Оптимальные инструменты*

Станислав Анатольев†

Российская экономическая школа, Москва

Настоящее эссе содержит краткий обзор оптимального инструментирования в линейных и нелинейных моделях, как кросс-секционных, так и на стационарных временных рядах. Разобраны примеры разумного построения инструментов.

1 Безусловные кросс-секционные модели

Рассмотрим привычное линейное уравнение с инструментальными переменными

$$y = x'\beta + e, \quad \mathbb{E}[ze] = 0,$$

где x является $k \times 1$ вектором объясняющих переменных («регрессоров»), β – оцениваемым $k \times 1$ вектором параметров, z – $l \times 1$ вектором *базовых* экзогенных инструментов, $l \geqslant k$. Инструменты используются из-за потенциальной эндогенности объясняющих переменных. В частном случае линейной регрессии имеем $z = x$ и $l = k$. Пусть имеется случайная выборка $\{(x_i, y_i, z_i)\}_{i=1}^n$, т.е. рассматриваются кросс-секции.

Вспомним необходимые условия качества базовых инструментов z.

- Условие «годности» инструментов, т.е. их некоррелированности с ошибками:

$$\mathbb{E}[ze] = 0.$$

- Условие «релевантности» инструментов, т.е. их коррелированности с регрессорами:

$$\underset{k\times l}{Q_{xz}} = \mathbb{E}[xz'] \text{ – матрица полного ранга } k.$$

- Условие линейной «неповторяемости» инструментов:

$$\underset{l\times l}{Q_{zz}} = \mathbb{E}[zz'] \text{ – невырожденная матрица.}$$

В случае точной идентификации ($l = k$) используется простая оценка инструментальных переменных (ИП)

$$\widehat{\beta}_{IV} = \left(\sum_{i=1}^n z_i x_i'\right)^{-1} \sum_{i=1}^n z_i y_i.$$

В случае сверхидентификации ($l > k$) мы из определенных проекционных соображений строим оценку двухшагового метода наименьших квадратов (2ШМНК):

$$\widehat{\beta}_{2SLS} = \left(\sum_{i=1}^n x_i z_i' \left(\sum_{i=1}^n z_i z_i'\right)^{-1} \sum_{i=1}^n z_i x_i'\right)^{-1} \sum_{i=1}^n x_i z_i' \left(\sum_{i=1}^n z_i z_i'\right)^{-1} \sum_{i=1}^n z_i y_i.$$

*Цитировать как: Анатольев, Станислав (2007) «Оптимальные инструменты», Квантиль, №2, стр. 61–69. Citation: Anatolyev, Stanislav (2007) "Optimal instruments," Quantile, No.2, pp. 61–69.

†Адрес: 117418, г. Москва, Нахимовский проспект, 47, офис 1721(3). Электронная почта: *sanatoly@nes.ru*

Заметим, что эту оценку можно переписать как

$$\widehat{\beta}_{2SLS} = \left(\sum_{i=1}^{n} \widehat{\zeta}_i x_i'\right)^{-1} \sum_{i=1}^{n} \widehat{\zeta}_i y_i,$$

где $\widehat{\zeta}_i$ представляет собой i-е «наблюдение» вектора

$$\underset{k\times 1}{\widehat{\zeta}} = \frac{1}{n}\sum_{j=1}^{n} x_j z_j' \left(\frac{1}{n}\sum_{j=1}^{n} z_j z_j'\right)^{-1} z = \underset{k\times l}{\widehat{Q}_{xz}}\ \underset{l\times l}{\widehat{Q}_{zz}^{-1}}\ \underset{l\times 1}{z},$$

являющегося определенным линейным преобразованием базового инструмента z. Сама же 2ШМНК-оценка выглядит как обычная (т.е. как будто бы при точной идентификации) ИП-оценка с инструментом $\widehat{\zeta}_i$ для i-го наблюдения.

