Лекция 6
Нелинейные стохастические условно-гауссовские модели

Артемов А. В.
мФТиАД ФКН ВШЭ

1 марта 2018

1 Нелинейная модель условной неоднородности ARCH

Казалось бы, линейные модели всем хороши: имеют широкое применение на практике и устроены очень просто. Однако всё не так радужно. На практике в данных могут возникать самые разные феномены, которые линейная модель описать не может. Например, если мы смотрим на цены, то могут возникать: кластеризация, катастрофическое изменение, тяжёлые хвосты распределений величин $h = (h_n)$ (см. выше), наличие «долгой памяти» у цен и её свойств и так далее. Для того, чтобы как-то описать их, обращаются к *нелинейным* моделям. Таких моделей много, а особенностей в данных ещё больше, поэтому перед исследователями возникает далеко не самая тривиальная задача подбора «подходящей» модели.

Как и раньше, пусть $(\Omega, \mathcal{F}, \mathsf{P})$ — исходное вероятностное пространство, а $\varepsilon = (\varepsilon_n)$ — последовательность iid случайных величин с стандартным нормальным распределением (они будут моделировать «случайность» в рассматриваемых далее моделях). Далее, введём фильтрацию $(\mathcal{F}_n)_{n\geqslant 0}$ по правилу: $\mathcal{F}_0 = \{\varnothing, \Omega\}$, $\mathcal{F}_n = \sigma(\varepsilon_1, \dots, \varepsilon_n)$.

Модель ARCH(p) была введена Робертом Энглем следующим образом:

Определение 1. Будем называть последовательность случайных величин $h = (h_n)$ ARCH(p)-моделью, если

$$h_n = \sigma_n \varepsilon_n, \text{ где } \sigma_n^2 = \alpha_0 + \sum_{k=1}^{p} \alpha_i h_{n-k}^2,$$

$\alpha_0 > 0$, $\alpha_i \geqslant 0$, $h_0 = h_0(\omega)$ — случайная величина, не зависящая от $\varepsilon = (\varepsilon_n)_{n\geqslant 1}$.

Обычно h_0 полагают либо константой, либо случайной величиной, для которой второй момент выбирается из соображений «стационарности» значений $\mathsf{E}[h_n^2]$.

Из формулы для σ_n^2 видна явная зависимость от $h_{n-1}^2, \dots, h_{n-p}^2$. При этом ясно, что большие (малые) значения h_{n-k}^2 приводят к большим (малым) значениям σ_n^2. Возникновение же большого значения h_n^2 при условии, что $h_{n-1}^2, \dots, h_{n-p}^2$ были мылыми, происходит из-за возникновения большого значения ε_n. Это объясняет то, почему нелинейные модели могут помочь в описании событий наподобие кластерности, то есть группирования значений в пачки «больших» и пачки «маленьких» значений.

ARCH расшифровывается, как *авторегрессионная модель условной неоднородности*, или AutoRegressive Conditional Heteroskedastic. Смысл каждого слова весьма понятен:

авторегрессивная = прямая зависимость от своих предыдущих значений, условная = задаётся условное распределение $\mathrm{Law}(h_n \mid \mathcal{F}_{n-1})$, неоднородность = σ_n^2 ведёт себя весьма неоднородно.

Теперь рассмотрим модель ARCH(1). Для неё

$$h_n = \sigma_n \varepsilon_n, \text{ где } \sigma_n^2 = a_0 + \alpha_1 h_{n-1}^2.$$

Несложно понять, что выполнены следующие условия:

$$\mathsf{E}[h_n] = 0, \quad \mathsf{E}\big[h^2\big] = \alpha_0 + \alpha_1 \mathsf{E}\big[h_{n-1}^2\big], \quad \mathsf{E}\big[h^2 \mid \mathcal{F}_{n-1}\big] = \sigma_n^2 = a_0 + \alpha_1 h_{n-1}^2.$$

Если предположить, что $\alpha_1 \in (0,1)$, то рекуррентное соотношение на матожидание квадрата будет иметь единственное «стационарное» решение: $\mathsf{E}\big[h_{n-1}^2\big] = \alpha_0/(1-\alpha_1)$. Если взять $h_0^2 = \alpha_0/(1-\alpha_1)$, то матожидание квадрата будет постоянно.

