13 Теорема Гартнера-Эллиса

Мы рассматривали такие векторы $\vec{Z}_n$, что

$$R_{Z_n}(\vec{h}) = \mathbf{E}\exp(\langle\vec{h}, Z_n\rangle).$$

предполагая, что

$$\ln R(\vec{h}) = \lim_{n\to\infty}\frac{1}{n}\ln R_{Z_n}(n\vec{h}),$$

причем $R(\vec{h})$ дифференцируема на всем $\mathbb{R}^d$. Мы начали доказывать теорему:

Теорема 1 (Гартнер-Эллис)**.** . *Если* $\ln R$ *существенно гладкая, непрерывная снизу функция, то* $P(\vec{Z}_n \in A)$ *удовляет ПБУ с* $\Lambda(\vec{\theta})$*.*

13.1 Окончание доказательства

Доказательство. 1.2) Мы доказали для компактного F и теперь хотим доказать для произвольного замкнутого.

Будем считать, что $\inf_F \Lambda(\vec{\theta})$ конечный, в противном случае доказа Введем $F_M = F \cap [-M, M]^d$ и для компакта F_M воспользуемся предыдущим свойством, откуда

$$\mathbf{P}(\vec{Z}_n \in F_M) \le \exp(-\inf_{F_M}\Lambda(\vec{\theta})n + \varepsilon n) \le \exp(-I_F n + 2\varepsilon n)$$

при достаточно большом M, где $I_F = -\inf_{\theta\in F}\Lambda(\vec{\theta})$.

Теперь нам нужно оценить вероятности $\mathbf{P}(\vec{Z}_n \notin F_M)$. Оценим их следующим образом:

$$\mathbf{P}(\vec{Z}_n \notin F_M) \le \mathbf{P}(\exists i : Z_{n,i} > M) + \mathbf{P}(\exists i : Z_{n,i} < -M),$$

где $Z_{n,i}$ — координаты $\vec{Z}_n$.

В силу неравенства Маркова при $\vec{e}_i = (0, ...0, 1, 0, ...0)$

$$\mathbf{P}(Z_{n,i} > M) \le \mathbf{E}e^{nZ_{n,i}}e^{-nM} = R_n(n\vec{e}_i)e^{-nM},$$

Аналогично

$$\mathbf{P}(Z_{n,i} < -M) \le \mathbf{E}e^{-nZ_{n,i}}e^{-nM} = R_n(-n\vec{e}_i)e^{-nM}.$$

При этом в силу условия Гартнера-Эллиса

$$R_n(n\vec{e}_i) \le (R(\vec{e}_i) + \varepsilon)^n, \quad R_n(-n\vec{e}_i) \le (R(-\vec{e}_i) + \varepsilon)^n.$$

При достаточно большом M выполнено неравенство $\ln(R(\vec{e}_i)+\varepsilon)-M<(I_F+2\varepsilon)$, откуда

$$\mathbf{P}(\vec{Z}_n \in F) \le (2d+1)\exp(-I_F n + 2\varepsilon n).$$

Отсюда имеем требуемое соотношение

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(\vec{Z}_n \in F) \ge -I_F + 2\varepsilon,$$

откуда из произвольности $\varepsilon > 0$ имеем требуемое.

2) Докажем нижнюю оценку, т.е. для любого открытого G покажем, что

$$\liminf \frac{1}{n} \ln \mathbf{P}(\vec{Z}_n \in G) \ge -\inf_{\vec{\theta}\in G} \Lambda(\vec{\theta}).$$

Как и прежде, достаточно доказать, что при любых $\vec{x}$ и всех достаточно малых δ

$$\liminf \frac{1}{n} \ln \mathbf{P}(\vec{Z}_n \in U_\delta(\vec{x})) \ge -\Lambda(\vec{x}).$$

2.1) Допустим, что найдется супремум $\Lambda(\vec{x})$ достигается при некотором $\vec{h}$. Тогда $grad(\ln R(\vec{h})) = \vec{x}$ и

$$\mathbf{P}(\vec{Z}_n \in U_\delta(\vec{x})) = \int_{U_\delta(\vec{x})} \mathbf{P}(\vec{Z}_n \in d\vec{y}) = e^{-n(\vec{h},\vec{x})} R_n(n\vec{h}) \int_{U_\delta(\vec{x})} e^{-n(\vec{h},\vec{y}-\vec{x})} \mathbf{P}(\vec{Z}_n^{(n\vec{h})} \in d\vec{y}) \ge$$
$$e^{-n(\vec{h},\vec{x})} R_n(n\vec{h}) e^{-|\vec{h}|\delta} \mathbf{P}(\vec{Z}_n^{(n\vec{h})} \in U_\delta(\vec{x})).$$

