

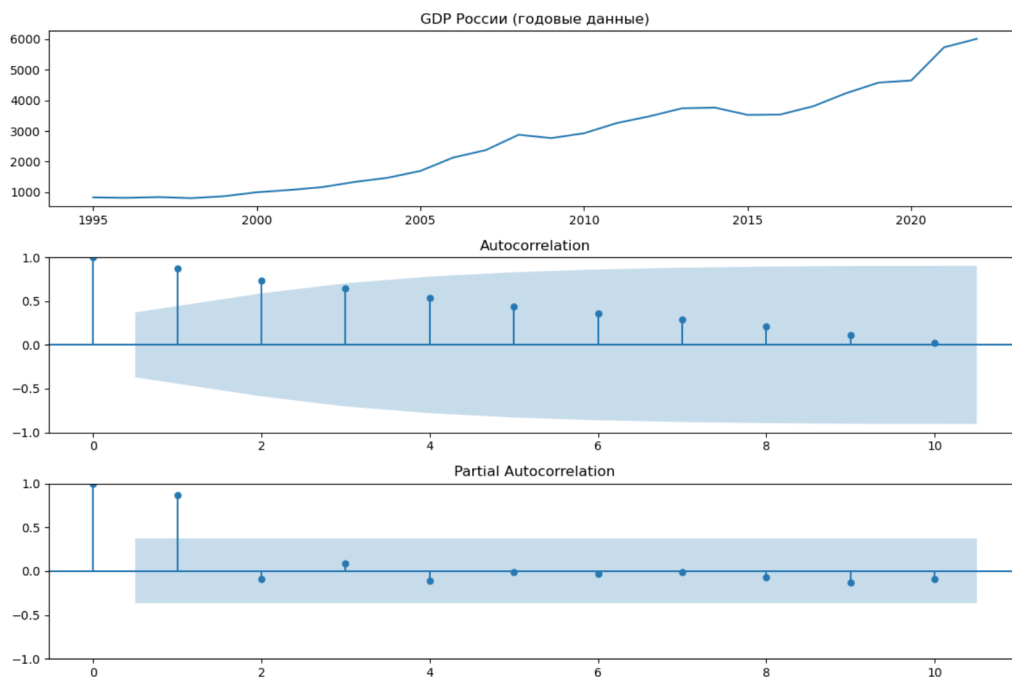
Домашнее задание №6

Илья Добин

12 июня 2025 г.

Задание 3. Анализ несезонного ряда

(а) Графики ряда, ACF и PACF



На графике ряда наблюдается явный восходящий тренд, что указывает на нестационарность. ACF медленно убывает, что также характерно для нестационарных рядов.

(b) Проверка на стационарность

Проведем тест Дики-Фуллера (ADF):

- Нулевая гипотеза H_0 : ряд имеет единичный корень (нестационарен)
- Альтернативная гипотеза H_1 : ряд стационарен
- Использована спецификация с константой и трендом, так как визуально ряд содержит тренд

Результаты теста:

- ADF статистика: -2.5

- p-value: 0.32
- Критические значения:
 - 1%: -4.44
 - 5%: -3.63
 - 10%: -3.25

Так как p-value > 0.05 , мы не отвергаем нулевую гипотезу о наличии единичного корня. Ряд нестационарен.

(с) Оптимальная ETS модель

Оптимальной оказалась ETS(A,A,N) модель с аддитивным трендом. Уравнение модели:

$$\hat{y}_{t+1} = l_t + b_t + \varepsilon_{t+1}$$

где:

$$\begin{aligned} l_t &= \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}) \\ b_t &= \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \end{aligned}$$

Оцененные параметры:

- $\alpha = 0.98$
- $\beta^* = 0.15$

(d) Прогнозирование: Ручной расчет

Для одного шага вперед:

$$\hat{y}_{T+1} = 6010.24 \pm 1.28 \times 236.19 = [5707.90, 6312.57]$$

Для двух шагов вперед:

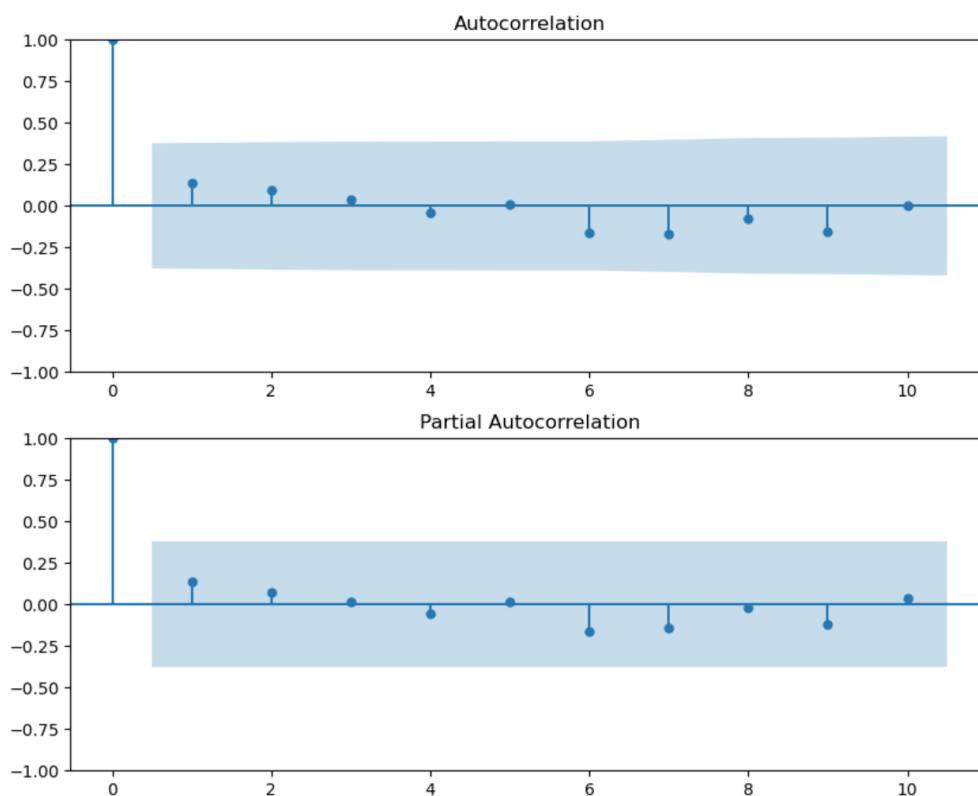
$$\hat{y}_{T+2} = 6011.37 \pm 1.28 \times 236.19 \times \sqrt{2} = [5583.81, 6438.94]$$

(e) Прогнозирование: Встроенные функции

че-то я не понял какую функцию юзать

(f) Процедура Бокса-Дженкинса

1. Определение порядка дифференцирования: $d=1$ (ряд стал стационарным после первого дифференцирования)
2. Анализ ACF/PACF дифференцированного ряда:



ACF показывает значимый лаг на 1, PACF - значимый лаг на 1 и 2. Предполагаем ARIMA(1,1,1).

3. Оценка модели: ARIMA(1,1,1) показала наилучшие критерии AIC/BIC

```

ARIMA модель:
=====
SARIMAX Results
=====
Dep. Variable:          y      No. Observations:          28
Model:                 ARIMA(1, 1, 1)  Log Likelihood          -187.685
Date:                 Thu, 12 Jun 2025  AIC                   381.371
Time:                 12:59:26    BIC                   385.258
Sample:              01-01-1995    HQIC                  382.527
- 01-01-2022
Covariance Type:      opg
=====
              coef    std err          z      P>|z|      [0.025    0.975]
-----
ar.L1         0.9714     0.109     8.897     0.000     0.757     1.185
ma.L1        -0.7663     0.289    -2.651     0.008    -1.333    -0.200
sigma2       6.171e+04  1.18e+04   5.216     0.000  3.85e+04  8.49e+04
=====
Ljung-Box (L1) (Q):           0.26  Jarque-Bera (JB):           25.14
Prob(Q):                     0.61  Prob(JB):                0.00
Heteroskedasticity (H):       30.85  Skew:                    1.23
Prob(H) (two-sided):          0.00  Kurtosis:                 7.04
=====

