Lab 2 econometria CA 16/03/2020

Aggiungo le dummy periodiche. Destagionalizziamo p, lo facciamo regredendo yt sulle dummy. Eliminiamo quella variabilità della serie che dipendono solo dalle fluttuazioni stagionali.

Destagionalizzazione

↓

$$y_t = \sum_{i=1}^{12} \beta_i d_{i,t} + \varepsilon_t$$

$$d_{i,t} = \begin{cases} 1 \text{ se } t \text{ corrisponde al mese } i \\ 0 \text{ altrimenti} \end{cases}$$

1) stimare $\beta_i, i = 1, \ldots, 12$ con OLS, ottenendo $\hat{\beta}_i$;

2) serie destagionalizzata $\hat{\varepsilon}_t$:

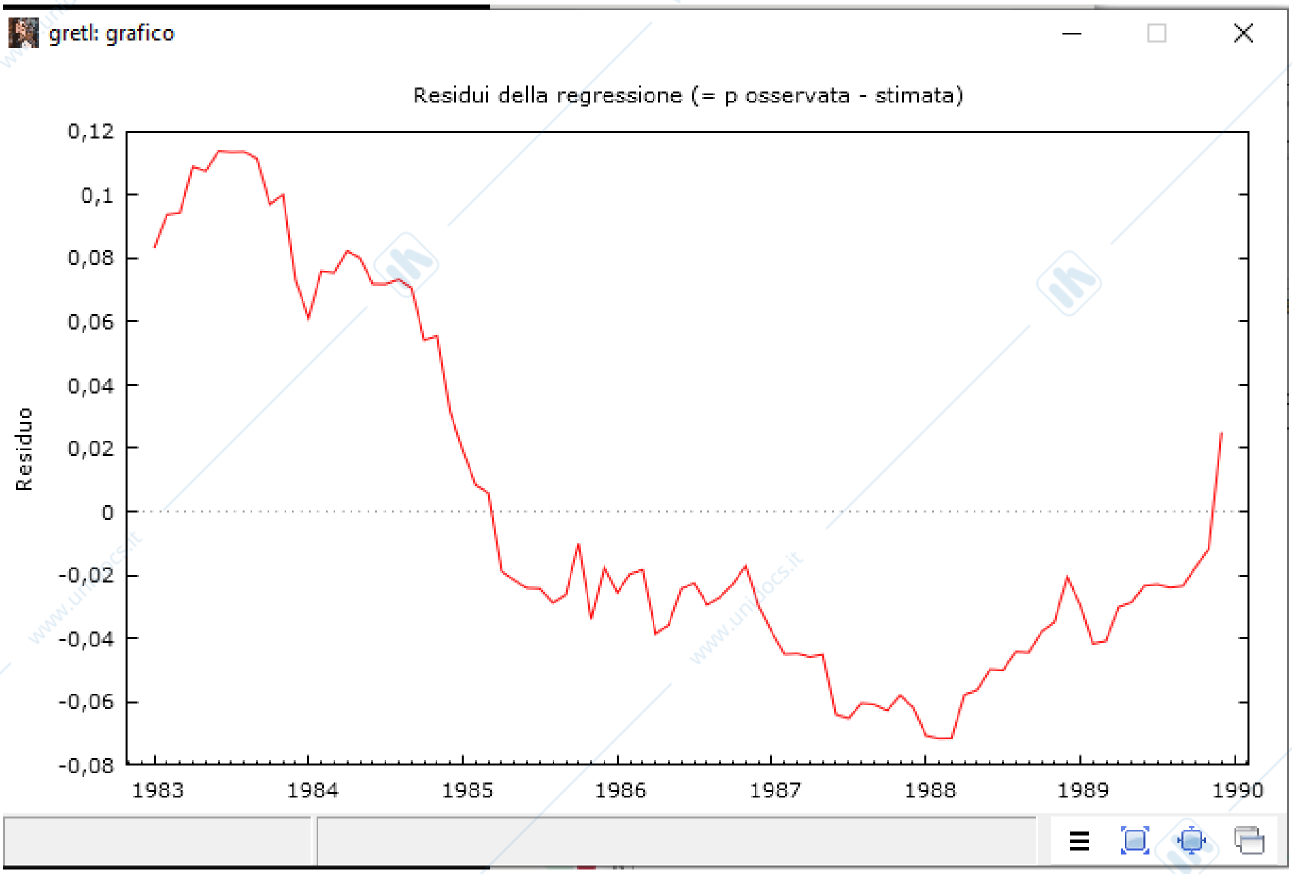
$$\hat{\varepsilon}_t = y_t - \sum_{i=1}^{12} \hat{\beta}_i d_{i,t}$$

