Epi-Convergence of M-estimators When Objective Functions are Convex

Kosaku Takanashi
Graduate School of Economics
Keio University

The main purpose of this paper is to present an approach, based on Mosco-convergence, for proving the almost sure convergence of the estimation problem defined by convex minimization. This is done under weaker hypotheses than those usually assumed. Mosco-convergence that our approach in this study is based on is weaker topology than uniform convergence. Mosco-convergence ensures the convergence of empirical minimizer to the exact minimizer. Unlike bracketing condition, maximal entropy and sieve, epi-convergence dose not require compactness assumptions on the parameter spaces.

Let $(\Omega, \mathcal{F}, \mathbb{P})$be probability triple and $\omega \in \Omega$. Let $\Theta \subseteq \mathscr{X}$ be a parameter set in Hilbert space and $\theta \in \Theta$. To show the consistency, we need the convergence of empirical minimizer

$$\arg\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \rho\left(\omega_i, \theta\right) \to \arg\min_{\theta} \mathbb{E}\left[\rho\left(\omega, \theta\right)\right].$$

Consider the uniform convergence

$$\sup_{\theta} \left| \frac{1}{n} \sum_{i=1}^{n} \rho\left(\omega_i, \theta\right) - \mathbb{E}\left[\rho\left(\omega, \theta\right)\right] \right| \xrightarrow{p} 0$$

that implies consistency and model selection optimality. For uniform convergence, we need impose one of either assumptions (van der Vaart and Wellner, 96, Springer); 1:Totally boundedness of parameter space, 2:Random entropy condition, 3:ρ is finite convex.

These assumptions are little bit restrictive for a functional estimation. Instead of uniform convergence, we shall consider a weaker topology than uniform convergence. We focus on the Mosco topology. Mosco convergence ensures the convergence of empirical minimizer.

To obtain Mosco-convergence of the objective function, we use the theorem of the equivalences between Mosco-convergence, Graph convergence(G-convergence) of subdifferential operators, pointwise convergence of Moreau-Yosida approximation and pointwise convergence of resolvent. Using these equivalences, we can establishe the consistency and weak convergence of a estimatior defined on infinite dimensioanl parameter space.

Our main results are as follows. Let $f_1, f_2, \cdots$ and f_0 be a sequence of proper lsc convex function defined on $(\Omega, \mathscr{X})$such that

$$f_n\left(\omega, \theta\right) \triangleq \frac{1}{n} \sum_{i=1}^{n} \rho_i\left(\omega, \theta\right)$$

$$f_0\left(\omega, \theta\right) \triangleq \mathbb{E}\left[\rho\left(\omega, \theta\right)\right].$$

We can obtain the following results of the consistency and achivement of infimum.
There exists a $\mathbb{P} - \textit{negligible}$ subset $\mathcal{N}$ of Ω. For almost all $\omega \in \Omega \setminus \mathcal{N}$, we have

$$\limsup_{i\to\infty} \left(\arg\min_{\theta\in} f_i\left(\omega, \theta\right) \right) \subset \arg\min_{\theta\in\Theta} f_0\left(\omega, \theta\right)$$

where the lim sup is in the weak topology. For all $\omega \in \Omega \setminus \mathcal{N}$ if there is a weakly compact set $K \subset \mathscr{X}$ such that $\arg\min_{\theta} f_i\left(\omega, \theta\right) \subset K$ and $\arg\min f_i \neq \emptyset$ for all i , then

$$\lim_{i\to\infty} \left(\inf f_i\left(\omega, \theta\right)\right) = \inf f_0\left(\omega, \theta\right).$$

This can be interapted as consistency and risk optimality in infinite dimensional parameter spaces. And We establish the existence of minimum defined on Hilbert spaces.

If the set of $\arg\min f$ is sigleton, we can obtain a convergence in distribution of the sequences$\arg\min f_i$ and $\inf f_i$

$$\arg\min_{\theta\in} f_i\left(\omega, \theta\right) \rightsquigarrow \arg\min_{\theta\in\Theta} f_0\left(\omega, \theta\right)$$

$$\lim_{i\to\infty} \left(\inf f_i\left(\omega, \theta\right)\right) \rightsquigarrow \inf f_0\left(\omega, \theta\right).$$