Далее, посчитаем четвёртый момент, пользуясь независимостью σ_n и ε_n и тем, что $\mathsf{E}[\varepsilon_n^4] = 3$ (проверьте!):

$$\begin{aligned}
\mathsf{E}\big[h_n^4\big] &= \mathsf{E}\big[\sigma_n^4\big]\,\mathsf{E}\big[\varepsilon_n^4\big] = 3\,\mathsf{E}\big[(\alpha_0 + \alpha_1 h_{n-1}^2)^2\big] = \\
&= 3(\alpha_0^2 + 2\alpha_0\alpha_1\,\mathsf{E}\big[h_{n-1}^2\big] + \alpha_1^2\,\mathsf{E}\big[h_{n-1}^4\big]) = \\
&= \frac{3\alpha_0^2(1+\alpha_1)}{1-\alpha_1} + 3\alpha_1^2\,\mathsf{E}\big[h_{n-1}^4\big].
\end{aligned}$$

Если предположить, что $\alpha_1 \in (0,1)$ и $3\alpha_1^2 < 1$, то можно найти «стационарное решение»:

$$\mathsf{E}\big[h_n^4\big] = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)}.$$

Из полученных формул несложно получить, что стационарное значение коэффициента эксцесса равно

$$K \equiv \frac{\mathsf{E}[h_n^4]}{(\mathsf{E}[h_n^2])^2} - 3 = \frac{3(1-\alpha_1^2)}{1-3\alpha_1^2} - 3 = \frac{6\alpha_1^2}{1-3\alpha_1^2}.$$

Его положительность говорит о том, что плотность «установившегося» распределения $h = (h_n)$ в окрестности среднего значения «вытянута» вверх. Напомним, что для нормального распределения эксцесс равен нулю.

Теперь заметим, что наша модель задаёт *мартингал-разность*, то есть $\mathsf{E}[h_n \mid \mathcal{F}_{n-1}] = 0$. Отсюда следует, что для любого $k < n$

$$\mathsf{E}[h_n h_k] = \mathsf{E}[\mathsf{E}[h_n h_k \mid \mathcal{F}_{n-1}]] = \mathsf{E}[h_k\,\mathsf{E}[h_n \mid \mathcal{F}_{n-1}]] = 0.$$

Это означает *ортогональность* значений: $\mathsf{cov}(h_n, h_m) = 0$ при $n \neq m$. Но из ортогональности этих значений не следует независимость, так как совместное распределение $\mathrm{Law}(h_m, h_n)$ не является гауссовским при $\alpha_1 > 0$ (почему?). Но если они не независимы, то между ними есть какая-то зависимость. Как её исследовать? Для этого посмотрим на корреляционную зависимость *квадратов* h_n^2 и h_m^2 в «стационарном» случае. Посчитаем дисперсию и ковариацию для соседних значений:

$$\begin{aligned}
\mathsf{D}\big[h_n^2\big] &= \mathsf{E}\big[h_n^4\big] - (\mathsf{E}\big[h_n^2\big])^2 = \frac{3\alpha_0^2(1+\alpha_1)}{(1-\alpha_1)(1-3\alpha_1^2)} - \frac{\alpha_0^2}{(1-\alpha_1)^2} = \\
&= \left(\frac{\alpha_0}{1-\alpha_1}\right)^2 \left(\frac{3-3\alpha_1^2}{1-3\alpha_1^2} - 1\right) = \frac{2}{1-3\alpha_1^2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^2. \\
\mathsf{E}\big[h_n^2 h_{n-1}^2\big] &= \mathsf{E}\big[(\alpha_0+\alpha_1 h_{n-1}^2)\varepsilon_n^2 h_{n-1}^2\big] = \mathsf{E}\big[\alpha_0 h_{n-1}^2\big] + \mathsf{E}\big[\alpha_1 h_{n-1}^4\big] = \\
&= \frac{\alpha_0^2}{1-\alpha_1} + \frac{2\alpha_1}{1-3\alpha_1^2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^2 = \frac{\alpha_0^2}{1+\alpha_1}\left(1 - \frac{2\alpha_1}{(1-\alpha_1)(1-3\alpha_1^2)}\right) = \\
&= \frac{\alpha_0^2}{1-\alpha_1}\left(\frac{1-\alpha_1-3\alpha_1^2+3\alpha_1^2+2\alpha_1}{(1-3\alpha_1^2)(1-\alpha_1)}\right) = \frac{\alpha_0^2}{1-\alpha_1}\frac{1+3\alpha_1}{1-3\alpha_1^2}, \\
\mathsf{cov}(h_n^2, h_{n-1}^2) &= \mathsf{E}\big[h_n^2 h_{n-1}^2\big] - \mathsf{E}\big[h_n^2\big]\,\mathsf{E}\big[h_{n-1}^2\big] = \frac{1+3\alpha_1}{1-3\alpha_1^2}\frac{\alpha_0^2}{1-\alpha_1} - \frac{\alpha_0^2}{(1-\alpha_1)^2} = \\
&= \left(\frac{\alpha_0}{1-\alpha_1}\right)^2\left(\frac{(1-\alpha_1)(1+3\alpha_1)}{1-3\alpha_1^2} - 1\right) = \frac{2\alpha_1}{1-3\alpha_1^2}\left(\frac{\alpha_0}{1-\alpha_1}\right)^2, \\
r(1) &\equiv \frac{\mathsf{cov}(h_n^2, h_{n-1}^2)}{\sqrt{\mathsf{D}\big[h_n^2\big]\,\mathsf{D}\big[h_{n-1}^2\big]}} = \alpha_1.
\end{aligned}$$