1)

Modello 1: OLS, usando le osservazioni 1983:01-1989:12 (T = 84)
Variabile dipendente: p

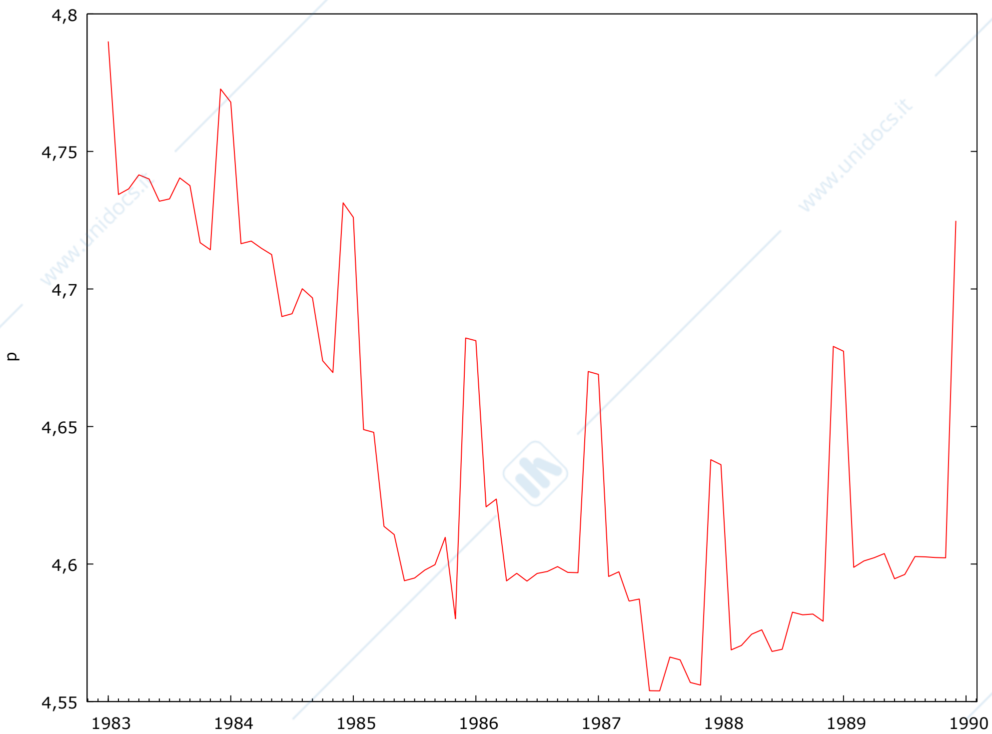
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
dm1	4,70680	0,0233112	201,9	<0,0001	***
dm2	4,64051	0,0233112	199,1	<0,0001	***
dm3	4,64201	0,0233112	199,1	<0,0001	***
dm4	4,63245	0,0233112	198,7	<0,0001	***
dm5	4,63242	0,0233112	198,7	<0,0001	***
dm6	4,61807	0,0233112	198,1	<0,0001	***
dm7	4,61920	0,0233112	198,2	<0,0001	***
dm8	4,62671	0,0233112	198,5	<0,0001	***
dm9	4,62607	0,0233112	198,4	<0,0001	***
dm10	4,61979	0,0233112	198,2	<0,0001	***
dm11	4,61405	0,0233112	197,9	<0,0001	***
dm12	4,69972	0,0233112	201,6	<0,0001	***