Известно, что при условной гетероскедастичности асимптотически эффективной является не 2ШМНК, а эффективная ОММ-оценка

$$\begin{aligned}\widehat{\beta}_{GMM} &= \left(\sum_{i=1}^{n} x_i z_i' \left(\sum_{i=1}^{n} z_i z_i' \widehat{e}_i^2\right)^{-1} \sum_{i=1}^{n} z_i x_i'\right)^{-1} \sum_{i=1}^{n} x_i z_i' \left(\sum_{i=1}^{n} z_i z_i' \widehat{e}_i^2\right)^{-1} \sum_{i=1}^{n} z_i y_i \\ &= \left(\sum_{i=1}^{n} \widehat{\zeta}_i x_i'\right)^{-1} \sum_{i=1}^{n} \widehat{\zeta}_i y_i,\end{aligned}$$

которая также выглядит как обычная ИП-оценка с инструментом $\widehat{\zeta}$, который является линейной комбинацией базового инструмента:

$$\underset{k\times 1}{\widehat{\zeta}} = \frac{1}{n}\sum_{j=1}^{n} x_j z_j' \left(\frac{1}{n}\sum_{j=1}^{n} z_j z_j' \widehat{e}_j^2\right)^{-1} z = \underset{k\times l}{\widehat{Q}_{xz}}\ \underset{l\times l}{\widehat{Q}_{zze^2}^{-1}}\ \underset{l\times 1}{z}.$$

Отсюда следует естественная интерпретация: эффективная ОММ-оценка (или 2ШМНК-оценка в случае условной гомоскедастичности) берет первоначальные инструменты z (их l штук) и умножает их слева на выписанную выше последовательность матриц, имеющую в совокупности размерность $k \times l$. Таким образом берутся, пусть и сложные, линейные комбинации от базовых инструментов так, чтобы получился точно идентифицирующий инструмент. Используемое взвешивание является оптимальным, ибо мы рассматриваем эффективный ОММ. Таким образом, эффективный ОММ неявно «ужимает» ячейки с информацией до необходимого размера оптимальным образом. Приведенный выше $\widehat{\zeta}$ называется *оптимальным инструментом.* Используя популяционные аналоги, можно получить (конечно же, недоступный) *идеальный оптимальный инструмент*, который асимптотически эквивалентен $\widehat{\zeta}$:

$$\zeta = Q_{xz} Q_{zze^2}^{-1} z,$$

где $Q_{zze^2} = \mathbb{E}[zz'e^2]$.

2 Условные кросс-секционные модели

Большинство линейных моделей в современной эконометрике имеют вид

$$y = x'\beta + e, \quad \mathbb{E}[e|z] = 0.$$

Здесь те же обозначения, что использовались ранее; единственное отличие – в условности ограничения $\mathbb{E}[e|z] = 0$. Действительно, в той же линейной регрессии (т.е. при $z = x$) мы имеем ограничение $\mathbb{E}[e|x] = 0$, а не $\mathbb{E}[xe] = 0$; последнее условие определяет всего лишь линейную проекцию и моделью не является. Мы продолжаем рассматривать случай кросс-секций.

В такой модели не только сами базовые инструменты z и их линейные комбинации могут использоваться в качестве инструментов, но и любые нелинейные функции от z. Действительно, для любой функции $f : \mathbb{R}^l \to \mathbb{R}^k$ инструмент $f(z)$ является годным: $\mathbb{E}[f(z)e] = 0$. Естественно, подразумеваются существование всех используемых моментов и релевантность $f(z)$. Какую функцию f выбрать наиболее выгодно с точки зрения максимизации асимптотической эффективности? Тождественную, логарифмическую, синусоидальную?

