Summary of Variational Convergence and Their Applications to Econometrics and Statistics

Kōsaku Takanashi
Keio University

0.1. Introduction

Given a real valued functional F on a parameter set Θ, one of the main problems of econometrics and statistical inference is to find a minimizer together with the minimum value:

$$\arg\min_{\theta\in\Theta} F(\theta),$$
$$\inf_{\theta\in\Theta} F(\theta).$$

The purpose of thesis is to study the dependence of $\arg\min_{\theta\in\Theta} F(\theta)$ and $\inf_{\theta\in\Theta} F(\theta)$ on the data in particular perturbations of the objective function F. It is clear that, if we consider a sequence $\}F_n\langle$ of perturbations of F, which converges to F in a very strong way, i.e., uniform convergence then, in general, we can prove by elementary arguments that the minimum values of the functionals $\}F_n\langle$ converge to the minimum value of F. If, in addition, Θ is a compact topological space, and if each function $\}F_n\langle$ is lower semicontinuous(lsc) on Θ, then it is easy to see that every sequence of $\arg\min_{\theta\in\Theta} F_n(\theta)$ has a subsequence which converges to a point of $\arg\min_{\theta\in\Theta} F(\theta)$. In particular, if F has a unique minimum point, then the whole sequence $\arg\min_{\theta\in\Theta} F_n(\theta)$ converges to $\arg\min_{\theta\in\Theta} F(\theta)$ in strongly (Newey and McFadden (1994)).

They are not suitable for many applications to Econometrics and nonparametric statistics, characterized by perturbations of minimum problems for integral functionals of the form

$$F(\phi) = \int_S \rangle\phi(x), \mathcal{H}\phi(x)| \, dx,$$

where $\mathcal{H}$ is an elliptical differential operator, S is a subset of $\mathbb{R}^d$, $\rangle\times\nparallel$ denote an inner product on Euclidean space and $\phi : \mathbb{R}^d \infty \mathbb{R}$ is a function satisfying some properties. Suppose that we have a sequence $\}F_n\langle$ of functionals of this form, corresponding to a sequence of functionals $\rangle\phi_n(x), \mathcal{H}_n\phi_n(x)|$. If the usual coerciveness and growth conditions are satisfied uniformly with respect to n, and if for every the sequence $\rangle\phi_n(x), \mathcal{H}_n\phi_n(x)|$ converges to $\rangle\phi(x), \mathcal{H}\phi(x)|$ pointwise a.e. on S, then $\}F_n\langle$ converges to F pointwise, but not uniformly.

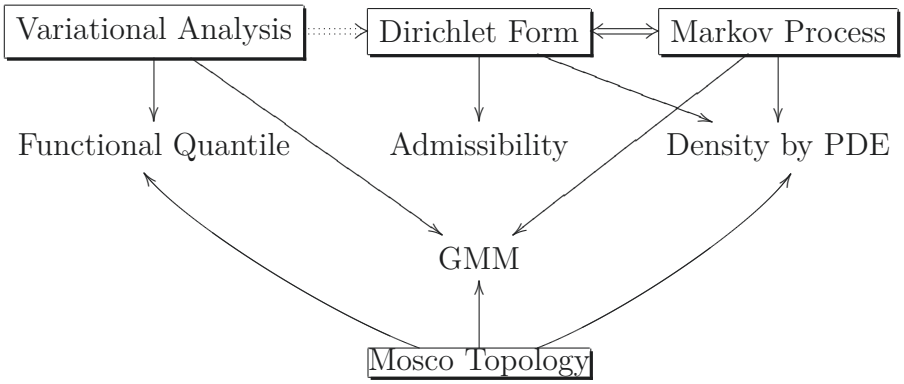
In order to make the objective function satisfy the uniform convergence, we have to impose some compactness of the parameter space or entropy conditions (e.g.,van der Vaart and Wellner (1996) and van der

Vaart (1998)). These assumptions are rather restrictive for the above integral functionals with differential operator, fully nonparametric and non-differentiable convex objective function settings. Although since the objective function is convex, it seems that we may use the convexity lemma (e.g., Pollard (1991) and Theorem 10.8 of Rockafellar (1970)) to ensure that point-wise convergence of convex functions implies uniform convergence, however, in the infinite-dimensional case, this argument for uniform convergence may fail . Let $\pi_n,\ n = 1, 2, \rtimes\rtimes$ be the sequence of projection operators on $\mathscr{H}$ onto $E_n \to \mathscr{H}$ where $E_n \subsetneqq E_{m>n}$. Consider a quadratic form $\rangle \pi_n \theta, \theta|$ for $\mathcal{D}\theta\ \forall\ \mathscr{H}$ that is considered as a convex function of θ. Then, as $n\ \infty \in$, $\rangle \pi_n \theta, \theta|$ converges point-wise to $\rangle \theta, \theta|$ but not uniformly.