Media var. dipendente	4,639817	SQM var. dipendente	0,064702
Somma quadr. residui	0,273879	E.S. della regressione	0,061676
R-quadro	0,211785	R-quadro corretto	0,091364
F(11, 72)	1,758698	P-value(F)	0,077731
Log-verosimiglianza	121,2964	Criterio di Akaike	−218,5928
Criterio di Schwarz	−189,4230	Hannan-Quinn	−206,8668
rho	0,972812	Durbin-Watson	0,031265



serie storica destagionalizzata

Seria storica originale



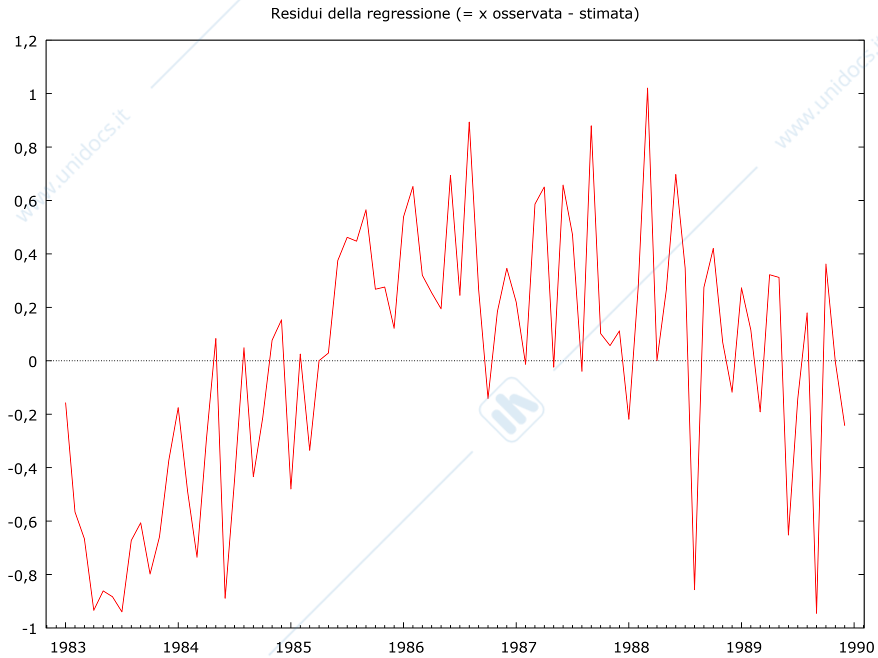
Destagionalizzazione per x

Modello 2: OLS, usando le osservazioni 1983:01-1989:12 (T = 84)
Variabile dipendente: x

