general class of EL and its related methods. Chapter 4 is to verify this conjecture and to give a specific class for dually representing GEL.