Далее посчитаем корреляцию в общем случае. Заметим, что для $k < n$

$$\begin{aligned}
\mathsf{E}\big[h_n^2 h_{n-k}^2\big] &= \mathsf{E}\big[\mathsf{E}\big[h_n^2 h_{n-k}^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \mathsf{E}\big[h_{n-k}^2\,\mathsf{E}\big[h_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \\
&= \mathsf{E}\big[h_{n-k}^2\,\mathsf{E}\big[\sigma_n^2\varepsilon_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \mathsf{E}\big[h_{n-k}^2\sigma_n^2\,\mathsf{E}\big[\varepsilon_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \\
&= \mathsf{E}\big[h_{n-k}^2(\alpha_0+\alpha_1 h_{n-1}^2)\big] = \alpha_0\,\mathsf{E}\big[h_{n-k}^2\big] + \alpha_1\,\mathsf{E}\big[h_{n-1}^2 h_{n-k}^2\big].
\end{aligned}$$

Тогда в стационарном случае это равенство можно преобразовать следующим образом:

$$\mathsf{E}\big[h_n^2 h_{n-k}^2\big] - \left(\frac{\alpha_0}{1-\alpha_1}\right)^2 = \frac{\alpha_0^2}{1-\alpha_1} - (1-\alpha_1)\left(\frac{\alpha_0}{1-\alpha_1}\right)^2 + \alpha_1\left(\mathsf{E}\big[h_{n-1}^2 h_{n-k}^2\big] - \left(\frac{\alpha_0}{1-\alpha_1}\right)^2\right).$$

Отсюда следует, что

$$r(k) = \alpha_1 r(k-1) \implies r(k) = \alpha_1^k.$$

В названии модели ARCH(p) фигурирует слово «авторегрессионая». Оказывается, что модель ARCH(p) сводится к AR(p)-модели. Действительно, пусть $\nu_n = h_n^2 - \sigma_n^2$. Если $\mathsf{E}[h_n^2] < \infty$, то $\mathsf{E}[\nu_n \mid \mathcal{F}_{n-1}] = 0$ и $\nu = (\nu_n)$ образует мартингал-разность относительно $(\mathcal{F}_n)_{n\geqslant 0}$. Далее, введём обозначение $x_n = h_n^2$. Тогда

$$x_n = \alpha_0 + \alpha_1 x_{n-1} + \ldots + \alpha_p x_{n-p} + \nu_n.$$

2 Обобщенная нелинейная модель GARCH

Успех условно-гауссовской модели ARCH(p), давшей объяснение многим феноменам в поведении финансовых индексов, породил целую кучу различных её обобщений, преследующих цель «ухватить», дать описание ряда других эффектов. Исторически первое обобщение было введено Тимом Боллерслевом в 1986-м году: так называемая *обобщённая* ARCH-*модель*, характеризуемая двумя параметрами (p,q). Её принято обозначать GARCH(p,q). Определяется она следующим образом:

$$h_n = \sigma_n\varepsilon_n, \text{ где } \sigma_n^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i h_{n-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{n-j}^2.$$

Основное преимущество GARCH(p, q)-моделей перед их прародительницей, ARCH(p)-моделью, состоит в подборе параметров модели. На практике периодически может оказаться так, что при подгонке статистических данных моделями ARCH(p) параметр p становится слишком большим (что усложняет анализ модели), в то время как при подгонке GARCH(p, q)-моделями можно ограничиваться небольшими значениями p и q (экспериментальный факт!).