However, in this case it is still possible to prove that, for any reasonable choice of the weaker topology, the minimum points and the minimum values of the functionals $\}F_n\langle$ converge to the minimum point and to the minimum value of F. In this thesis, as a reasonable choice of the topology we choose the mosco-convergence, that is the "weakest" notion of convergence for sequences of convex functional which allows to approach the limit on corresponding minimization problems. On this way, various limit problems are analyzed: some, such as a functional linear quantile regression, generalized method of moments estimate of diffusion processes, a kernel density estimate by partial differential equation method, convergence of invariant measure of computed dynamics with unbounded shocks and a relation between admissibility of statistical estimator and recurrence of Markov processes. For all these examples Mosco-convergence provides a flexible tool and a deep insight.

0.2. Examples

In this section we present a number of examples, in which we show how a notion of variational convergence must be sensible. These examples will be dealt with in detail in the next chapters.



0.2.1. Functional Linear Quantile Regression. (Chapter.4). Let $Z = (Y, X)$ be a pair of a scalar response variable Y and a square integrable random function $X = \}X(t)\langle_{t\in[0,1]}$ on a interval $[0,1]$. Let $Q_\tau(Y\|X)$ be the τth conditional quantile function of Y given X for any $\tau \forall (0,1)$ that is away from 0 and 1. The τth conditional quantile $Q_\tau(Y\|X)$ can be written as a linear functional of X:

$$Q_\tau(Y\|X) = \alpha_\tau + \int_0^1 X(t)\,\beta_\tau(t)\,dt, \quad \tau \forall (0,1),$$

where $\bar{X}(t) = X(t) \quad \mathbb{E}[X(t)]$, α_τ is a scalar constant and $\beta_\tau(t)$ is a scalar function in $L^2[0,1]$. Hereafter, we consider estimating the slope function β_τ. The unknown parameter $\theta_0 = (\alpha, \beta)$ belongs to $\{ \ = \mathbb{R} * L^2[0,1]$. We suppose $X(t)$ satisfies $\mathbb{E}\left[\int_0^1 \|X(t)\|^2 dt\right] < \in$. We take the slope function $\beta_\tau(t)$ to be an RKHS, $\{$, a subspace of the Hilbert space of square integrable functions $L^2[0,1]$. We denote the inner product and the associated norm in $\{$ by $\rangle \times \nparallel$ and $\times$ respectively. Suppose we observe data $(Y_i, X_i(t))$, $1 \geq i \geq n$ consisting of n independent copies of $(Y, X(t))$. With them, we may estimate α_τ, β_τ via by penalization method :

(0.2.1)

$$\left(\hat{\alpha}_{\tau,n,\lambda}, \hat{\beta}_{\tau,n,\lambda}\right) = \arg\min_{\alpha\in\mathbb{R},\beta\in\mathcal{H}} F_{\tau,n,\lambda}(\theta)$$

$$\triangleq \arg\min_{\alpha\in\mathbb{R},\beta\in\mathscr{H}} \frac{1}{n}\sum_{i=1}^n \rho_\tau\left(Y_i \quad \alpha \quad \int_0^1 X_i(t)\,\beta(t)\,dt\right) + \lambda_n J(\beta),$$

where $\rho_\tau(u) = \left\{\tau \quad \mathbf{1}_{(u\leq 0)}\right\} u$ is the check function (Koenker and Bassett Jr (1978)), λ_n is the smoothing parameter that converges to zero as $n \ \infty \in$ and $J(\beta)$ is a convex penalty functional on β. Obviously, the criterion function $\rho_\tau(\times)$ is not continuously differentiable. When $\rho = \|\times\|$ and (α, β) is in an infinite-dimensional space, we have