Конечно, не все так просто. Как минимум, форма функции должна зависеть от свойств рассматриваемой задачи. Обратимся вновь к случаю $z = x$. При условной гомоскедастичности мы, естественно, применяем МНК, что соответствует выбору тождественной функции в качестве f; при условной же гетероскедастичности мы применили бы ОМНК (отвлекаясь пока от проблематичности его реализации), что соответствует выбору $f(x) = x/\mathbb{E}[e^2|x]$. Оказывается, что и в общем случае оптимальный инструмент имеет похожую простую форму (Chamberlain, 1987)

$$\zeta = \frac{\mathbb{E}\left[x|z\right]}{\mathbb{E}\left[e^2|z\right]}.$$

Несмотря на простоту формы оптимального инструмента, его практическая реализация проблематична: в формуле фигурируют условные матожидания, о которых в данной задаче наверняка ничего не известно, поэтому их надо оценивать, причем если это делать честно, следует применять непараметрические методы. В данном контексте было предложено применять метод ближайших соседей (Robinson, 1987) и разложения в ряды (Newey, 1990), помимо прочего. Конечно, это не очень приятно. К тому же у исследователя должна быть солидного размера выборка для получения терпимой точности непараметрического оценивания, иначе асимптотика будет «работать» плохо.

3 Оптимальные инструменты в моделях временных рядов

В моделях временных рядов теория и практика оптимальных инструментов сильно усложняются, но и становятся более интересными. Обычно модель имеет вид

$$y_t = x_t'\beta + e_t, \quad \mathbb{E}\left[e_t|\Im_t\right] = 0,$$

где x_t is $k \times 1$, z_t – $\ell \times 1$ базовый инструмент, а $\Im_t$ содержит информацию в $z_t, z_{t-1}, \ldots$. Все ряды предполагаются стационарными и эргодическими. Максимальное множество годных инструментов «дважды бесконечно»: во-первых, как в предыдущем разделе, оно включает все нелинейные функции от z_t, а во-вторых, и все лаги (также вместе с нелинейными функциями от них) z_{t-j}, $j > 0$. Скомбинировать оптимальным образом имеющуюся в инструментах и их предыстории информацию является в общем случае сверхзадачей. Мы кратко опишем имеющиеся теоретические результаты в разной сложности задачах; подробности можно узнать в Anatolyev (2007) и в цитируемых там источниках.

Если задача относится к категории «однопериодных», т.е. ошибка e_t представляет собой мартингальное приращение по отношению к своему прошлому (грубо говоря, серийно некоррелирована), то форма оптимального инструмента следующая:

$$\zeta_t = \frac{\mathbb{E}\left[x_t|\Im_t\right]}{\mathbb{E}\left[e_t^2|\Im_t\right]}.$$

Заметим, что эта форма в точности та же, какую мы наблюдали в предыдущем разделе. Например, рассмотрим условно гомоскедастичную авторегрессию $\mathsf{AR}(k)$

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \cdots + \rho_k y_{t-k} + \varepsilon_t,$$

где ε_t – строгий белый шум. Очевидно, оптимальный инструмент ζ_t пропорционален вектору $x_t = (y_{t-1}\ y_{t-2}\ \cdots\ y_{t-k})'$, что сводит дело к оцениванию с помощью МНК.

Если задача относится к категории «многопериодных», т.е. ошибка e_t серийно коррелирована до ненулевого конечного порядка q, и наличествует условная гомоскедастичность, то форма оптимального инструмента (точнее, рекурсивное соотношение для него) следующая (Hansen, 1985):

$$\Theta(L)\zeta_t = \mathbb{E}\left[\Theta(L^{-1})^{-1}x_t|\Im_t\right],$$

где $\Theta(L) = \theta_0+\theta_1 L+\cdots+\theta_q L^q$ – лаговый полином из разложения Вольда ошибки: $e_t = \Theta(L)\varepsilon_t$. Например, в «двухпериодной» (т.е. $q = 1$) условно гомоскедастичной ARMA(1,1)-модели

$$y_t = \rho y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1},$$

где ε_t – строгий белый шум, оптимальный инструмент представим в виде

$$(1-\theta L)\,\zeta_t = \frac{1}{1-\theta L}y_{t-2},$$

или в конечном счете

$$\zeta_t = \sum_{i=1}^{\infty} i\theta^{i-1}y_{t-1-i}.$$

Интересно, что оптимальный инструмент в этом случае задействует всю предысторию «объясняющей» переменной, причем взвешивает ее старые значения с затуханием, чуть более медленным, чем экспоненциальное. Задействование всей предыстории – это последствие многопериодности задачи. Можно сравнить данный случай с рассмотренным выше случаем AR-модели, где оптимальный инструмент игнорирует предысторию вовсе.

