Краткий конспект лекций по курсу «Игры среднего поля»

Лекция 12

Оценки решений уравнений ФПК

На прошлой лекции обсуждалась задача Коши

$$\partial_t\mu_t = \frac{1}{2}\Delta\mu_t - \mathrm{div}(b\mu_t), \quad \mu_0 = \nu,$$

где ν — вероятностная мера на $\mathbb{R}^d$ и b — борелевское векторное поле, ограниченное на $[0,T]\times B(0,R)$ для всякого $R>0$. Было показано, что существует субвероятностное решение μ_t, причем мера $\mu_t\,dt$ имеет непрерывную положительную плотность ϱ относительно меры Лебега. Более того, если существует функция Ляпунова V, то субвероятностное решение является вероятностным и кривая $t\to\mu_t$ является непрерывной кривой в $\mathcal{P}(\mathbb{R}^d)$. Типичный пример:

$$|x|^2\in L^1(\nu), \quad \langle b(x,t),x\rangle \le C + C|x|^2.$$

В случае существования функции Ляпунова вероятностное решение оказывается единственным. Однако проблема единственности является частным случаем более общей проблемы зависимости решений от коэффициентов.

Положим

$$L_1u = \frac{1}{2}\Delta u + \langle b_1, \nabla u\rangle, \quad L_2u = \frac{1}{2}\Delta u + \langle b_2, \nabla u\rangle.$$

Пусть $\mu_t^1(dx) = \varrho^1(x,t)\,dx$ и $\mu_t^2(dx) = \varrho^2(x,t)\,dx$ — решения уравнений

$$\partial_t\mu_t^1 = L_1^*\mu_t^1, \quad \partial_t\mu_t^2 = L_2^*\mu_t^2,$$

и $\mu_0^1=\mu_0^2=\nu$. Напомним, что при $t>0$ плотности ϱ^1 и ϱ^2 непрерывны и положительны. Пусть

$$v(x,t) = \frac{\varrho^2(x,t)}{\varrho^1(x,t)}.$$

Теорема 1. *Предположим, что $|b_1-b_2|\in L^2(\mu_t^2\,dt)$ и выполняется хотя бы одно из условий*

(i) $(1+|x|)^{-1}|b_1(x,t)|\in L^1(\mu_t^1\,dt)$

(ii) $\langle b_1(x,t),x\rangle \le C + C|x|^2$.

Тогда

$$\int v(x,t)\ln v(x,t)\mu_t^1(dx) \le \frac{1}{2}\int_0^t\int |b_1(x,s)-b_2(x,s)|^2\,d\mu_s^2\,ds.$$

Обсудим лишь идею доказательства в предположении гладкости ϱ^1, ϱ^2, b_1, b_2.

Пусть $f\in C^2(0,+\infty)$. Можно проверить, что верны следующие равенства:

$$L_1^*(uv) = vL_1^*u + uL_1^*v + \langle\nabla u,\nabla v\rangle + uv\mathrm{div}b_1,$$

$$L_1^*f(u) = f'(u)L_1^*u + \frac{1}{2}f''(u)|\nabla u|^2 + (uf'(u)-f(u))\mathrm{div}b_1.$$

Заметим, что

$$\partial_t\varrho^2 = L_2^*\varrho^2 \quad\Leftrightarrow\quad \partial_t(v\varrho^1) = L_1^*(v\varrho^1) - \mathrm{div}\big((b_2-b_1)v\varrho^1\big).$$

Используя эти наблюдения и уравнение $\partial_t\varrho^1 = L^*\varrho^1$ выводим равенство

$$\partial_t(f(v)\varrho^1) = L_1^*(f(v)\varrho^1) - \frac{1}{2}f''(v)|\nabla v|^2\varrho^1 - f'(v)\mathrm{div}\big((b_2-b_1)v\varrho^1\big).$$

Умножая это равенство на функцию $\psi\in C_0^\infty(\mathbb{R}^d)$ и интегрируя, получаем

$$\int_{\mathbb{R}^d} f(v(x,t))\psi(x)\varrho^1(x,t)\,dx + \int_0^t\int_{\mathbb{R}^d}\frac{1}{2}\psi f''(v)|\nabla v|^2\varrho^1\,dx\,ds =$$
$$= f(1)\int_{\mathbb{R}^d}\psi\,d\nu + \int_0^t\int_{\mathbb{R}^d}\big[f(v)L_1\psi\big]\varrho^1\,dx\,ds +$$
$$+\int_0^t\int_{\mathbb{R}^d}\big[\langle b_2-b_1,\nabla v\rangle f''(v)\psi v\varrho^1 + f'(v)\langle b_2-b_1,\nabla\psi\rangle v\varrho^1\big]\,dx\,ds.$$