$$\sup_{\theta\in\mathbb{R}\times\mathcal{H}} \|F_{\tau,n,\lambda}(\theta) \quad \mathbb{E}\left[F_{\tau,n,\lambda}(\theta)\right]\| \subset \gamma$$

for some constant γ. The left-hand side of the inequality does not converge uniformly. The convexity lemma argument to ensure that point-wise convergence of convex functions implies uniform convergence may fail. Let $\pi_n,\ n = 1, 2, \times\times\times$ be the sequence of projection operators on $\mathscr{H}$ onto $E_n \rightarrow \mathscr{H}$ where $E_n \subsetneqq E_{m>n}$. Consider a quadratic form $\rangle\pi_n\theta, \theta|$ for $\mathcal{D}\theta \forall \mathscr{H}$ that is considered as a convex function of θ. Then, as $n \ \infty \in$, $\rangle\pi_n\theta, \theta|$ converges point-wise to $\rangle\theta, \theta|$ but not uniformly.

0.2.2. Conditional Moment Estimate of Markov Process. (Chapter.5). In many economic and finance applications it is common to start with a stochastic differential equation. The dynamics are usually described by an Itô-type stochastic differential equation in the time-homogeneous case reading

$$dx_t = \mu(x_t)\, dt + \Sigma^{1/2}(x_t)\, dW_t$$

where $\mu(\times)$ is a local mean (drift) and $\Sigma(\times)$ is a local variance (diffusion coefficient, or volatility) on the state space $S \rightarrow \mathbb{R}^d$. We study the problem of estimating the coefficients of a diffusion $\}X_t\langle_{t\geq 0}$; the estimation is based on discrete data $\}X_{n\Delta}; n = 1, 2, \times\times\times N\langle$. The sampling frequency is constant, and asymptotics are taken as the number N of observations tends to infinity. We prove that the weak asymptotics of estimating both the diffusion coefficient (the volatility) and the drift in a nonparametric setting.

There are well known connections between the coefficients of the stochastic differential equation and the infinitesimal generator. Hansen and Scheinkman (1995) studied how to generate moment conditions for continuous-time Markov processes with discretely sampled data by using the infinitesimal generator. Hansen and Scheinkman (1995) derived moment conditions for estimating the parameters of continuous-time Markov processes using discrete time data. The central thrust of their

approach can be illustrated through a simple example with one dimensional $x_t \ \forall \ \mathbb{R}$ and constant coefficients $\mu, \sigma \ \forall \ \mathbb{R}$. Applying the infinitesimal generator to any well behaved transformation $f(\times)$ of X_t and by Ito's Lemma we have:

$$\mathbb{E}\left[\mathcal{H} \times f(x)\right] = \mathbb{E}\left[\quad f'(x)\,\mu\,(x \quad \mu) + \frac{\sigma^2}{2} f''(x)\right] = 0$$

which yields an infinite number of moment conditions, one for each f, related to the marginal distribution of $f(x)$. From these moment conditions, and under the regularity conditions specified by Hansen and Scheinkman (1995) GMM estimation may be performed in the usual way.

These above theorems imply that the drift and diffusion coefficients of stationary scalar diffusions can be identified up to a common scale factor. Additional a conditional expectation operator allows one to identify fully the generator of reversible processes likewise

$$\frac{\lambda_1}{\hat{\lambda}_1}\hat{\mathcal{H}}$$

where λ_1 is the second largest eigenvalue of a conditional expectation and $\hat{\lambda}_1$ is the second largest eigenvalue of the estimated infinitesimal generator $\hat{\mathcal{H}}$. To estimate the second largest eigenvalue, we must solve the stochastic optimization problem of the unknown maximum eigenvalue $\hat{\lambda}$:

$$\hat{\lambda} = \max_{\int \phi dQ = 0, \|\phi\| = 1} \left\langle \phi, \hat{\mathcal{H}}\phi \right\rangle$$

which is the stochastic optimization problem including a differential operator. Therefore, the left-hand side does not converge uniformly.