Наконец, в самой общей категории задач – «многопериодных», т.е. с серийно коррелированной до ненулевого конечного порядка q ошибкой e_t, в присутствии условной гетероскедастичности – форма оптимального инструмента следующая (Anatolyev, 2003):

$$\Phi_t(L)\zeta_t = \mathbb{E}\left[\Psi_t(L^{-1})x_t|\Im_t\right],$$

где лаговые полиномы $\Phi_t(L)$ и $\Psi_t(L)$ на этот раз зависят от момента времени t. Определяются последние из сложной нелинейной системы (фактически функциональных) уравнений, имеющей множественные решения. Дополнительное условие стабильности рекурсии, приведенной выше, выделяет из множества решений искомое, определяющее оптимальный инструмент. Более того, увы, искомое решение не выражаемо в явном виде. Чтобы взглянуть, как эта система выглядит, рассмотрим случай «двухпериодной» задачи (т.е. $q = 1$), когда $\omega_t = \mathbb{E}[e_t^2|\Im_t]$, $\gamma_t = \mathbb{E}\left[e_t e_{t-1}|\Im_t\right]$, $\mathbb{E}\left[e_t e_{t-j}|\Im_t\right] = 0$ при $j > 1$. Тогда рекурсия для оптимального инструмента выглядит как

$$\zeta_t = \phi_t(\zeta_{t-1} - \gamma_t^{-1}\delta_t),$$

а система имеет следующий вид:

$$\gamma_t + \phi_t\left(\omega_t + \mathbb{E}_t\left[\phi_{t+1}\gamma_{t+1}|\Im_t\right]\right) = 0,$$

$$\delta_t = \mathbb{E}\left[x_t + \phi_{t+1}\delta_{t+1}|\Im_t\right].$$

Наконец, условие стабильности имеет вид

$$\mathbb{E}[\log|\phi_t|] < 0.$$

Систему можно разрешить в явном виде лишь в случае условной гомоскедастичности, и оптимальный инструмент тогда сводится к рассмотренному в предыдущей категории задач.

Очевидно, что оптимальные инструменты, имеющие такую сложную форму, почти невозможно реализовать на практике, по крайней мере в чистом виде (не сильно успешные попытки описаны в Anatolyev, 2007). Поэтому большее распространение получила идея сократить множество, в котором ищется оптимальный инструмент, таким образом жертвуя какой-то долей асимптотической эффективности, зато приобретая надежду, что результирующий инструмент будет иметь не такую сложную форму, и его будет не так трудно воплотить в жизнь на практике.

Наиболее интуитивным подмножеством всего множества годных инструментов является пространство линейных комбинаций базовых инструментов. То есть жертвуются все нелинейные инструменты, зато используется вся предыстория. Очевидно, что жертва может быть значительной, если, например, вновь вспомнить МНК и ОМНК в линейной регрессии. Ведь МНК как раз использует линейным образом регрессоры, а вот использование ОМНК требует нелинейного взвешивания регрессора. С другой стороны, мы знаем, что реализовать МНК гораздо проще, чем ОМНК (ведь форму условной гетероскедастичности мы, конечно же, не знаем). А если условная гетероскедастичность несильная, то, возможно, мы не очень-то много и теряем.

