Instrumental Variable Selection in Generalized Calibration

Bertarelli G, Donelli N, Di Brisco A, Sora V and Valli I.

PhD in statistics - DEMS - Milano Bicocca

Jun. 26th, 2013

Abstract

Calibration is a procedure to incorporates auxiliary information for estimation of finite population parameters. The target of calibration is to obtain a system of weights for constructing estimators of finite population parameters. Calibration weights can be derived by:

- the *minimum distance method* (Deville & Sarndal (1992)),
- the *instrument vector method* (Generalized Estimator) (Devilee, 1998 - Estevao & Sandal, 2006).

The Generalized Estimator depends on the choice of the calibration funztion and of the instrument.
Choosen a linear calibration function, we use typical econometric method for variables selection in order to select the instruments. Caner & Fan (2010) proposed to use the adaptive lasso for variable selection because of its *oracle properties* when instruments are irrelevant.
We show adaptive lasso can select the instruments for which the MSE is smaller.

The Functional Form

In the functional approach we consider weights of the form

$$w_k = d_k F_k(\lambda' \mathbf{z}_k)$$

where $\mathbf{z}_k$ is a vector of instruments with values defined for $k \in s$, sharing the dimension of the specific auxiliary vector $\mathbf{x}_k$, and λ is a vector to be determined from the calibration equation $\sum_s w_x \mathbf{x}_k = \sum_U \mathbf{x}_k$. Apparently, in this context, the geometric interpretation in terms of the minimization of a distance is lost, but if $d_k F_k(\cdot)$ is the inverse mapping of a function $g_k(\cdot, d_k)$ we can write

$$g_k(w_k, d_k) - \lambda' \mathbf{z}_k = 0$$

Under sufficient regularity conditions on $g_k(\cdot, d_k)$, we can find a function $G_k(w_k, d_k)$, differentiable with respect to w_k, which, for every fixed $d_k > 0$, is defined on an interval $D_k(d_k)$ containing d_k so that $g_k(w_k, d_k) = \frac{\partial G_k(w_k, d_k)}{\partial w_k}$ and:

- $G_k(\cdot, d_k)$ is non negative;
- $G_k(\cdot, d_k)$ is strictly convex;
- $G_k(d_k, d_k) = 0$.

Hence, the function $G_k(\cdot, d_k)$ is a distance and the calibration weights obtained through the functional approach can be seen as solutions of a constrained minimum-distance problem, with constraints applied to the instruments $\mathbf{z}$.

Why econometrics instruments approaches?

In econometrics instrumentals variables are used in the context of regression models with stochastic covariates that are correlated. In fact, *IVs* are exogenous variables, correlated with covariates, that are used to "scale" the model in order to achieve a new model that allows consistent (or even asymptotically normal) estimation of its parameters.
The calibration estimators are assisted by a linear regression superpopulation model, hence the IVs $\mathbf{z}_k$ can be seen as IVs for the regression model and so shares all the usual properties of the IVs used in econometrics and all their problems, the main of which is how to choose them.

The Adaptive Lasso

The Lasso is a weighted ℓ_1 penalization method for simultaneous estimation and model selection. It has the oracle properties of asymptotic normality with optimal convergence rate and model selection consistency. Adaptive Lasso is essentially a l_1 penalization method where the degree of penalty is adaptively chosen. Using adaptive lasso we can do more robust inference. In this work we choose adaptive lasso over Lasso (or other econometrics methods for variables selection) because of its advantage in instrumental variables model selection consistency.
Lasso could have bias in estimation and is inconsistent on model selection under certain circumstances (Zou,2006). It's necessary to point out that when the dimesion is high and possible severe collinearity are present, adaptive lasso may fail too (Zou and Zhang, 2009).

Simulation Study

$$y = x_1 + x_2 + \epsilon$$

For possible instruments we consider five trasformations of covariates:

$$x,\ x^2,\ 1 + x,\ 2x,\ e^x$$

From a population of 1000 individuals we extract a sample of 200 individuals with $\pi_i = 0.2$.

Model	x	x^2	$1 + x$	$2x$	e^x	True Value

Conclusions

In our study the adaptive lasso for the selection of the instruments tends to choose always the same instruments for different distributions.
The proposed method always choose the same function as instrumental variables for the same distribution and the choice is independent from the order of the instruments.
The generalized estimator built with instrumental variables selected is closer to the true value.
Anyway, it is worth emphasizing that the oracle properties do not automatically result in optimal prediction performanc.Probably for this reason, sometimes this method fails. It's necessary to study when does this method fails
Moreover a geometric interpretation when $g_k(\cdot, d_k)$ is not invertible have to be find.