0.2.3. Adaptive Density Estimate by Partial Differential Equation Method. (Chapter.6). Consider a Gaussian kernel estimate based on sample data $\}x_1, x_2, \times\times\times, x_N\langle$ with $x_i \ \forall \ \mathbb{R}$:

$$\hat{f}(x;t) = \frac{1}{N}\sum_{i=1}^{N} \phi(x, X_i; t), \ x \ \forall \ \mathbb{R},$$

where ϕ is the Gaussian kernel function:

$$\phi(x, X_i; t) = \nabla\frac{1}{\sqrt{2\pi t}} \exp\left(\quad \frac{1}{2t}(x \quad X_i)^2\right)$$

and $\sqrt{t}$ is its bandwidth. Observe that the Gaussian kernel density estimator is the unique solution to the heat equation(diffusion partial differential equation (PDE))

$$\frac{\partial}{\partial t}\hat{f}\left(x;t\right)=\frac{1}{2}\frac{\partial^2}{\partial x^2}\hat{f}\left(x;t\right),\ x\ \forall\ \mathscr{X},t>0,$$
$$\hat{f}\left(x;0\right)=\frac{1}{N}\sum_{i=1}^{N}\delta\left(x-X_i\right)$$

where $\delta\left(x-X_i\right)$ is the Dirac measure at X_i. Considering the heat equation provides an interpretation that the Gaussian kernel is the so-called Green' s function for the diffusion PDE. If we treat the bandwidth as time, the smoothness of the estimated density will increase as the time increases. The value of t has the same effect as h in the kernel estimate, and therefore, the time parameter in heat diffusion resembles the bandwidth in kernel density estimation. Thus, the Gaussian kernel density estimator can be obtained by solving the parabolic PDE up to time t.

The advantage of the PDE formulation is the case where the domain of the data is to be bounded. The most traditional kernel estimators suffer from boundary value problem. Suppose the data is known to be in a finite domain $[0,1]$. Impose the Neumann boundary condition

$$\left.\frac{\partial}{\partial x}\hat{f}\left(x;t\right)\right|_{x=1}=\left.\frac{\partial}{\partial x}\hat{f}\left(x;t\right)\right|_{x=0}=0.$$

on PDE. The PDE with the initial condition and Neumann boundary condition has an analytical solution.

For multi-dimensional case, there is, in general, no analytical expression for the diffusion kernel. Computational approximation is needed. Let A be an elliptic differential operator which is a non-negative definite self-adjoint operator on $\mathscr{H}$. There is a one to one correspondence between the family of closed symmetric forms $(\mathcal{F},\mathcal{E}\ (\mathcal{F}))$ and the family of non-negative definite self-adjoint operators A on $\mathscr{H}$. The correspondence is determined by

$$\mathcal{F}(f,g)=\left\langle\sqrt{A}f,\sqrt{A}g\right\rangle$$
$$\mathcal{G}\ =\mathscr{D}\left(\sqrt{A}\right)$$

where$\rangle, |$ is a inner product in $L^2(S;m)$:

$$\rangle f, g| = \int_E f(x)\, g(x)\, m(dx).$$

This$(\mathcal{F}, \mathcal{G})$ is called Dirichlet form of m-symmetric processes. Define the resolvent $G_\alpha f$ such that

$$\mathcal{F}_\alpha(G_\alpha f, v) = \rangle f, v| \,, \quad {}^\forall f \,\forall\, \mathscr{H},\ {}^\forall v \,\forall\, \mathcal{E}(\mathcal{F})$$
$$\mathcal{F}(G_\alpha f, v) + \alpha \rangle G_\alpha f, v| = \rangle f, v| \,.$$

The above defined $\}G_\alpha\langle$ is a strongly continuous resolvent generated by $(\mathcal{F}, \mathcal{E}(\mathcal{F}))$ and define a Laplace transformation of $\kappa(x, y, t)$:

$$G_\alpha(x, y, \lambda) = \int_0^\infty e^{-\alpha s} \kappa(x, y, t)\, ds.$$

From this resolvent $\}G_\alpha\langle$, by the inverse Laplace transform:

$$\kappa(x, y, t) \times f = \lim_{\beta\to\infty} e^{-t\beta} \sum_{n=0}^{\infty} \frac{(t\beta)^n}{n!} (\beta G_\beta)^n f, \qquad f \,\forall\, \mathbb{H}.$$
$$= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{\beta t} G_\beta f d\beta.$$

An approximation solution of the resolvent $G_\alpha^m f$ is defined by the variational form of

$$G_\alpha^m f = \arg\min_{v\in\mathcal{H}_m} \left\{ \mathcal{F}(v, v) + \frac{1}{2\alpha} \rangle v \quad f, v \quad f| \right\}.$$

Since the above optimization problem contains an elliptic differential operator, uniform convergence of this objective function may fails.