Теория так называемых *линейных оптимальных инструментов* в наиболее общем виде разработана в West, Wong & Anatolyev (2002). Например, рассмотрим условно гетероскедастичную авторегрессию $\mathsf{AR}(1)$

$$y_t = \rho y_{t-1} + \varepsilon_t,$$

где ε_t – симметрично распределенное мартингальное приращение. Тогда оптимальный инструмент равен

$$\zeta_t = \sum_{i=1}^{\infty} \frac{\rho^{r-1}}{\tau_r} \varepsilon_{t-i},$$

где $\tau_r = \mathbb{E}[\varepsilon_t^2 \varepsilon_{t-r}^2]\mathbb{E}[\varepsilon_t^2]^{-2}$. В условно гетероскедастичной же $\mathsf{ARMA}(1{,}1)$-модели

$$y_t = \rho y_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1},$$

где ε_t – строгий белый шум, оптимальный инструмент выглядит еще более громоздко:

$$\zeta_t = \sum_{i=1}^{\infty} \phi_i \varepsilon_{t-1-i},$$

где

$$\phi_1 = \left(1 + \sum_{i=1}^{\infty} \theta^{2i} \frac{\tau_1}{\tau_{i+1}}\right)^{-1} \sum_{i=1}^{\infty} \theta^{i-1} \frac{\rho^i - \theta^i}{\tau_{i+1}}, \quad \phi_r = \sum_{i=0}^{\infty} \theta^i \frac{\rho^{j+i} - \theta^{j+i}\left(1 + \theta\phi_1\tau_1\right)}{\tau_{r+i+1}}.$$

Теперь в обоих примерах оба оптимальных инструмента задействует всю предысторию регрессора. К сожалению, правда заключается в том, что такой «пассивный» способ эксплуатировать условную гетероскедастичность не дает большого выигрыша по сравнению с ее неэксплуатацией. То есть, например, в $\mathsf{AR}(1)$-модели оптимальная ИП-оценка по асимптотическим свойствам гораздо ближе к МНК, чем к ОМНК (см. Anatolyev, 2007). С другой стороны, вся система весов зависит от сравнительно легко оцениваемого набора параметров (по крайней мере, если ввести параметризацию для ε_t^2 типа ARCH), не содержащих условных матожиданий.

4 Нелинейные модели

Всю представленную выше теорию с определенными модификациями можно применять и к нелинейным моделям регрессионного типа

$$y = m(x, \beta) + e, \quad \mathbb{E}[e|z] = 0,$$

а также к еще более общим моделям

$$\mathbb{E}[m(x, \beta)|z] = 0.$$

Разумеется, при этом после расчета оптимальных или «почти оптимальных» инструментов необходимо применять численную минимизацию ОММ-критерия, что в общем-то приходится делать независимо от выбора инструментов.

Более того, на этапе построения и расчета оптимальных инструментов необходим дополнительный этап получения предварительных состоятельных оценок параметров β. Причина в том, что объект, выполняющий в нелинейной модели функцию регрессора x – это так называемый *квазирегрессор* $\partial m(x, b)/\partial b$, оцененный в точке $b = \beta$ (в линейном случае $m(x, \beta) = x'\beta$, и квазирегрессором является сам регрессор). Дополнительный этап получения предварительных оценок обуславливается как раз зависимостью квазирегрессора от неизвестного параметра. Но состоятельную оценку обычно получить относительно легко (например, прогнать ОММ с неоптимально подобранными инструментами, скажем, с базовыми). Получив состоятельную оценку для β и подставив ее в выражение для квазирегрессора, следует оцененный квазирегрессор и использовать вместо x во всех формулах предыдущих разделов. В разделах 6–7 приводится последовательность действий для двух актуальных нелинейных моделей.

5 Выводы для практической работы

Что полезного может вынести эконометрист-практик из представленной выше теории оптимальных инструментов? Очевидно, в сложных задачах использование теории «по полной программе» затруднительно ввиду сложности конструируемых объектов даже на популяционном уровне, а тем более из-за трудностей в реализации в конкретной выборке. Тем не менее, некоторые упрощенные версии оптимальных или «почти оптимальных» инструментов заслуживают пристального внимания. В некоторых ситуациях промежуточной целью может служить построение «хороших», то есть сильных инструментов, с целью избежать проблем, связанных с использованием слабых инструментов (см. Цыплаков, 2007, раздел 6). В любом случае теорию полезно знать для лучшей ориентации во многих практических ситуациях.