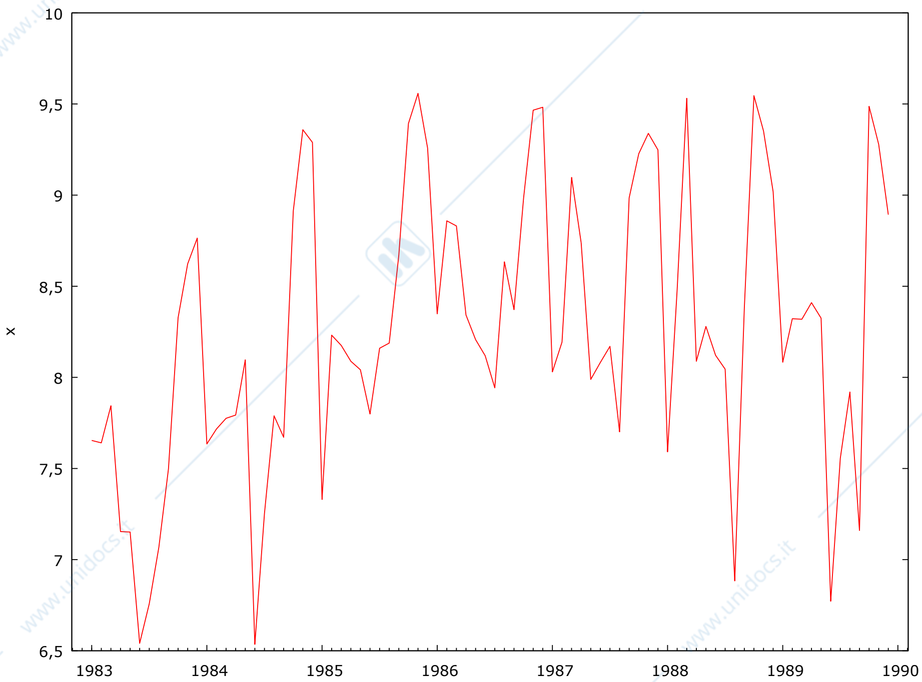
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
dm1	7,81039	0,194926	40,07	<0,0001	***
dm2	8,20645	0,194926	42,10	<0,0001	***
dm3	8,51039	0,194926	43,66	<0,0001	***
dm4	8,08781	0,194926	41,49	<0,0001	***
dm5	8,01281	0,194926	41,11	<0,0001	***
dm6	7,42423	0,194926	38,09	<0,0001	***
dm7	7,69787	0,194926	39,49	<0,0001	***
dm8	7,74052	0,194926	39,71	<0,0001	***
dm9	8,10537	0,194926	41,58	<0,0001	***
dm10	9,12502	0,194926	46,81	<0,0001	***
dm11	9,28205	0,194926	47,62	<0,0001	***
dm12	9,13601	0,194926	46,87	<0,0001	***

Media var. dipendente	8,261577	SQM var. dipendente	0,766298
Somma quadr. residui	19,15010	E.S. della regressione	0,515726
R-quadro	0,607086	R-quadro corretto	0,547057
F(11, 72)	10,11329	P-value(F)	7,61e-11
Log-verosimiglianza	−57,09346	Criterio di Akaike	138,1869
Criterio di Schwarz	167,3567	Hannan-Quinn	149,9129
rho	0,500018	Durbin-Watson	0,998687

