Спецкурс "Большие уклонения"

Шкляев А.В.

27 ноября 2017 г.

Теорема Гартнера-Эллиса

Пример 5.1. Применим теорему Гартнера-Эллиса к однородной марковской цепи X_n с переходной матрицей P с положительными элементами и конечным числом состояний r.
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

$$\vec{Z}_n = \left(\frac{1}{n} \sum_{i=0}^{n} I_{X_i = j}, j = 1...r \right).$$

Тогда

$$\mathbf{E} e^{n(\vec{h}, \vec{Z}_n)} = \mathbf{E} \exp \left(\sum_{j=1}^{r} \sum_{i=0}^{n} I_{X_i = j} h_j \right) = \sum_{i_0, ..., i_n} p_{i_0} e^{h_{i_0}} \prod_{j=1}^{n} p_{i_{j-1}, i_j} e^{h_{i_j}}.$$

Положим $Q = (p_{i,j} e^{h_j})$, $\vec{q} = (p_i e^{h_i})$. 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 вещественно и единственно, то можно ожидать, что

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

В нашем случае это верно, так как верна теорема Перрона-Фробениуса:
Теорема Если P — матрица с положительными элементами, 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 любых векторов $\vec{x}$ с положительными координатами найдутся $0 < a < b$, такие что

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

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

$$\ln \lambda \leq \lim_{n \to \infty} \frac{1}{n} \ln \sum_{j=1}^{n} P_{i,j}^n x_j \leq \ln \lambda,$$

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 (кроме, возможно, дифференцируемости функции λ).
Таким образом, $\Lambda(\theta) = \sup_h ((\theta, h) - \ln \lambda(h))$. Докажем, что на самом деле она равна

$$\sup_{\vec{u} : u_i > 0} \sum_{j=1}^{k} \theta_j \ln \frac{u_j}{(uP)_j}.$$

Действительно, при любом $\vec{u} > 0$ рассмотрим $h_i = \ln(u_i / (uP)_i)$. Тогда

$$(uQ)_j = \sum_{i=1}^{k} u_i p_{i,j} \frac{u_j}{(uP)_j} = u_j,$$

откуда u — главный собственный вектор, а 1 — главное собственное значение и $\lambda(h) = 1$. При этом

$$\Lambda(\theta) \geq (\theta, h) - \ln \lambda(h) = \sum_{j=1}^{k} \theta_j \ln \frac{u_j}{(uP)_j}.$$

В обратную сторону — рассмотрим $\vec{h}$, $\vec{u}$ — левый собственный вектор матрицы Q с максимальным собственным значением.
Тогда

$$(\theta, \vec{h}) - \ln \lambda(\vec{h}) = \sum_{j=1}^{k} \theta_j \left(h_j - \ln \frac{(uQ)_j}{u_j} \right) = \sum_{j=1}^{k} \theta_j \left(\ln \frac{u_j e^{h_j}}{(uQ)_j} \right) = \sum_{j=1}^{k} \theta_j \ln \frac{u_j}{(uP)_j},$$

откуда имеем нужное соотношение.

Таким образом,

$$\Lambda(\theta) = \sup_{\vec{u}:u_i>0} \sum_{j=1}^{k} \theta_j \ln \frac{u_j}{(uP)_j}.$$

Достаточно очевидно, что если мы вместо n везде поставим какую-либо другую целочисленную последовательность $1/a_n \to \infty$, то получим аналогичный результат с условием

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

и результатом, где n будет заменена на a_n. Вопрос сведения одного к другому есть просто вопрос перепараметризации Z_n в виде Z_{a_n}. Да и целочисленность здесь не столь важна, поскольку замена a_n на $[a_n]$ не меняет ни условия, ни утверждения. С помощью этого нехитрого рассуждения удается получить теорему об умеренных уклонениях:
Теорема 5.1. Пусть X_i — н.о.р. случайные векторы, $\mathbf{E}X_1 = 0$, Σ_{X_1} — матрица ковариации X_1 обратима, $R(h)$ конечна в некоторой окрестности нуля. Тогда $Z_n = \frac{1}{\sqrt{na_n}} S_n$, $a_n \to \infty$, $a_n = o(n)$ удовлетворяет

$$-\frac{1}{2} \inf_{A_{int}} \tilde{\Lambda}(\theta) \le \liminf_{n\to\infty} \frac{1}{a_n} \ln \mathbf{P}(Z_n \in A) \le \limsup_{n\to\infty} \frac{1}{a_n} \ln \mathbf{P}(Z_n \in A) \le -\frac{1}{2} \inf_{\overline{A}} \tilde{\Lambda}(\theta),$$