Например, пусть есть основания считать, что в конкретной одноопериодной (или кросс-секционной) задаче очень сильная условная гетероскедастичность. Значит, в такой ситуации использование оптимального инструмента должно дать большую отдачу, чем в случае слабой гетероскедастичности, при которой идея оптимального инструментирования не стоит выделки. Другая ситуация – известно, что ошибка в задаче сильно автокоррелирована, а условная гетероскедастичность, наоборот, слабая. В таком случае мудрым решением будет построить инструмент, который был бы оптимальным в отсутствии гетероскедастичности вообще, ибо пренебрежение гетероскедастичностью существенно упрощает процесс реализации оптимального инструмента. Конечно, такой инструмент не будет исконно оптимальным, ибо мы используем неверную посылку при его построении, но он несильно, скорей всего, будет отличаться от исконно оптимального, если условная гетероскедастичность и впрямь слабая.

Далее мы разбираем две актуальные модели, для которых применяем теорию оптимальных инструментов и высказанные выше соображения для построения если не совсем оптимальных, то по крайней мере удачных инструментов. Один из примеров, попроще, – кросс-секционный; другой, посложнее, – на временных рядах.

6 Пример: нелинейная функция потребления

Обратимся к задаче оценивания нелинейной функции потребления:

$$C = \mu + \delta Y^{\gamma} + \varepsilon,$$

где C – потребление, Y – доход, а ε – ошибка со свойством $\mathbb{E}[\varepsilon|Y] = 0$. Подразумевается наличие случайной выборки. Вектором истинных параметров является $\beta = (\mu, \delta, \gamma)'$. В данной ситуации входящие в регрессионную функцию переменные экзогенны, поэтому, скажем, НЛМНК, который неявно подразумевает в качестве инструмента квазирегрессор $(1, Y^{\gamma}, \delta Y^{\gamma}\ln(Y))'$, дает состоятельные и асимптотически нормальные оценки. Их-то и можно использовать в качестве предварительных; назовем их $\hat{\beta}_0 = (\hat{\mu}_0, \hat{\delta}_0, \hat{\gamma}_0)'$.

В оцениваемом уравнении наверняка имеется сильная гетероскедастичность, например, из-за эффекта размера. Оптимальный инструмент согласно разделу 2 равен отношению квазирегрессора к скедастичной функции. Поэтому напрашивается параметризация скедастичной регрессии, например, как

$$(C - \hat{\mu}_0 - \hat{\delta}_0 Y^{\hat{\gamma}_0})^2 = \kappa + \lambda Y^{\theta} + \eta,$$

и оценивание ее с помощью НЛМНК. Расчетные значения этой скедастичной регрессии можно считать хорошим приближением для значений скедастичной функции. Возможно, что вышеуказанная параметризация избыточна, и простая квадратичная функция хорошо бы приблизила истинную скедастичную функцию; тогда бы для оценивания скедастичной регрессии хватило бы обычного МНК. Впрочем, о качестве параметрической спецификации лучше судить по обычной регрессионной диагностике. Если же данных много, то скедастичную регрессию лучше оценить непараметрически, тем более что это несложно сделать ввиду наличия всего одного регрессора; в таком случае конструируемые инструменты претендуют на звание оптимальных. Мы же, параметризуя скедастичную регрессию, получим в конечном счете «почти оптимальные» инструменты.

Итак, оценив квазирегрессор, используя предварительные НЛМНК-оценки, и оценив скедастичную функцию, мы делим первое на второе и получаем серию оптимальных или «почти оптимальных» инструментов, после чего применяем обычный нелинейный метод инструментальных переменных (см. Цыплаков, 2007, стр. 36). Внимательный читатель должен заметить, что описанные в этом примере действия – это фактически алгоритм обобщенного нелинейного метода наименьших квадратов. Это не совпадение, а следствие кросс-секционного контекста задачи и отсутствия эндогенности в регрессорах.