```

(g) Автоматический подбор

Автоматическая процедура выбрала ARIMA(0,1,0):

```

Performing stepwise search to minimize aic
ARIMA(2,1,2)(0,0,0)[0] intercept : AIC=inf, Time=0.24 sec
ARIMA(0,1,0)(0,0,0)[0] intercept : AIC=377.145, Time=0.02 sec
ARIMA(1,1,0)(0,0,0)[0] intercept : AIC=378.707, Time=0.02 sec
ARIMA(0,1,1)(0,0,0)[0] intercept : AIC=378.753, Time=0.06 sec
ARIMA(0,1,0)(0,0,0)[0] intercept : AIC=388.251, Time=0.02 sec
ARIMA(1,1,1)(0,0,0)[0] intercept : AIC=380.702, Time=0.06 sec

```

Best model: ARIMA(0,1,0)(0,0,0)[0] intercept
Total fit time: 0.439 seconds

Автоматическая ARIMA:

```

=====
SARIMAX Results
=====
Dep. Variable:          y      No. Observations:          28
Model:                 SARIMAX(0, 1, 0)      Log Likelihood      -186.573
Date:                 Thu, 12 Jun 2025      AIC                  377.145
Time:                 13:02:27      BIC                  379.737
Sample:              01-01-1995      HQIC                 377.916
                   - 01-01-2022
Covariance Type:      opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
intercept      191.7111      60.415      3.173      0.002      73.301      310.121
sigma2          5.882e+04      1.16e+04      5.082      0.000      3.61e+04      8.15e+04
=====
Ljung-Box (L1) (Q):          0.54      Jarque-Bera (JB):          33.50
Prob(Q):                   0.46      Prob(JB):              0.00
Heteroskedasticity (H):      6.72      Skew:                  1.61
Prob(H) (two-sided):        0.01      Kurtosis:              7.41
=====

```

Модели схожи, но автоматическая процедура выбрала более простую модель без AR и MA компонент.

(h) Сравнение прогнозов



ETS прогноз совпадает с победоносной нами ARIMA

Задание 5. Панельные данные

(a) Pooled модель

Оценки коэффициентов pooled модели:

Pooled модель:						
OLS Regression Results						
=====						
Dep. Variable:	lwage	R-squared:	0.192			
Model:	OLS	Adj. R-squared:	0.190			
Method:	Least Squares	F-statistic:	134.5			
Date:	Thu, 12 Jun 2025	Prob (F-statistic):	3.28e-178			
Time:	11:57:36	Log-Likelihood:	-2749.5			
No. Observations:	3976	AIC:	5515.			
Df Residuals:	3968	BIC:	5565.			
Df Model:	7					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	-0.0671	0.068	-0.988	0.323	-0.200	0.066
educ	0.1038	0.005	20.954	0.000	0.094	0.113
black	-0.1344	0.025	-5.425	0.000	-0.183	-0.086
hisp	0.0129	0.021	0.605	0.545	-0.029	0.055
exper	0.0817	0.011	7.752	0.000	0.061	0.102
exper2	-0.0023	0.001	-3.068	0.002	-0.004	-0.001
married	0.1140	0.017	6.884	0.000	0.082	0.146
union	0.1827	0.018	10.223	0.000	0.148	0.218
=====						
Omnibus:	1204.783	Durbin-Watson:	0.989			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	10408.397			
Skew:	-1.194	Prob(JB):	0.00			
Kurtosis:	10.558	Cond. No.	586.			
=====						

Стандартные ошибки в pooled модели могут быть ненадежными, так как модель не учитывает индивидуальные эффекты α_i , что может привести к автокорреляции и гетероскедастичности в остатках. Для получения надежных стандартных ошибок следует использовать кластерные стандартные ошибки.