с $\tilde{\Lambda}(\theta) = (\theta, \Sigma^{-1}\theta)$.
Доказательство Теоремы 5.1. Положим $R(h) = e^{(h,\Sigma h)/2}$. Тогда

$$\Lambda(\theta) = \sup((h,\theta) - (h,\Sigma h)/2) = -(\theta, \Sigma^{-1}\theta)/2,$$

поскольку максимум достигается при $((h,\theta) - (h,\Sigma h)/2)' = \theta - \Sigma h = 0$. Таким образом, нам достаточно доказать, что

$$\lim_{n\to\infty} a_n \ln R_{Z_n}(a_n^{-1} h) \to \ln R(h) = (h, \Sigma h)/2.$$

При этом

$$\ln R_{Z_n}(a_n^{-1} h) = \ln \mathbf{E} e^{(S_n,h) a_n^{-1/2} n^{-1/2}} = n \ln R_{X_1}(h(na_n)^{-1/2}).$$

Но для $b_n = (na_n)^{-1/2} \to 0$

$$R_{X_1}(hb_n) = 1 + (R'_{X_1}(0), h) b_n + \frac{(R''_{X_1}(0)h, h)}{2} b_n^2 + O(b_n^3),$$

где $R'_{X_1}(0) = \vec{0}$, $R''_{X_1}(0) = \Sigma$, откуда

$$\ln R_{X_1}(hb_n) = \ln \left(1 + \frac{(R''_{X_1}(0)h, h)}{2} b_n^2 + O(b_n^3)\right) = O(b_n^3) + \frac{(\Sigma h, h) b_n^2}{2}.$$

Следовательно,

$$na_n \ln R_{X_1}(h(na_n)^{-1/2}) \to (\Sigma h, h)/2,$$

что и требовалось доказать.

Пример 5.2. Рассмотрим скалярные X_i, $\mathbf{E}X_i = 0$, $\mathbf{D}X_i = 1$ и поглядим на $P(S_n \ge n^\beta x)$, $x > 0$, $\beta \in (1/2, 1)$. Тогда $\tilde{\Lambda}(\theta) = \theta^2/2$ и теорема приобретает вид

$$\lim \frac{\mathbf{P}(n^{(\alpha-1)/2} S_n \ge \theta)}{n^\alpha} = -\theta^2/2.$$

Полагая $\alpha = 1 - 2\beta$, имеем

$$\lim \frac{\mathbf{P}(S_n \ge n^\beta \theta)}{n^{2\beta-1}} = -\theta^2/2.$$

То есть

$$\mathbf{P}(S_n \ge n^\beta \theta) \approx e^{-\frac{n^{2\beta-1}\theta^2}{2}}.$$

Стоит отметить, что если $S_n/\sqrt{n}$ была бы $\mathcal{N}(0,1)$, то

$$\mathbf{P}(S_n \ge n^\beta \theta) = \mathbf{P}(S_n n^{-1/2} \ge n^{\beta-1/2}\theta) \sim e^{-n^{2(\beta-1/2)}\theta^2/2} = e^{-n^{2\beta-1}\theta^2/2}$$

Таким образом, нормальная аппроксимация в ЦПТ достаточна для грубой асимптотики вероятностей $\mathbf{P}(S_n \ge n^\beta \theta)$. Позже мы покажем, что более тонкая асимптотика в ЦПТ недостаточна.

Перейдем к доказательству теоремы Гартнера-Эллиса. Обсудим два значимых вопроса.
Замечание 1. 1) Функция R гладкая выпуклая ф-ия, Λ выпуклая ф-ия роста.
2) Если найдется h: $grad \ln R(\vec{h}) = \vec{\theta}$, то $\Lambda(\vec{\theta}) = (\vec{\theta}, \vec{h}) - \ln R(\vec{h})$.
Доказательство части 2) и второй половины 1) повторяет уже проведенные доказательства для теоремы Крамера. Выпуклость R следует из выпуклости R_{Z_n}.
Замечание 2. Можно заметить, что

$$\Lambda(\vec{\theta}) \le \liminf \Lambda_n(\vec{\theta}),$$

т.к.

$$\Lambda(\vec{\theta}) = \sup_{\vec{h}}((\vec{h}, \vec{\theta}) - \ln R(\vec{h})) \le \liminf \sup_{\vec{h}} (\vec{h}, \vec{\theta}) - n^{-1} \ln R(\vec{h}n) = \liminf \Lambda_n(\vec{\theta}).$$

Может возникнуть подозрение, что $\Lambda(\vec{\theta}) = \lim \Lambda_n(\vec{\theta})$. Но это не так. Для примера можно рассмотреть $Z_n = 1/n$. Тогда $R_n(h) = e^{h/n}$, $R(h) = 1$, $\Lambda_n(\theta) = \infty$, $\Lambda(\theta) = \infty$, $\theta \ne 0$ и $\Lambda(0) = 0$.