Как и всегда, подробнее рассмотрим частный случай: GARCH$(1, 1)$. Для него

$$h_n = \sigma_n \varepsilon_n, \text{ где } \sigma_n^2 = \alpha_0 + \alpha_1 h_{n-1}^2 + \beta_1 \sigma_{n-1}^2.$$

Отсюда ясно, что

$$\mathsf{E}[h_n] = 0, \quad \mathsf{E}\big[h_n^2\big] = \mathsf{E}\big[\sigma_n^2\big] = \alpha_0 + \alpha_1 \mathsf{E}\big[h_{n-1}^2\big] + \beta_1 \mathsf{E}\big[\sigma_{n-1}^2\big] = \alpha_0 + (\alpha_1 + \beta_1) \mathsf{E}\big[h_{n-1}^2\big].$$

Если $\alpha_1 + \beta_1 < 1$, то существует «стационарное» значение $\mathsf{E}[h_n^2]$, равное

$$\mathsf{E}\big[h_n^2\big] = \frac{\alpha_0}{1 - \alpha_1 - \beta_1}.$$

Далее, посчитаем четвёртый момент.

$$\begin{aligned}
\mathsf{E}\big[h_n^4\big] &= \mathsf{E}\big[\sigma_n^4\big] \mathsf{E}\big[\varepsilon_n^4\big] = 3\,\mathsf{E}\big[(\alpha_0 + \alpha_1 h_{n-1}^2 + \beta_1 \sigma_{n-1}^2)^2\big] = \\
&= 3(\alpha_0^2 + \alpha_1^2 \mathsf{E}\big[h_{n-1}^4\big] + \beta_1^2 \mathsf{E}\big[\sigma_1^4\big] + 2\alpha_0\alpha_1 \mathsf{E}\big[h_{n-1}^2\big] + 2\alpha_0\beta_1 \mathsf{E}\big[\sigma_{n-1}^2\big] + \\
&+ 2\alpha_1\beta_1 \mathsf{E}\big[h_{n-1}^2 \sigma_{n-1}^2\big]) = 3\alpha_0^2 + 3\alpha_1^2 \mathsf{E}\big[h_{n-1}^4\big] + \beta_1^2 \mathsf{E}\big[h_{n-1}^4\big] + \\
&+ 6\alpha_0\alpha_1 \mathsf{E}\big[h_{n-1}^2\big] + 6\alpha_0\beta_1 \mathsf{E}\big[h_{n-1}^2\big] + 2\alpha_1\beta_1 \mathsf{E}\big[h_{n-1}^4\big].
\end{aligned}$$

Если предположить, что $3\alpha_1^2 + \beta_1^2 + 2\alpha_1\beta_1 < 1$, то можно найти «стационарное» решение:

$$(1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1) \mathsf{E}\big[h_n^4\big] = 3\alpha_0^2 \left(1 + \frac{2(\alpha_1 + \beta_1)}{1 - \alpha_1 - \beta_1}\right),$$

$$\mathsf{E}\big[h_n^4\big] = \frac{3\alpha_0^2 (1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1)}.$$

Отсюа можно получить коэффициент эксцесса:

$$K = \frac{3(1 - (\alpha_1 + \beta_1)^2)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1} - 3 = \frac{6\alpha_1^2}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}.$$

Теперь посчитаем корреляционную функцию для этой модели. Заметим, что

$$\begin{aligned}
\mathsf{E}\big[h_n^2 h_{n-1}^2\big] &= \mathsf{E}\big[(\alpha_0 + \alpha_1 h_{n-1}^2 + \beta_1 \sigma_{n-1}^2) \varepsilon_n^2 h_{n-1}^2\big] = \\
&= \alpha_0 \mathsf{E}\big[h_{n-1}^2\big] + (\alpha_1 + \beta_1/3) \mathsf{E}\big[h_{n-1}^4\big] = \\
&= \frac{\alpha_0^2}{1 - \alpha_1 - \beta_1} + \frac{\alpha_0^2 (3\alpha_1 + \beta_1)(1 + \alpha_1 + \beta_1)}{(1 - \alpha_1 - \beta_1)(1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1)} = \\
&= \frac{\alpha_0^2}{1 - \alpha_1 - \beta_1} \left(1 + \frac{(3\alpha_1 + \beta_1)(1 + \alpha_1 + \beta_1)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}\right) \\
\mathsf{cov}(h_n^2, h_{n-1}^2) &= \mathsf{E}\big[h_n^2 h_{n-1}^2\big] - \mathsf{E}\big[h_n^2\big] \mathsf{E}\big[h_{n-1}^2\big] = \\
&= \frac{\alpha_0^2}{1 - \alpha_1 - \beta_1} \left(1 + \frac{(3\alpha_1 + \beta_1)(1 + \alpha_1 + \beta_1)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}\right) - \frac{\alpha_0^2}{(1 - \alpha_1 - \beta_1)^2} = \\
&= \frac{\alpha_0^2}{(1 - \alpha_1 - \beta_1)^2} \left(1 - \alpha_1 - \beta_1 + \frac{(3\alpha_1 + \beta_1)(1 - (\alpha_1 + \beta_1)^2)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1} - 1\right) = \\
&= \frac{\alpha_0^2}{(1 - \alpha_1 - \beta_1)^2} \left(\frac{(3\alpha_1 + \beta_1)(1 - \alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1} - (\alpha_1 + \beta_1)\right) = \\
&= \left(\frac{\alpha_0}{1 - \alpha_1 - \beta_1}\right)^2 \frac{2\alpha_1 (1 - \alpha_1\beta_1 - \beta_1^2)}{1 - 3\alpha_1^2 - \beta_1^2 - 2\alpha_1\beta_1}
\end{aligned}$$

