Лекция №13

Прогнозирование по оценённой эконометрической модели и проверка её адекватности

План

1. Завершение темы "Характеристики качества спецификации эконометрических моделей";
2. Оптимальный точечны прогноз и характеристика точности прогноза (стандартная ошибка прогноза);
3. Интервальное прогнозирование по модели и проверка её адекватности;

В конце предыдущей лекции мы обсудили $F-$тест качества спецификации модели. Подчеркнем, что если в итоге $F-$теста принимается гипотеза $H_0: a_1 = a_2 = \ldots = a_k = 0$, то это означает, то ни одна из объясняющих переменных не содержит в себе информацию о эндогенной переменной модели. Если же H_0 отклоняется в пользу гипотезы $H_1 = \overline{H_0}$, это значит, что хотя бы одна экзогенная переменная содержит информацию об эндогенной переменной модели.

Скорректированнный коэффициент детерминации, как интструмент модификации модели

Скоррективонная коэффицент детерминации рассчитывается по следующему правилу:

$$\overline{R}^2 = 1 - \frac{\frac{ESS}{n-(k+1)}}{\frac{TSS}{n-1}}$$

$$\overline{R}^2 \leqslant R^2$$

Числитель и позволяет отбирать в модель объясняющие перменные. Если при включении в модель новой объясняющей переменной велечина $\overline{R}^2$ возрастает (лучше сказать не убывает), то включение этой переменной в модель полезно, если убывает, то бесмысленно.

Замечание. Первое: В числителе вычитаемого $\dfrac{ESS}{n-(k+1)}$ размещается оценка дисперсии случайного возмущения. В знаменателе находится дисперсия эндогенной переменной. При добавлении в модель новой объясняющей переменной меняется только числитель (знаменатель остаётся всегда неизменным) и поэтому увеличение $\overline{R}^2$ равносильно снижению дисперсии случайного возмущения и значит уменьшению экзогенных переменных. Добавим, что всегда имеет место следующее неравенство, при чём $\overline{R}^2$ может быть меньше 0.

Эконометрические модели создаются в частности для прогнозирования неизвестных объясняемых переменных (эндогенных переменных). Поясним на примере базовой модели эконометрики (ЛММР) существо задачи прогнозирования:

$$\begin{cases} y = a_0 + a_1 \cdot x_1 + a_2 \cdot x_2 + \ldots a_k \cdot x_k + u_t \\ E(u_t) = 0; E\left(u_t^2\right) = \sigma_u^2 \end{cases} \quad (1)$$

Пусть при заданных значениях $(1, x_{1,0}, x_{2,0}, \dots, x_{k,0}) = \vec{x}_0^{\,T}$ нужно знать неизвестное значение y_0 эндогенной переменной y. Величина y_0, по предположению, задана экзогенными переменными уравнениям модели (1):

$$y_0 = a_0 \ + a_1 \cdot x_{1,0} + a_2 \cdot x_{2,0} + \ldots a_k \cdot x_{k,0} + u_0 \qquad (2)$$

Обозначим $\tilde{y}_0$ прогноз значения y_0. Прогноз y_0 принято называть оптимальным, если оказываются справедливыми следующие два свойства этого прогноза:

$$\begin{cases} E\left(\tilde{y}_0 - y_0\right) = E\left(\triangle \tilde{y}_0\right) = 0, \\ E\left(\tilde{y}_0 - y_0\right)^2 = \sigma^2\left(\triangle \tilde{y}_0\right) \to \min \end{cases} \qquad (3)$$

Прокомментируем оба требования: первое требование означает равенство 0 математичкого ожидания истинной ошибки прогноза $\left(\tilde{y}_0 - y_0\right)$, несмящённая оценка прогноза, второе требование означает, что среднее растояние между прогнозом и истинной минимально. Подчеркнём, что эти два требования являются аналогом требованиям оптимальности к статистической процедуре оценивания моделей.

Справедлива следующая теорема:

Пусть модель (1) оценивается методом наименьших квадратов по уравнениям наблюдений, где справедливы все предпосылки теоремы Гаусса-Маркова. Тогда оптимальный прогноз рассчитывается по формуле (4):

$$A)\ \tilde{y}_0 = \tilde{a}_0 + \tilde{a}_1 \cdot x_{1,0} + \ldots + \tilde{a}_k \cdot x_{k,0} \qquad (4)$$

Характеристика точности прогноз:

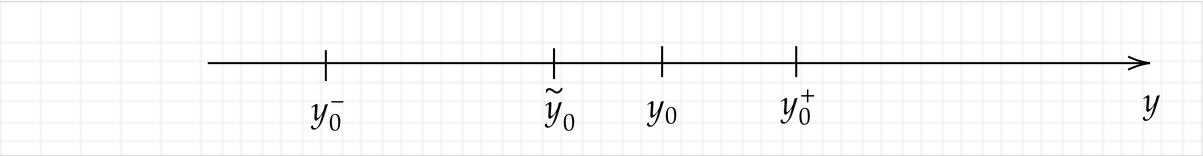
$$\sigma^2\left(\triangle \tilde{y}_0\right) = \sigma_u^2 + \sigma_u^2 \cdot \vec{x}_0^{\,T} \cdot Q \cdot \vec{x}_0 = \sigma_u^2(1 + q_0)$$

Замечание. Присутствие в характеристике точности прогноза слагаемого $\vec{x}_0$ обусловлен случайными ошибками в оценке коэффициентов модели. Первое слагаемое в характеристике обусловлено отсутвием случайного возмущения u_0.