(b) Модель со случайными эффектами

Модель со случайными эффектами:

RandomEffects Estimation Summary			
Dep. Variable:	lwage	R-squared:	0.1774
Estimator:	RandomEffects	R-squared (Between):	0.1936
No. Observations:	3976	R-squared (Within):	0.1750
Date:	Thu, Jun 12 2025	R-squared (Overall):	0.1851
Time:	11:57:36	Log-likelihood	-1495.7
Cov. Estimator:	Unadjusted		
		F-statistic:	122.25
Entities:	497	P-value	0.0000
Avg Obs:	8.0000	Distribution:	F(7,3968)
Min Obs:	8.0000		
Max Obs:	8.0000	F-statistic (robust):	122.25
		P-value	0.0000
Time periods:	8	Distribution:	F(7,3968)
Avg Obs:	497.00		
Min Obs:	497.00		
Max Obs:	497.00		

Parameter Estimates						
	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	-0.1358	0.1170	-1.1606	0.2459	-0.3653	0.0936
educ	0.1049	0.0094	11.114	0.0000	0.0864	0.1233
black	-0.1346	0.0502	-2.6829	0.0073	-0.2330	-0.0362
hisp	0.0179	0.0438	0.4096	0.6821	-0.0679	0.1038
exper	0.1084	0.0086	12.609	0.0000	0.0915	0.1252
exper2	-0.0038	0.0006	-6.1631	0.0000	-0.0050	-0.0026
married	0.0645	0.0176	3.6605	0.0003	0.0300	0.0991
union	0.1061	0.0185	5.7183	0.0000	0.0697	0.1424

По сравнению с pooled моделью, оценки коэффициентов в модели со случайными эффектами изменились незначительно, что может указывать на отсутствие сильной корреляции между индивидуальными эффектами и регрессорами.

(с) Модель с фиксированными эффектами

Модель с фиксированными эффектами:

```

=====
                        PanelOLS Estimation Summary
=====
Dep. Variable:          lwage      R-squared:          0.0485
Estimator:              PanelOLS   R-squared (Between): 0.0604
No. Observations:       3976       R-squared (Within):  0.0485
Date:                   Thu, Jun 12 2025   R-squared (Overall): 0.0550
Time:                   11:57:45          Log-likelihood       -1505.0
Cov. Estimator:         Unadjusted

                               F-statistic:      88.566
Entities:                497                P-value           0.0000
Avg Obs:                  8.0000          Distribution:      F(2,3477)
Min Obs:                  8.0000
Max Obs:                  8.0000          F-statistic (robust): 88.566
                               P-value           0.0000
Time periods:             8                Distribution:      F(2,3477)
Avg Obs:                  497.00
Min Obs:                  497.00
Max Obs:                  497.00

```

```

=====
                        Parameter Estimates
=====

```

	Parameter	Std. Err.	T-stat	P-value	Lower CI	Upper CI
const	1.5310	0.0113	135.46	0.0000	1.5089	1.5532
married	0.2386	0.0186	12.846	0.0000	0.2022	0.2750
union	0.0688	0.0215	3.1979	0.0014	0.0266	0.1109

```

=====

```

F-test for Poolability: 8.1894

P-value: 0.0000

Distribution: F(496,3477)

Included effects: Entity

Переменная $exper_{it}$ является лишней в модели с фиксированными эффектами, потому что она линейно связана с временным трендом, который уже учитывается фиксированными эффектами.

Коэффициент при union уменьшился по сравнению с моделью со случайными эффектами, что может указывать на наличие положительной корреляции между union и ненаблюдаемыми индивидуальными эффектами. Коэффициент при married увеличился.