Предположим, что $\psi\ge 0$ и $f''\ge 0$. С помощью неравенства Коши-Буняковского приходим к оценке

$$\int_{\mathbb{R}^d} f(v(x,t))\psi(x)\varrho^1(x,t)\,dx \le f(1)\int_{\mathbb{R}^d}\psi\,d\nu +$$
$$+\frac{1}{2}\int_0^t\int_{\mathbb{R}^d}|b_2-b_1|^2 f''(v)v^2\varrho^2\,dx\,ds +$$
$$+\int_0^t\int_{\mathbb{R}^d}\big[f(v)L_1\psi\big]\varrho^1\,dx\,ds +$$
$$+\int_0^t\int_{\mathbb{R}^d}\big[f'(v)\langle b_2-b_1,\nabla\psi\rangle v\varrho^1\big]\,dx\,ds.$$

Пусть последовательность функций ψ_N удовлетворяет условиям:

$$L_1\psi_N\to 0,\quad |\nabla\psi_N|\to 0,\quad \psi_N\to 1.$$

Заменяя в полученном выше неравенстве ψ на ψ_N и устремляя $N\to\infty$ приходим к неравенству

$$\int_{\mathbb{R}^d} f(v(x,t))\varrho^1(x,t)\,dx \le f(1) + \frac{1}{2}\int_0^t\int_{\mathbb{R}^d}|b_2-b_1|^2 f''(v)v^2\varrho^1\,dx\,ds.$$

При $f(v)=v\ln v$ получаем требуемую оценку.

Оценка энтропии уже является оценкой расстояния между вероятностными плотностями, но удобнее работать с L^1–расстоянием или полной вариацией.

Неравенство Пинскера–Кульбака–Чезара

Пусть $(X,\mathcal{A})$ — измеримое пространство, μ и ν — вероятностные меры, причем $\nu=f\mu$ и $f>0$.

Предложение 1. *Имеет место неравенство*

$$\|\mu-\nu\|_{TV}^2 = \Big(\int|f-1|\,d\mu\Big)^2 \le 2\int f\ln f\,d\mu.$$

Доказательство. Положим $E=\{x\colon f(x)\le 1\}$, $\mu(E)=t$ и $\nu(E)=a$. Заметим, что $0<a\le t$ и равенство $t=1$ влечет равенство $f(x)=1$ для μ почти всех x. Пусть $0<a<t<1$. Имеем

$$\int|f-1|\,d\mu = \int_E(1-f)\,d\mu + \int_{X\setminus E}(f-1)\,d\mu = 2(t-a).$$

Применим информационное неравенство

$$\int f\ln f\,d\mu \ge \int f\ln g\,d\mu$$

с вероятностной плотностью $g(x)=\frac{a}{t}$ при $x\in E$ и $g(x)=\frac{1-a}{1-t}$ при $x\notin E$. Получаем

$$\int f\ln f\,d\mu \ge a\ln\tfrac{a}{t} + (1-a)\ln\tfrac{1-a}{1-t}.$$

Остается заметить, что правая часть оценивается снизу выражением $2(t-a)^2$. □

Следствие 1. *В условиях теоремы верна оценка*

$$\|\varrho^1(\,\cdot\,,t)-\varrho^2(\,\cdot\,,t)\|_{L^1(\mathbb{R}^d)}^2 \le \int_0^t\int|b_1(x,s)-b_2(x,s)|^2\varrho^2(x,s)\,dx\,ds.$$

Эта оценка имеет много разнообразных приложений: 1) единственность вероятностного решения, 2) разрешимость нелинейных уравнений Фоккера–Планка–Колмогорова, когда коэффициент b имеет вид $b(x,\mu_t)$, 3) непрерывная и дифференцируемая зависимость решений от параметра, 4) сходимость к стационарному решению, 5) разрешимость системы уравнений теории игр среднего поля.

Стохастическое оптимальное управление

Пусть $b_t(x,u)$, $\sigma_t(x,u)$ непрерывны по совокупности переменных и удовлетворяют условию

$$|b_t(x,u)-b_t(y,v)|+|\sigma_t(x,u)-\sigma_t(x,v)|\le C\bigl(|x-y|+|u-v|\bigr).$$

Пусть $(\Omega,\mathcal{F},P)$ — вероятностное пространство, $\mathcal{F}_\sqcup$ — фильтрация, причем $\mathcal{F}_0$ содержит все события вероятности нуль, w_t — $\mathcal{F}_t$–броуновское движение. Для каждого прогрессивно измеримого процесса u_t (со значениями в некотором ограниченном множестве $U\subset\mathbb{R}^m$) можно построить решение x_t стохастического уравнения

$$dx_t=b_t(x_t,u_t)\,dt+\sigma_t(x_t,u_t)\,dw_t,\quad x_0=x.$$

Рассмотрим задачу о минимизации выражения

$$J(x,t,u)=\mathbb{E}\Bigl(\int_t^T L_s(x_s,u_s)\,ds+g(x_T)\Bigr)$$

с помощью выбора управления u_t. Функции $L_s(x,u)$ и $g(x)$ предполагаются непрерывными по совокупности переменных и липшицевыми по x.

