

Многомерная авторегрессия

Жолобов В. А.

Zholobov.VA@phystech.edu

Ключевые слова: *многомерная авторегрессия; VAR;*

1 Цель

Прогнозирование временных рядов методом VAR.

2 Описание модели

Многомерная авторегрессия является эконометрической моделью, которую используют для прогнозирования временными рядами. Представляет собой обобщение модели авторегрессии (AR).

$$Y_{1,t} = \alpha_1 + \beta_{11,1}Y_{1,t-1} + \beta_{12,1}Y_{2,t-1} + \varepsilon_{1,t}$$

$$Y_{2,t} = \alpha_2 + \beta_{21,1}Y_{1,t-1} + \beta_{22,1}Y_{2,t-1} + \varepsilon_{2,t}$$

При заданных временном ряде $s_1 = [x_1, \dots, x_{T-1}]^T, x_i \in R^1$ и матрице признаков, столбцами которой являются временные ряды s_2, \dots, s_m . Необходимо спрогнозировать следующую величину x_T ряда s_1 . Предполагается, что

- Ряд не имеет пропущенных значений,
- Отчеты сделаны через равные промежутки времени

Составляется $(m \times k)$ матрица значений временного ряда:

$$S = \left(\begin{array}{c|ccc} x_T & x_{T-1} & \dots & x_{T-k+1} \\ \hline x_{(m-1)k} & x_{(m-1)k-1} & \dots & x_{(m-2)k+1} \\ \dots & \dots & \dots & \dots \\ x_k & x_{k-1} & \dots & x_1 \end{array} \right)$$

В случае, когда учитываются временные ряды s_1, s_2, \dots, s_m , для каждого j -го временного ряда строится авторегрессионная матрица S_j и присоединяется справа. Полученная матрица

$$A = [S_1 \quad S_2 \quad \dots \quad S_m]$$

3 Пути решения задачи

Требуется решить задачу линейной регрессии $\|Xw - y\|^2 \rightarrow \min$ Искомый вектор параметров имеет вид

$$w = (X^T X)^{-1} (X^T y)$$

В терминах линейной регрессии

$$y = Xw$$

Путь к решению задачи: последовательная генерация прогнозируемых значений. В качестве x^T выбираются $K - 1$ предыдущих значений временного ряда, и для каждого следующего значения $x_{T+1}, x_{T+2}, \dots, x_{T+l}$ необходимо заново строить авторегрессионную матрицу X . Результат работы алгоритма представлен в вычислительном эксперименте

4 Функция ошибки

Для измерения точности предсказания используются следующие метрики.
MAPE

$$M = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

MSE

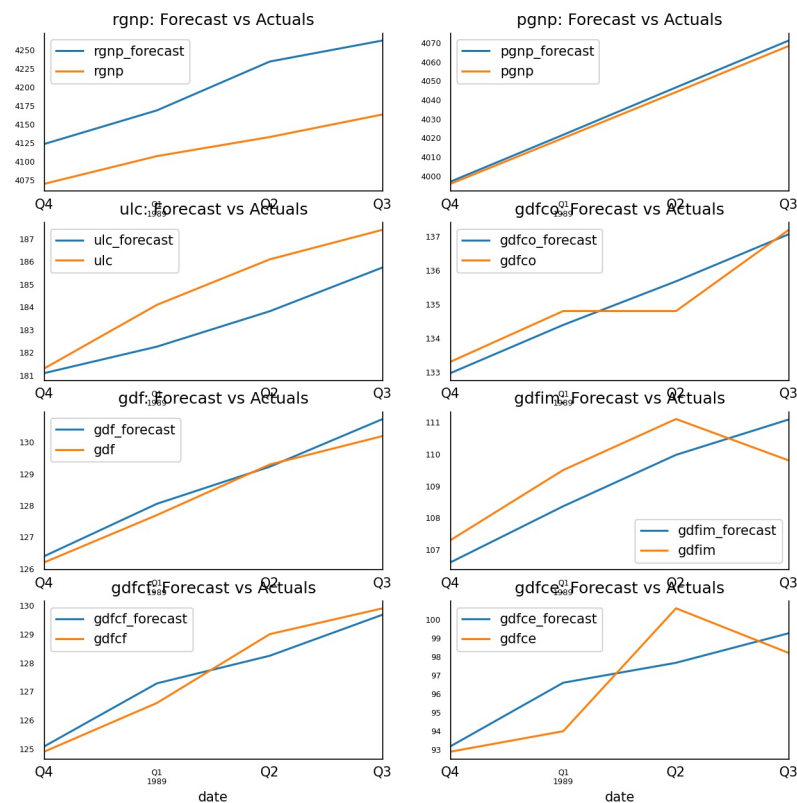
$$\sqrt{\frac{1}{n} \sum_{t=1}^n (A_t - F_t)^2}$$

5 Вычислительный эксперимент

Используется выборка из статьи [1].

	rgnp	pgnp	ulc	gdfco	gdf	gdfim	gdfcf	gdfce
Count	123	123	123	123	123	123	123	123
Mean	2747.09	2758.46	98.28	69.46	68.49	59.39	66.26	52.90
Std	702.44	735.02	47.74	31.72	30.18	31.08	31.64	31.60

Вначале проводим тест Грейнджера и убеждаемся в том, что каждый признак не зависит от другого, так как p_value достаточно высок.



	rgnp	pgnp	ulc	gdfco	gdf	gdfim	gdfcf	gdfce
MAPE	0.0192	0.0005	0.0081	0.0033	0.0023	0.0097	0.0036	0.0177
MSE	82.0245	2.146	1.6856	0.5169	0.3392	1.0826	0.5286	2.034

Видим, что в целом восстанавливаются значения достаточно точно.

Литература

- [1] Yash P Mehra. Wage growth and the inflation process: an empirical approach. In *Cointegration*, pages 147–159. Springer, 1994.

Поступила в редакцию