Раньше мы пользовались для оценки последней вероятности ЗБЧ, но теперь последовательность Z более сложная. Оценим ее снизу следующим образом:

$$\mathbf{P}(\vec{Z}_n^{(n\vec{h})} \in U_\delta(\vec{x})) = 1 - \mathbf{P}(\vec{Z}_n^{(n\vec{h})} \notin U_\delta(\vec{x})) \ge$$
$$1 - \left(\sum_{i=1}^{d} \left(\mathbf{P}(\vec{Z}_{n,i}^{(n\vec{h})} - x_i > \delta d^{-1/2}) + \mathbf{P}(\vec{Z}_{n,i}^{(n\vec{h})} - x_i < -\delta d^{-1/2})\right)\right).$$

В силу неравенства Маркова при любом $\tilde{h} > 0$

$$\mathbf{P}(\vec{Z}_{n,i}^{(n\vec{h})} - x_i > \delta d^{-1/2}) \le \frac{\mathbf{E} e^{n\tilde{h} Z_{n,i}^{(n\vec{h})}}}{e^{n\tilde{h}(x_i + \delta d^{-1/2})}} = \frac{R_n(n(\tilde{h}\vec{e}_i + \vec{h}))}{R_n(n\vec{h}) e^{n\tilde{h}(x_i+\delta d^{-1/2})}}$$

и

$$\limsup_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(\vec{Z}_{n,i}^{(n\vec{h})} - x_i > \delta d^{-1/2}) \le \ln R(\tilde{h}\vec{e}_i + \vec{h}) - \ln R(\vec{h}) - \tilde{h} x_i + \delta d^{-1/2}.$$

Но $\ln R(\tilde h\vec e_i + \vec h) - \ln R(\vec h) - (e_i, grad \ln R(\vec h))\tilde h = o(\tilde h)$, откуда при достаточно малом $\tilde h$ величина $\limsup_{n\to\infty} \frac{1}{n}\ln \mathbf{P}(\vec Z_{n,i}^{(n\vec h)} - x_i > \delta d^{-1/2})$ отрицательна, а значит

$$\mathbf{P}(\vec Z_n^{(n\vec h)} \in U_\delta(\vec x)) \geq \frac{1}{2}$$

при достаточно большом n. Таким образом,

$$\liminf_{n\to\infty} \frac{1}{n}\ln \mathbf{P}(\vec Z_n \in U_\delta(\vec x)) \geq \Lambda(\vec x) - |\vec h|\delta,$$

откуда в силу произвольности δ имеем требуемое.

2.2) Как и в теореме Крамера предположим, что супремум в определении Λ недостижим. Тогда рассмотрим $\vec X_n = \vec Y_n + \vec Z_n$, $\vec Y_n \sim \mathcal{N}(0, \mathbf{E}/(Mn))$ и не зависит от $\vec Z_n$, где M — некоторый параметр, E — единичная матрица. Тогда

$$\ln R_{\vec X_n}(n\vec h) = \ln R_n(n\vec h) + \frac{n}{2M}|\vec h|^2 \geq \ln R_n(n\vec h)$$

и

$$\ln \tilde R(\vec h) = \lim_{n\to\infty} \frac{1}{n}\ln R_{X_n}(n\vec h) = \ln R(\vec h) + \frac{n}{2M}|\vec h|^2 \geq \ln R(h),\ \tilde\Lambda(\vec\theta) \leq \Lambda(\vec\theta),$$

где $\tilde\Lambda(\vec\theta) = \sup_{\vec h}((\vec\theta, \vec h) - \ln \tilde R(\vec h))$. При этом $\ln R(\vec h) \geq (\vec h, \vec\mu)$, где $\vec\mu = grad \ln R(\vec 0)$, существующий в силу дифференцируемости $\ln R$. , а значит

$$(\vec h, \vec x) - \ln \tilde R(\vec h) \leq (\vec h, (\vec x - \vec\mu)) - \frac{1}{2M}|\vec h|^2 \to -\infty,$$

$h \to \infty$. Следовательно, супремум в $\tilde\Lambda(\vec\theta)$ достижим в конкретной точке $\vec h$, удовлетворяющей условию $\vec x = grad \ln \tilde h(\vec h)$.
В силу 2.1)

$$\liminf_{n\to\infty} \frac{1}{n}\ln \mathbf{P}(\vec X_n \in U_{\delta/2}(\vec x)) \geq -\tilde\Lambda(\vec x) \geq -\Lambda(\vec x).$$

При этом

$$\mathbf{P}(\vec Z_n \in U_\delta(\vec x)) \geq \mathbf{P}(\vec X_n \in U_{\delta/2}(\vec x)) - \mathbf{P}(|\vec Y_n| > \delta/2).$$