0.2.4. Convergence of Computed Dynamic Markov Model. (Chapter.7). Most dynamic economic models do not have a closed-form solution. Model's policy functions are approximated by numerical methods. Therefore, the researcher can only evaluate an approximated invariant measure associated with the approximated transition function rather than the exact invariant measure implied by the exact transition function.

The equilibrium law of motion of the state variables can be specified by a dynamical system of the form

$$s_{t+1} = \varphi(s_t, \varepsilon_{t+1}), \qquad t = 0, 1, 2, \ldots$$

Here, s_t is a vector of state variables that characterize the evolution of the system. The variable ε is an independent and identically distributed shock. In most case, researcher dose not know the exact form of transition equations φ. He only access to numerical approximation to the transition equations φ_j with index j. The index j indicate the approximation and imply that as j goes to infinity the approximation φ_j converge to their exact values (the metric of convergence is defined later).

Note that each φ_n defines the associated pair (P_j, T_j) : Markov operator associated with φ_j is defined as

$$\begin{aligned} T_j f(s) &\triangleq \int f(t) P_j(s, dt) \\ &= \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon) \end{aligned}$$

The adjoint T_j^* of T_j is as

$$\begin{aligned} \rangle T_j f, \mu_j| &= \int \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon) d\mu_j(s) \\ \left\langle f, T_j^* \mu_j \right\rangle &= \int \int f(\varphi_j(s, \varepsilon)) Q(s, d\varepsilon) d\mu_j(s) \end{aligned}$$

Moreover there always exists an invariant distribution $\mu_j^* = T_j^* \mu_j^*$. The purpose is the convergence of invariant measure obtained from numerical simulations to the exact invariant measure.

Santos and Peralta-Alva (2005) have studied the convergence of computed invariant measure of economic models which cannot be solved analytically and must be solved numerically or with some other form of approximation. Fernandez-Villaverde et al. (2006) have studied the convergence of the likelihood of computed economic models. However, they assume that the state space is compact and therefore, the support of the shock of dynamical system is assumed to be bounded. Although this assumption is standard in the numerical literature, but this assumption excludes from the dynamical model the normal distribution. To relax the compactness assumption for the convergence of the approximated invariant measure, we must relax the topology of uniform convergence.

0.2.5. Admissibility. (Chapter.8). Consider the problem of estimating the mean of a multivariate normal distribution. Consider squared error as the loss function. We are interested in determining necessary and sufficient conditions for an estimator, δ, to be admissible. Stein (1956) proved that the mean, $\delta(x) = x$ (the best invariant estimator), is admissible if the dimension m of the multivariate normal distribution satisfies $m \geq 2$ and is inadmissible if $m \subset 3$. There is another interesting division between dimensions $m = 2$ and $m = 3$. Brownian motion is recurrent in dimensions $m = 1, 2$ and is transient if $m \subset 3$. The mathematical link between the statistical decision problem and the stochastic process problems is a simple calculus of variational problem, i.e., Dirichlet form. This Dirichlet form involves the infinitesimal generator of the above mentioned Brownian motion, more generally a symmetric Markov processes. And it is known that the Markov process is recurrent if and only if the corresponding Dirichlet form has 0 infimum.

At the same time, subject to the regularity conditions mentioned above, we are able to exploit the mathematical link to the statistical problem to show that the statistical estimator is admissible also if and only if this exterior Dirichlet problem is insoluble.