$$\begin{aligned}
\mathsf{D}\big[h_n^2\big] &= \mathsf{E}\big[h_n^4\big] - (\mathsf{E}\big[h_n^2\big])^2 = \frac{3\alpha_0^2(1+\alpha_1+\beta_1)}{(1-\alpha_1-\beta_1)(1-3\alpha_1^2-\beta_1^2-2\alpha_1\beta_1)} - \frac{\alpha_0^2}{(1-\alpha_1-\beta_1)^2} = \\
&= \frac{\alpha_0^2}{(1-\alpha_1-\beta_1)^2}\left(\frac{3(1-\alpha_1^2-\beta_1^2-2\alpha_1\beta_1)}{1-3\alpha_1^2-\beta_1^2-2\alpha_1\beta_1} - 1\right) = \\
&= \left(\frac{\alpha_0}{1-\alpha_1-\beta_1}\right)^2 \frac{2(1-2\alpha_1\beta_1-\beta_1^2)}{1-3\alpha_1^2-\beta_1^2-2\alpha_1\beta_1}
\end{aligned}$$

Отсюда сразу же получаем, что

$$\rho(1) \equiv \frac{\mathsf{cov}(h_n^2, h_{n-1}^2)}{\sqrt{\mathsf{D}[h_n^2]\,\mathsf{D}\big[h_{n-1}^2\big]}} = \frac{\alpha_1(1-\alpha_1\beta_1-\beta_1^2)}{1-2\alpha_1\beta_1-\beta_1^2}.$$

Далее заметим, что

$$\begin{aligned}
\mathsf{E}\big[h_n^2 h_{n-k}^2\big] &= \mathsf{E}\big[\mathsf{E}\big[h_n^2 h_{n-k}^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \mathsf{E}\big[h_{n-k}^2\,\mathsf{E}\big[h_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \\
&= \mathsf{E}\big[h_{n-k}^2\,\mathsf{E}\big[\sigma_n^2\varepsilon_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \mathsf{E}\big[h_{n-k}^2\sigma_n^2\,\mathsf{E}\big[\varepsilon_n^2 \,\big|\, \mathcal{F}_{n-1}\big]\big] = \\
&= \mathsf{E}\big[h_{n-k}^2(\alpha_0+\alpha_1 h_{n-1}^2+\beta_1\sigma_{n-1}^2)\big] = \\
&= \alpha_0\,\mathsf{E}\big[h_{n-k}^2\big] + (\alpha_1+\beta_1)\,\mathsf{E}\big[h_{n-1}^2 h_{n-k}^2\big]
\end{aligned}$$

Следовательно, в стационарном случае это можно перезаписать следующим образом:

$$\begin{aligned}
\mathsf{E}\big[h_n^2 h_{n-k}^2\big] - \left(\frac{\alpha_0}{1-\alpha_1-\beta_1}\right)^2 &= \frac{\alpha_0^2}{1-\alpha_1-\beta_1} + (\alpha_1+\beta_1-1)\left(\frac{\alpha_0}{1-\alpha_1-\beta_1}\right)^2 + \\
&\quad + (\alpha_1+\beta_1)\left(\mathsf{E}\big[h_{n-1}^2 h_{n-k}^2\big] - \left(\frac{\alpha_0}{1-\alpha_1-\beta_1}\right)^2\right)
\end{aligned}$$

Тогда $\mathsf{cov}(h_n^2, h_{n-k}^2) = (\alpha_1+\beta_1)\,\mathsf{cov}(h_{n-1}^2, h_{n-k}^2)$ и

$$\rho(k) \equiv \frac{\mathsf{cov}(h_n^2, h_{n-k}^2)}{\sqrt{\mathsf{D}[h_n^2]\,\mathsf{D}\big[h_{n-k}^2\big]}} = \frac{\alpha_1(1-\alpha_1\beta_1-\beta_1^2)}{1-2\alpha_1\beta_1-\beta_1^2}(\alpha_1+\beta_1)^{k-1}.$$