7 Пример: высокочастотные межтрансакционные доходности

Второй пример касается моделирования высокочастотных доходностей на финансовом рынке. Пусть $t_0 < t_1 < ... < t_i < ...$ – моменты осуществления финансовых трансакций, t_i для i-ой трансакции. Величина $d_i = t_i - t_{i-1}$ назывется i-ой дюрацией. Далее, если за p_{t_i} обозначить цену финансового актива в момент i-ой трансакции, то доходность этой трансакции будет $R_i = \ln(p_{t_i}) - \ln(p_{t_{i-1}})$. Введем понятие нормированной (относительно дюрации) доходности $r_i = R_i/\sqrt{d_i}$. Рассмотрим недавно предложенную модель Meddahi, Renault & Werker (2006):

$$\mathbb{E}\left[(r_i^2 - \lambda) - (r_{i-1}^2 - \lambda)\exp(-\kappa d_{i-1})\frac{v(\kappa d_i)}{v(\kappa d_{i-1})}|\Im_{i-2}\right] = 0,$$

где $v(x) = (1 - \exp(-x))/x$ и $\Im_i = \{d_j, r_j, j \leq i\}$. Вектором истинных параметров является $\beta = (\lambda, \kappa)'$.

Понятно, что здесь можно найти уйму годных инструментов, но вопрос в том, как построить если не оптимальные, то хотя бы очень хорошие инструменты. В данной задаче мы имеем серийную корреляцию в ошибках (задача двухпериодная по построению) и наверняка сильную условную гетероскедастичность (присущую любым моделям для высокочастотных финансовых данных). Как мы знаем из раздела 3, явная формула для оптимального инструмента не выведена. Поэтому мы не рассчитываем на оптимальное инструментирование, но попробуем приблизиться к нему, привлекая соображения из раздела 5. Разумной выглядит следующая последовательность действий.

- Выбираем небольшой вектор инструментов, измеримых относительно $\Im_{i-2}$, например $(1, \ln(d_{i-2}), \ln(d_{i-3}), r_{i-2}, r_{i-3})'$ (здесь и далее дюрации прологарифмированы с целью симметризации распределения инструментов). Применяем ОММ и получаем предварительные оценки для λ и κ. Используя эти оценки, оцениваем квазирегрессоры $\partial m_i(\beta)/\partial\lambda$ и $\partial m_i(\beta)/\partial\kappa$, где $m_i(\beta)$ – выражение в уравнении модели под знаком условного матожидания. Кроме того, оценим дисперсию $\mathbb{E}[m_i(\beta)^2]$ и автоковариацию первого порядка $\mathbb{E}[m_i(\beta)m_{i-1}(\beta)]$, откуда получим состоятельную оценку для θ, коэффициента в разложении Вольда $m_i(\beta) = \varepsilon_i - \theta\varepsilon_{i-1}$. Последняя оценка пригодится позже.

- Линейно регрессируем оцененные квазирегрессоры на предикторах из $\Im_{i-2}$, а именно, на всяческих функциях от недавних d_{i-2} и r_{i-2}, например, на

$$(1, \ln(d_{i-2}), r_{i-2}, r_{i-2}\ln(d_{i-2}), ..., \ln(d_{i-4}), r_{i-4}, r_{i-4}\ln(d_{i-4}))',$$