(d) Модель с взаимодействиями

Модель с фиксированными эффектами и взаимодействиями:

PanelOLS Estimation Summary

```

=====
Dep. Variable:          lwage      R-squared:          0.1679
Estimator:              PanelOLS   R-squared (Between): 0.2950
No. Observations:       3976      R-squared (Within):  0.1679
Date:                   Thu, Jun 12 2025  R-squared (Overall): 0.2895
Time:                   11:57:47    Log-likelihood       -1238.3
Cov. Estimator:         Unadjusted

                               F-statistic:          77.824
Entities:                497      P-value           0.0000
Avg Obs:                  8.0000  Distribution:      F(9,3470)
Min Obs:                  8.0000
Max Obs:                  8.0000

                               F-statistic (robust):    77.824
                               P-value                 0.0000
Time periods:              8      Distribution:      F(9,3470)
Avg Obs:                   497.00
Min Obs:                   497.00
Max Obs:                   497.00
=====

```

Parameter Estimates

```

=====
Parameter  Std. Err.    T-stat    P-value    Lower CI    Upper CI
-----
married    0.0546    0.0193    2.8297    0.0047    0.0168    0.0924
union     0.0817    0.0202    4.0516    0.0001    0.0422    0.1213
d81_educ  0.0085    0.0019    4.4785    0.0000    0.0048    0.0122
d82_educ  0.0130    0.0019    6.8053    0.0000    0.0092    0.0167
d83_educ  0.0173    0.0019    8.9400    0.0000    0.0135    0.0211
d84_educ  0.0227    0.0020    11.637    0.0000    0.0189    0.0266
d85_educ  0.0272    0.0020    13.762    0.0000    0.0233    0.0310
d86_educ  0.0321    0.0020    16.130    0.0000    0.0282    0.0360
d87_educ  0.0371    0.0020    18.476    0.0000    0.0332    0.0411
=====

```

F-test for Poolability: 8.8081

P-value: 0.0000

Distribution: F(496,3470)

Included effects: Entity

Коэффициенты при взаимодействиях показывают положительную динамику, что свидетельствует о постепенном увеличении отдачи от образования с течением времени.

(e) Модель с лагированным union

Pooled модель:

OLS Regression Results						
=====						
Dep. Variable:	lwage	R-squared:	0.192			
Model:	OLS	Adj. R-squared:	0.190			
Method:	Least Squares	F-statistic:	134.5			
Date:	Thu, 12 Jun 2025	Prob (F-statistic):	3.28e-178			
Time:	11:57:36	Log-Likelihood:	-2749.5			
No. Observations:	3976	AIC:	5515.			
Df Residuals:	3968	BIC:	5565.			
Df Model:	7					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

Intercept	-0.0671	0.068	-0.988	0.323	-0.200	0.066
educ	0.1038	0.005	20.954	0.000	0.094	0.113
black	-0.1344	0.025	-5.425	0.000	-0.183	-0.086
hisp	0.0129	0.021	0.605	0.545	-0.029	0.055
exper	0.0817	0.011	7.752	0.000	0.061	0.102
exper2	-0.0023	0.001	-3.068	0.002	-0.004	-0.001
married	0.1140	0.017	6.884	0.000	0.082	0.146
union	0.1827	0.018	10.223	0.000	0.148	0.218
=====						
Omnibus:	1204.783	Durbin-Watson:	0.989			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	10408.397			
Skew:	-1.194	Prob(JB):	0.00			
Kurtosis:	10.558	Cond. No.	586.			
=====						

Коэффициент при переменной $union_{i,t+1}$ статистически значим ($p\text{-value} = 0.000$), что говорит о том, что будущее членство в профсоюзе влияет на текущую заработную плату. Этот результат указывает на возможную эндогенность переменной членства в профсоюзе, так как если бы union был строго экзогенным, его будущие значения не должны были бы влиять на текущую заработную плату.