Let X be an m-dimensional normally distributed random variable with unknown mean $\theta = (\theta_1, \times\times\times, \theta_m)'$ to be estimated and a variance-covariance matrix as the identity matrix I . We denote the estimator of $\theta = (\theta_1, \times\times\times, \theta_m)'$ as $\delta = (\delta_1, \times\times\times\delta_m)'$. Although a natural estimator is taking mean:$\delta = \overline{x}$, mean is inadmissible when $m \subset 3$. Define the loss function

$$L(\theta, \delta) = (\delta \quad \theta)' D (\delta \quad \theta)$$

where D is a fixed known $m*\ m$ diagonal matrix with elements $(d_1, d_2, \times\times\times, d_m)$ with $d_1 \geqq d_2 \geqq \times\times\times \geqq d_m > 0$. The measure for the goodness of estimator δ is the risk R defined by

$$R(\theta, \delta) = \mathbb{E}[L(\theta, \delta)].$$

It is said that δ is admissible if there is no δ_* such that for any unknown θ, $R(\theta, \delta_*) \geq R(\theta, \delta)$ and for some θ, $R(\theta, \delta_*) < R(\theta, \delta)$. Let G be any nonnegative Borel measure on $\mathbb{R}^m$ and be a finite measure. Define the

Bayes risk $B(G,\delta)$ by

$$B(G,\delta) = \int R(\theta,\delta)\, G(d\theta).$$

Whether or not G is finite measure, the generalized Bayes estimator(Pitmann estimator) is given by

$$\delta_G(x) = \frac{\int \theta p_\theta(x)\, G(d\theta)}{\int p_\theta(x)\, G(d\theta)}$$

where $p_\theta(x)$ is the normal density with mean θ:

$$p_\theta(x) = \frac{1}{\sqrt{(2\pi)^m}} \exp\left(\quad \frac{1}{2}\sum_{i=1}^{m} (x_i \quad \theta_i)^2 \right).$$

For notational convenience, let $\gamma_G(x)$ be

$$\gamma_G(x) = \delta_G(x) \quad x.$$

Denote the convolution

$$g^*(x) = \int p_\theta(x)\, G(d\theta)$$

by $g^* = p \bullet G$. g^* is the marginal likelihood. Let F be a given generalized prior including improper prior, and define $f^* = p \bullet F$. $\gamma_G(x)$ can be written by

$$\gamma_G(x) = \frac{g^*(x)}{g^*(x)}.$$

The fundamental tool for the necessary and sufficient condition for admissibility due to Stein et al. (1955). According to this, δ_F is admissible if and only if there is a sequence of nonnegative finite Borel measures, G_i, $i = 1, 2, \ldots$ satisfying

$G_i(\}0\langle) = 1.$
Each G_i has compact support and satisfy

$$B(G_i, \delta_F) \quad B(G_i, \delta_{G_i}) \infty 0.$$

Hence, we focus on $B(G_i, \delta_F) \quad B(G_i, \delta_{G_i})$. The important thing is that some algebra provides the Dirichlet form in the following way; Write

$$B\left(G_i,\delta_F\right) \quad B\left(G_i,\delta_{G_i}\right)=\int\left[R\left(\theta,\delta_F\right) \quad R\left(\theta,\delta_G\right)\right]g\left(\theta\right)d\theta$$

where $R\left(\theta,\delta_G\right)$ can be expressed by

$$R\left(\theta,\delta_G\right)=\int\sqrt{\delta_G} \quad \theta\sqrt{}^{2}p_\theta\left(x\right)dx$$

Next, we have

$$\begin{aligned}
B\left(G_i,\delta_F\right) \quad B\left(G_i,\delta_{G_i}\right) &= \int\sqrt{\delta_F\left(x\right)} \quad \delta_{G_i}\left(x\right)\sqrt{}^{2}g_i^*\left(x\right)dx \\
&= \int\sqrt{} \quad \log f^*\left(x\right) \qquad \log g^*\left(x\right)\sqrt{}^{2}g_i^*\left(x\right)dx \\
&= \int\left\| \quad \log\frac{f^*\left(x\right)}{g^*\left(x\right)}\right\|^2 g_i^*\left(x\right)dx \\
&= \int\left\|\frac{f^*\left(x\right)g_i^*\left(x\right) \quad f^*\left(x\right) \quad g_i^*\left(x\right)}{f^*\left(x\right)}\right\|^2\frac{1}{g_i^*\left(x\right)}dx \\
&= \int\left\|\frac{f^*\left(x\right) \quad g_i^*\left(x\right) \quad g_i^*\left(x\right) \quad f^*\left(x\right)}{\}f^*\left(x\right)\langle^2}\right\|^2\frac{\}f^*\left(x\right)\langle^2}{g_i^*\left(x\right)}dx
\end{aligned}$$