  руководствуясь обычными критериями качества регрессионной подгонки (такими как значения F-статистики), при этом ограничивая количество регрессоров; здесь можно проявить незаурядную фантазию. Расчетные значения этих двух регрессий можно использовать как приближения для $\mathbb{E}\left[\partial m_i(\beta)/\partial\lambda|\Im_{i-2}\right]$ и $\mathbb{E}\left[\partial m_i(\beta)/\partial\kappa|\Im_{i-2}\right]$. Аналогично, используя предварительные оценки, оцениваем условную дисперсию $\mathbb{E}[m_i(\beta)^2|\Im_{i-2}]$ и условную автоковариацию первого порядка $\mathbb{E}[m_i(\beta)m_{i-1}(\beta)|\Im_{i-2}]$, линейно регрессируя оцененные $m_i(\beta)^2$ и $m_i(\beta)m_{i-1}(\beta)$ на тех же (или других, в зависимости от качества подгонки) предикторах и получая расчетные значения.

- Довольно простой, но сильный инструмент можно теперь получить, как в предыдущем примере, разделив оценки проекции квазирегрессоров на оценку условной дисперсии. Этот инструмент статический, и его построение фактически повторяет алгоритм в предыдущем примере, игнорируя серийную корреляцию ошибок.

- Напротив, можно игнорировать условную гетероскедастичность, но учесть серийную корреляцию, что немного более трудоемко. Для этого мы строим проекции не только самого квазирегрессора, но и его J (скажем, 10) будущих значений, и модифицируем формулу, приведенную в разделе 3, следующим образом:

$$\zeta_i = \theta\zeta_{i-1} + \sum_{j=0}^{J}\theta^j\mathbb{E}\left[\frac{\partial m_{i+j}(\beta)}{\partial\beta}|\Im_{i-2}\right] + \frac{\theta^{J+1}}{1-\theta}\mathbb{E}\left[\frac{\partial m_i(\beta)}{\partial\beta}\right].$$

  Имея оценки для всех задействованных компонент, остается только раскрутить эту рекурсию, начав с первого наблюдения и полагая довыборочные значения всех компонент равными нулю.

- Итак, мы построили два хороших, хотя и не оптимальных инструмента. Заметим, что каждый из них состоит из двух компонент, т.к. квазирегрессор состоит из двух компонент. Для максимально выгодного использования построенных инструментов необходимо теперь прогнать ОММ со сверхидентифицирующим инструментом, состоящим, таким образом, из четырех компонент.

Результирующий инструмент, строго говоря, не являемся оптимальным. Настоящий оптимальный инструмент, как говорилось в разделе 3, имеет очень сложную структуру и задан в неявном виде. В Anatolyev (2007) описано, как построить так называемый «приблизительно оптимальный» инструмент, который будет наверняка лучше построенного по вышеприведенной схеме, хотя вряд ли, по мнению автора, намного.

Список литературы

Цыплаков, А. (2007). Экскурс в мир инструментальных переменных. *Квантиль* 2, 21–47.

Anatolyev, S. (2003). The form of the optimal nonlinear instrument for multiperiod conditional moment restrictions. *Econometric Theory* 19, 602–609.

Anatolyev, S. (2007). Optimal instruments in time series: a survey. *Journal of Economic Surveys* 21, 143–173.

Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics* 34, 305–334.

Hansen, L.P. (1985). A method for calculating bounds on the asymptotic variance-covariance matrices of generalized method of moments estimators. *Journal of Econometrics* 30, 203–228.

Meddahi, N., E. Renault & B. Werker (2006). GARCH and irregularly spaced data. *Economics Letters* 90, 200–204.

Newey, W.K. (1990). Efficient instrumental variables estimation of nonlinear models. *Econometrica* 58, 809–837.

Robinson, P. (1987). Asymptotically efficient estimation in the presence of heteroskedasticity of unknown form. *Econometrica* 55, 875–891.

West, K. D., K.-f. Wong & S. Anatolyev (2002). Instrumental variables estimation of heteroskedastic linear models using all lags of instruments. Working Paper, University of Wisconsin–Madison.

Optimal instruments

Stanislav Anatolyev

New Economic School, Moscow

This essay briefly surveys optimal instrumentation in linear and nonlinear models, both cross-sectional and stationary time series. Examples of judicious construction of instruments are given.