Подводим итог, оптимальный прогноз по модели вычисляется в итоге подстановки в оценку функции регресси заданых значений случайных переменных.

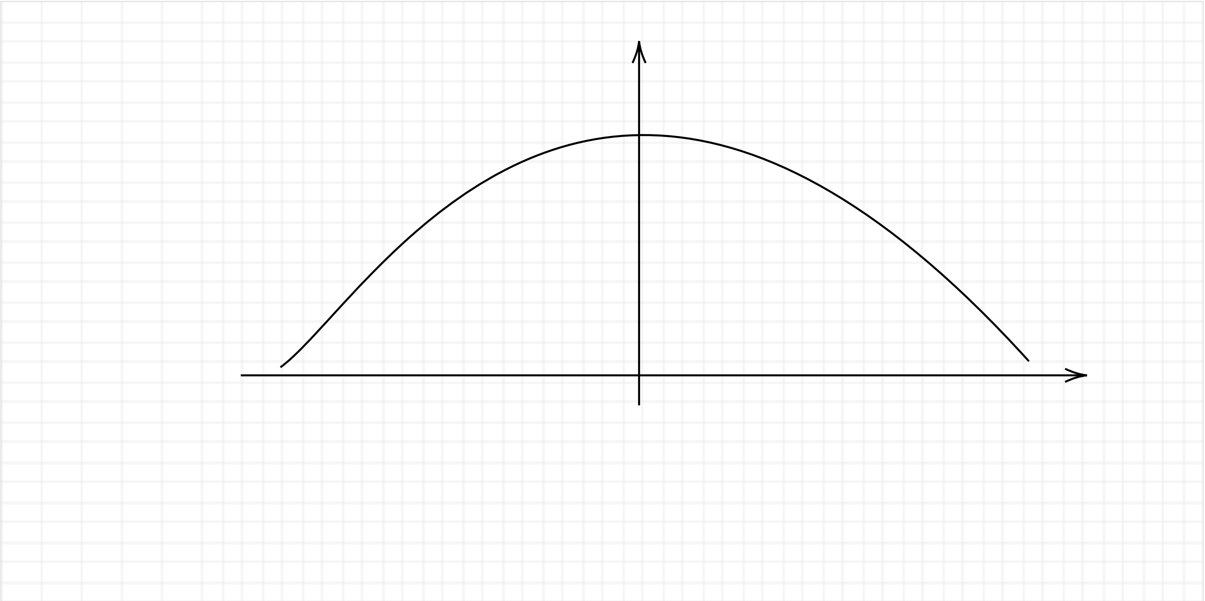
Интервальное прогнозирование по модели и проверка её адекватности

В пукте выше мы обсудили оптимальный прогноз $\tilde{y}_0$.



По мимо точечного прогноза в финансово-экономической сфере прогноз искомой величины y_0 часто строится ввиде интервала с левой границей y_0^- и правой границей y_0^+. Этот интервал накрывает неизвестное значение y_0 с заданной доверительной вероятностью. В основании лежит следующая теорема:

Пусть в моделе (1) выполнены все предпосылки теоремы Гаусса-Маркова и случайные возмущения имеют нормальный закон распределения



Тогда следующая дробь:

$$t = \frac{\widetilde{y}_0 - y_0}{\widetilde{\sigma}\left(\triangle \widetilde{y}_0\right)} \sim t(m),\ \text{где } m = m - k + 1$$

имеет закон распределения Стьюдента с количеством степеней свододы $n - k + 1$. Из теоремы выше вытекает равенство (3):

$$P\left(|\frac{\widetilde{y}_0 - y_0}{\widetilde{\sigma}\left(\triangle \widetilde{y}_0\right)}| \leqslant t_{\text{крит}} \right) = 1 - \alpha \qquad (3)$$

То есть существует такое $t_{\text{крит}}$ при котором справедливо равенство (3). $t_{\text{крит}}$ имеет значение двухсторонней квантили с кол-ом степеней свободы $n - k + 1$. Освобождаясь от модуля мы перепишем формулу в следующем виде:

$$P\left(\widetilde{y}_0 - t_{\text{крит}} \cdot \widetilde{\sigma}\left(\triangle \widetilde{y}_0\right) \leqslant y_0 \leqslant \widetilde{y}_0 + t_{\text{крит}} \cdot \widetilde{\sigma}\left(\triangle \widetilde{y}_0\right)\right) \qquad (3')$$

Из которого и следуют формулы границ доверительно интервала.