Положим

$$v(x,t)=\inf_u J(x,t,u).$$

Как и в случае детерминированной задачи оптимального управления в стохастическом случае выполняется принцип динамического программирования: для всех $0\le t\le\tau\le T$

$$v(x,t)=\inf_u\mathbb{E}\Bigl(\int_t^\tau L_s(x_s,u_s)\,ds+v(x_\tau,\tau)\Bigr).$$

Предположим, что функция $v\in C_b^{2,1}$ и проделаем неформальные выкладки, приводящие к уравнению Гамильтона–Якоби–Беллмана.

По формуле Ито

$$v(x_{t+h},t+h)-v(x,t)=\int_t^{t+h}\bigl[v_t+\frac{1}{2}\mathrm{tr}(A_s(x_s,u_s)D^2v)+\langle b_s(x_s,u_s),\nabla v\rangle\bigr]\,ds+\text{мартингал},\quad A_t=\sigma_t\sigma_t^*.$$

Следовательно, верно равенство

$$\inf_u\mathbb{E}\Bigl(\int_t^{t+h}L_s(x_s,u_s)+v_t+\frac{1}{2}\mathrm{tr}(A_s(x_s,u_s)D^2v)+\langle b_s(x_s,u_s),\nabla v\rangle\,ds\Bigr)$$

Делим на h и устремляем h к нулю. Получаем

$$\inf_{u\in U}\Bigl\{L_t(x,u)+v_t(x,t)+\frac{1}{2}\mathrm{tr}(A_t(x,u)D^2v(x,t))+\langle b_t(x,u),\nabla v(x,t)\rangle\Bigr\}=0.$$

Далее для простоты считаем, что $\sigma_t(x,u)=I$. Положим

$$H(x,t,p)=\sup_{u\in U}\Bigl\{-L_t(x,u)-\langle b_t(x,u),p\rangle\Bigr\}.$$

Функция v является решением уравнения Гамильтона–Якоби–Беллмана

$$-v_t-\frac{1}{2}\Delta v+H(x,t,\nabla v)=0$$

с условием $v(x,T)=g(x)$.

Достаточные условия оптимального контроля

Предположим, что w — решение класса $C_b^{2,1}$ задачи Коши

$$-w_t-\frac{1}{2}\Delta w+H(x,t,\nabla w)=0,\quad w(x,T)=g(x).$$

Предложение 2. *Всегда выполнено неравенство* $w(x,t) \le v(x,t)$. *Более того, если* $\sup$ *в определении функции* H *достигается на единственном значении* $u = u(x,t,p)$, *причем зависимость* u *от* x, t, p *липшицева, то* $u_t = u(x_t, t, \nabla w(x_t,t))$, *где*

$$dx_t = b_t(x_t, u(x_t, t, \nabla w(x_t,t)))\,dt + w_t, \quad x_0 = x,$$

является оптимальным контролем и $v(x,t) = w(x,t)$.

Доказательство. Пусть u_t — произвольный контроль и x_t — соответствующее решение стохастического уравнения. Применяя формулу Ито получаем равенство

$$\mathbb{E}\Big(\int_t^T L_s(x_s,u_s)\,ds + g(x_T) - w(x,t)\Big) = \mathbb{E}\Big(\int_t^T L_s(x_s,u_s) + w_t + \frac{1}{2}\Delta w + \langle b_s(x_s,u_s), \nabla w\rangle\,ds\Big).$$

В силу определения функции H правая часть оценивается снизу выражением

$$\mathbb{E}\Big(\int_t^T w_t + \frac{1}{2}\Delta w - H(x_s, s, \nabla w)\,ds\Big) = 0.$$

Следовательно, имеет место неравенство

$$w(x,t) \le \mathbb{E}\Big(\int_t^T L_s(x_s,u_s)\,ds + g(x_T)\Big),$$

из которого следует оценка $w(x,t) \le v(x,t)$.

Если $u_s = u(x_s, s, \nabla w(x_s,s))$, то в проделанном рассуждении вместо оценки снизу можно написать равенство и в итоге получить $w(x,t) = v(x,t)$. $\square$

Отметим, что в условиях предложения

$$b_t(x_t, u(x_t, t, \nabla w(x_t,t))) = -H_p(x_t, t, \nabla w(x_t,t)).$$

Следовательно, уравнение Фоккера–Планка–Колмогорова для распределения μ_t процесса x_t имеет вид

$$\partial_t \mu_t = \tfrac{1}{2}\Delta\mu_t + \mathrm{div}\big(H_p(x,t)\mu_t\big) = 0.$$