Модуль вектора больше $\delta/2$ только если одна из координат больше по модулю $\delta/(2\sqrt{d})$. Значит

$$\mathbf{P}(|\vec Y_n| > \delta/2) \leq 2d\left(1 - \Phi\left(\frac{\delta^2\sqrt{nM}}{2d}\right)\right).$$

В силу соотношения
$$\Phi(x) \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2},\ x \to \infty,$$
имеем
$$\mathbf{P}\left(\frac{Y_1 + ... + Y_n}{n} > \frac{\delta}{2}\right) \leq \exp\left(-\frac{M\delta^2 n}{2d}\right).$$
Выберем M так, что $M\delta^2/(2d)$ будет больше $\Lambda(\vec{x}) + 2\varepsilon$, тогда
$$\mathbf{P}(\vec{Z}_n \in U_\delta(\vec{x})) \geq \exp(-\Lambda(\vec{x})n)e^{-\varepsilon n}\left(1 - 2de^{-\varepsilon n}\right).$$
Следовательно,
$$\liminf_{n\to\infty} \frac{1}{n} \ln \mathbf{P}(\vec{Z}_n \in U_\delta(\vec{x})) \geq -\Lambda(\vec{x}) - \varepsilon.$$
В силу произвольности ε имеем требуемое. □

Пример 1. Применим теорему Гартнера-Эллиса к однородной марковской цепи X_n с переходной матрицей P с конечным числом состояний r. Будем считать, что P в некоторой степени положительна (это равносильно неразложимости и непериодичности цепи).

Итак, для марковской цепи
$$\mathbf{P}(X_0 = i_0, ..., X_n = i_n) = p_{i_0} p_{i_0,i_1} ... p_{i_{n-1},i_n},$$
где $P = p_{i,j} = \mathbf{P}(X_1 = j | X_0 = i)$ — переходная матрица, а $p_i = \mathbf{P}(X_0 = i)$ — начальное распределение.
Положим для некоторой функции g
$$\vec{Z}_n = \frac{1}{n}\sum_{i=0}^{n} g(X_i).$$
Тогда
$$\mathbf{E}e^{n(\vec{h},\vec{Z}_n)} = \mathbf{E}\exp\left(\sum_{i=0}^{n}\langle g(X_i), \vec{h}\rangle\right) = \sum_{i_0,\ldots,i_n} p_{i_0} e^{\langle g(i_0),\vec{h}\rangle} \prod_{j=1}^{n} p_{i_{j-1},i_j} e^{\langle g(i_j),\vec{h}\rangle}.$$
Положим $Q = (p_{i,j}e^{\langle g(j),\vec{h}\rangle})$, $\vec{q} = (p_i e^{\langle g(i),\vec{h}\rangle})$. Тогда $\mathbf{E}e^{n(\vec{h},\vec{Z}_n)} = \vec{q}^t Q^n \vec{e}$, где $\vec{e} = (1, 1, ..., 1)$.
Если максимальное по модулю собственное значение λ этой матрицы вещественно и единственно, то можно ожидать, что
$$\frac{1}{n}\ln \mathbf{E}e^{n(\vec{h},\vec{Z}_n)} \to \ln\lambda(\vec{h}).$$
В нашем случае это верно, так как верна теорема Перрона-Фробениуса:

Теорема 2. *Если Q — матрица, в какой-то степени имеющая только положительные элементы, то существует собственное значение $\lambda > 0$, такое что*

1. *Оно имеет кратность 1.*
2. *Остальные собственные значения не превосходят его по модулю.*
3. *У соответствующего ему собственного вектора $\vec{e}_\lambda$ все координаты положительны.*

Тогда для вектора $\vec{e}$ найдутся $0 < a < b$, такие что

$$a\vec{e}_\lambda \leq \vec{e} \leq b\vec{e}_\lambda,$$

где неравенство между векторами подразумевается в смысле неравенств между каждой компонентой.
Тогда

$$a\lambda^n \langle \vec{q}, \vec{e}_\lambda \rangle \leq \vec{q}^T Q^n \vec{e} \leq b\vec{q}^T Q^n \vec{e}_\lambda = b\lambda^n \langle \vec{q}, \vec{e}_\lambda \rangle,$$

откуда

$$\frac{1}{n} \ln \mathbf{E} e^{n(\vec{h}, \vec{Z}_n)} \to \ln \lambda.$$

Здесь $\ln \lambda = \ln \lambda(h)$ оказывается в роли $\ln R(h)$ из теоремы Гартнера-Эллиса. Она определена на всем пространстве, поэтому условия теоремы выполнены (хотя дифференцируемость функции λ неочевидна). Поэтому сумма