and letting $\hat{h}_i\left(x\right)=\frac{g_i^*(x)}{f^*(x)}$, we have

$$B\left(G_i,\delta_F\right) \quad B\left(G_i,\delta_{G_i}\right)=\int\frac{\left\| \quad \hat{h}\left(x\right)\right\|^2}{\hat{h}_i\left(x\right)}f^*\left(x\right)dx$$

and letting $\hat{j}_i\left(x\right)=2\sqrt{\hat{h}_i\left(x\right)}$ provides

$$\hat{j}_i\left(x\right)=\frac{\hat{h}_i\left(x\right)}{\sqrt{\hat{h}_i\left(x\right)}},$$

therefore,

$$B\left(G_i,\delta_F\right) \quad B\left(G_i,\delta_{G_i}\right)=\int\left\| \quad \hat{j}_i\left(x\right)\right\|^2 f^*\left(x\right)dx.$$

The above equation is the fundamental equation of this study: the close connection between the statistical problem of admissibility and recurrence of diffusion processes on $\mathbb{R}^m$. Theses above variational problems comprise the differential operator. One must be assiduous in the topological matters. Here is the main mathematical link between the risk boundedness below and recurrence of transience of Markov processes.

$\leq$ Let $\}T_t\langle_{t>0}$ be a strongly continuous Markovian semigroup on $L^2(S;m)$ and $\int \left\| \quad \hat{j}_i(x) \right\|^2 f^*(x)\,dx$ be the associated Dirichlet space relative to $L^2(S;m)$. Then $\}T_t\langle_{t>0}$ is transient if and only if there exists a bounded m-integrable strictly positive function g such that $g > 0,\ m \quad a.e.$ on S satisfying

$$\int_E \|\jmath_i\| g dm \leqq \int \left\| \quad \hat{j}_i(x) \right\|^2 f^*(x)\,dx, {}^{\forall} j_i \ \forall\ \mathcal{G}\,.$$

$\leq$ For each Dirichlet form $\int \left\| \quad \hat{j}_i(x) \right\|^2 f^*(x)\,dx$ on $L^2(S;m)$, the following is equivalent:

(i)$\}T_t\langle_{t>0}$ is recurrent.

(ii)There exists a sequence $\}j_n\langle$ satisfying

$$\}j_n\langle \to \mathcal{G}\,, \quad \lim_{n\to\infty} j_n = 1\,(m\text{-a.e.})\,, \quad \lim_{n\to\infty} \int \left\| \quad \hat{j}_i(x) \right\|^2 f^*(x)\,dx = 0.$$

(iii)In the case where $m(E) < \in$,

$$\int \sqrt{\ 1\ }^{2} \sqrt{f^*(x)}\,dx = 0.$$

0.3. Plan of this thesis

The rest of this paper is organized as follows. In Chapter 2, we present the general set-up and main results of a fully abstract model. We describe the Mosco convergence and introduce the narrow convergence in the Mosco topology. We derive the quadratic approximation of a convex objective function in an infinite-dimensional Hilbert space. We also provide the asymptotic distribution of the optimal value.

In Chapter 3, we present the general notion of Dirichlet form and their relationship with the symmetric Markov process. There is one-to-one correspondence between the Dirichlet form and semi-group of the symmetric Markov process. We apply the Mosco convergence to the perturbed Dirichlet form and describe the Mosco convergence of the Dirichlet form. We derive the Mosco convergence of the Dirichlet form implies the narrow convergence of the corresponding symmetric Markov process and vise versa.

In Chapter 4, we study an asymptotics of functional linear quantile regression in which the dependent variable is scalar while the covariate is a function. We apply a roughness regularization approach of a reproducing kernel Hilbert space framework. In the above circumstance,

narrow convergence with respect to uniform convergence fails to hold, because of the strength of its topology. A new approach we propose to the lack-of-uniform-convergence is based on Mosco-convergence that is weaker topology than uniform convergence. By applying narrow convergence with respect to Mosco topology, we develop an infinite-dimensional version of the convexity argument and provide a proof of an asymptotic normality of argmin processes. Our new technique also provides the asymptotic confidence intervals and the generalized likelihood ratio hypothesis testing in fully nonparametric circumstance.

In Chapter 5, we adopt the approach of Hansen and Scheinkman (1995), and provide an asymptotics of this approach. We begin by considering a Markov process specified in terms of its infinitesimal generator. Formally, this generator is defined as an operator on a function space, and, in effect, this operator stipulates the local evolution of the process. For the fully identification, one must estimate the second largest eigenvalue of the infinitesimal generator $\mathcal{H}$ which involves an optimization problem including differential operator. That is beyond an usual asymptotics of empirical process theory. We deal with this problem by the introduced Mosco topology.

In Chapter 6, We extend Botev et al. (2010)'s adaptive kernel density estimation method based on the smoothing properties of linear diffusion processes in two ways. First, we extend their proposed diffusion kernel method to kernel density estimators based on Lévy processes, which have the diffusion estimator as a special case. The kernels constructed via a Lévy process could be tailored for data for which smoothing with the diffusion estimator is not optimal. Second, we consider an asymptotics of the estimated diffusion differential operator that has a random fluctuate due to the estimated pilot density. This problem induces a variational problem, and in fact can be addressed by a straightforward application of Mosco convergence of Dirichlet form.

In Chapter 7, we provide the conditions for the convergence of invariant measure obtained from numerical simulations to the exact invariant measure. Most dynamic economic models do not have a closed-form solution. Model's policy functions are approximated by numerical methods. Therefore, the researcher can only evaluate an approximated invariant measure associated with the approximated transition function

rather than the exact invariant measure implied by the exact transition function. However, previous study assumed that the state space is compact and therefore, the support of the shock of dynamical system is assumed to be bounded. We relax the compactness assumption for the convergence of the approximated invariant measure.

In Chapter 8, we generalize and reformulate Brown (1971) idea of the admissibility question for more general distributions, for more general Bayesian decisions and for more general variational form, i.e., Dirichlet form. This connection goes far beyond the diffusion processes case that Brown (1971) consider. The relation between admissibility of a general Bayesian decision which is based on general distributions and recurrence of the other symmetric Markov processes is established. Since general distributions include Lévy type(infinitely divisible) distributions as a special case of a much more general phenomenon, we give a striking result on a maximum likelihood estimate(MLE) of Cauchy distribution that MLE of Cauchy distribution with dimension $d = 1$ is admissible but is inadmissible with $d \subset 2$. This phenomenon is compatible with the transiency of Cauchy processes with the division between dimensions $m = 1$ and $m = 2$.

Bibliography

Botev, Z. I., J. F. Grotowski, D. P. Kroese, et al. (2010): "Kernel density estimation via diffusion," *The annals of Statistics*, 38, 2916–2957.

Brown, L. D. (1971): "Admissible estimators, recurrent diffusions, and insoluble boundary value problems," *The Annals of Mathematical Statistics*, 42, 855–903.

Fernandez-Villaverde, J., J. Rubio-Ramirez, and M. Santos (2006): "Convergence Properties of the Likelihood of Computed Dynamic Models," *Econometrica*, 74, 93–119.

Hansen, L. P. and J. A. Scheinkman (1995): "Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes," *Econometrica: Journal of the Econometric Society*, 767–804.

Koenker, R. and G. Bassett Jr (1978): "Regression quantiles," *Econometrica: journal of the Econometric Society*, 33–50.

Newey, W. K. and D. McFadden (1994): "Large sample estimation and hypothesis testing," *Handbook of econometrics*, 4, 2111–2245.

Pollard, D. (1991): "Asymptotics for Least Absolute Deviation Regression Estimators," *Econometric Theory*, 7, 186–199.

Rockafellar, R. T. (1970): *Convex analysis*, Princeton university press.

Santos, M. and A. Peralta-Alva (2005): "Accuracy of simulations for stochastic dynamic models," *Econometrica*, 73, 1939–1976.

Stein, C. (1956): "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution," Tech. rep., STANFORD UNIVERSITY STANFORD United States.

Stein, C. et al. (1955): "A necessary and sufficient condition for admissibility," *The Annals of Mathematical Statistics*, 26, 518–522.

van der Vaart, A. (1998): *Asymptotic statistics*, Cambridge University Press.

van der Vaart, A. and J. Wellner (1996): *Weak convergence and empirical processes*, Springer.