Serie destagionalizzata



Serie originale



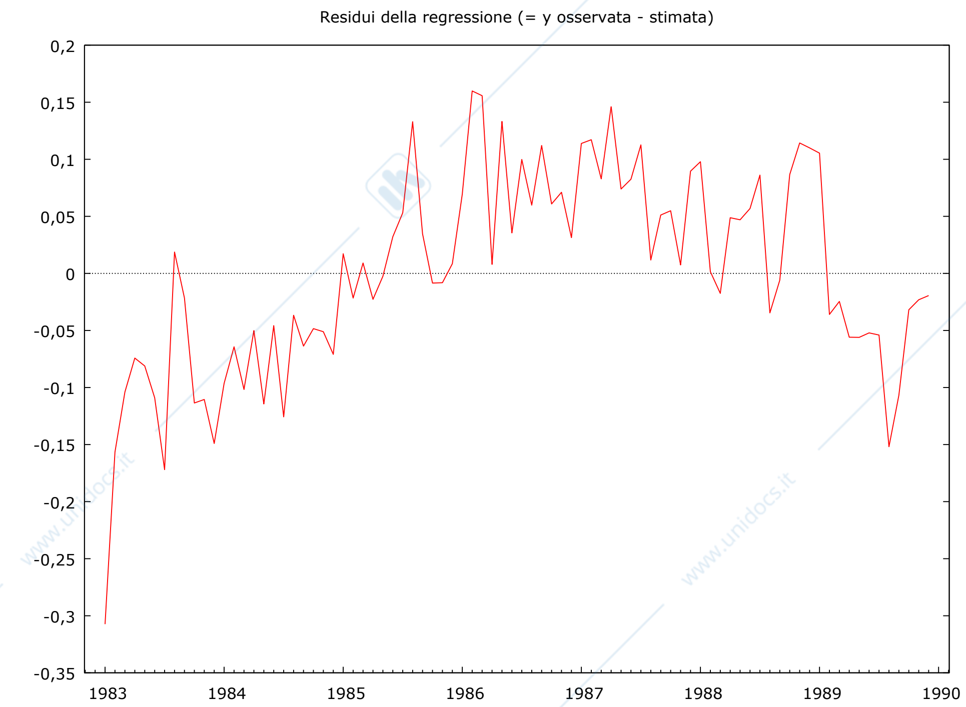
Destagionalizzazione per y

Modello 3: OLS, usando le osservazioni 1983:01-1989:12 (T = 84)
Variabile dipendente: y

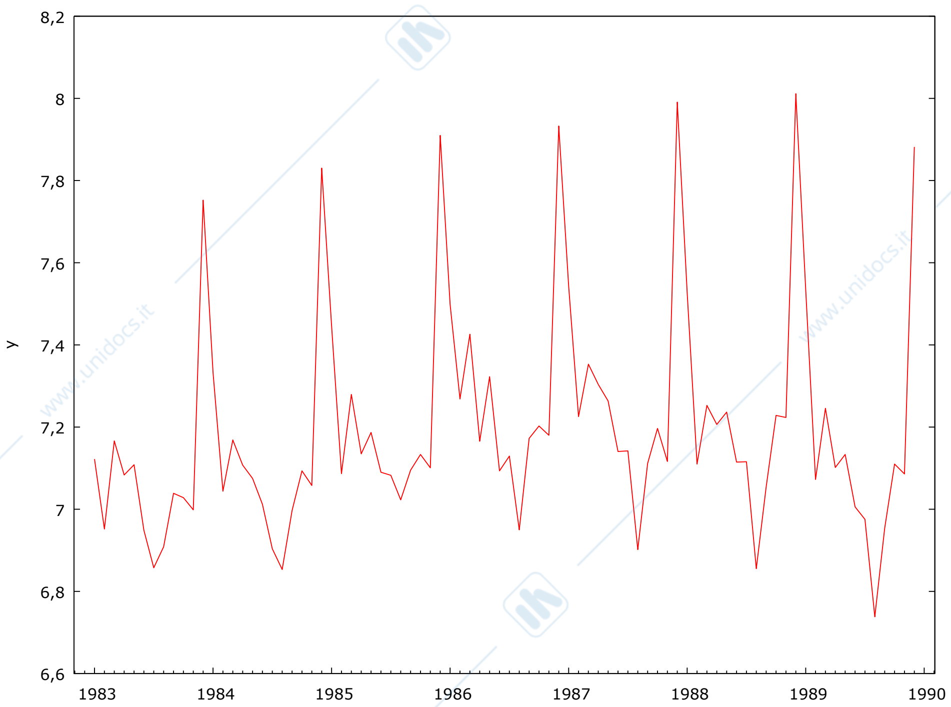
	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
dm1	7,42940	0,0359660	206,6	<0,0001	***
dm2	7,10836	0,0359660	197,6	<0,0001	***
dm3	7,27023	0,0359660	202,1	<0,0001	***
dm4	7,15758	0,0359660	199,0	<0,0001	***
dm5	7,18937	0,0359660	199,9	<0,0001	***
dm6	7,05793	0,0359660	196,2	<0,0001	***
dm7	7,02945	0,0359660	195,4	<0,0001	***
dm8	6,89001	0,0359660	191,6	<0,0001	***
dm9	7,06037	0,0359660	196,3	<0,0001	***
dm10	7,14179	0,0359660	198,6	<0,0001	***
dm11	7,10899	0,0359660	197,7	<0,0001	***
dm12	7,90136	0,0359660	219,7	<0,0001	***

Media var. dipendente	7,195405	SQM var. dipendente	0,264447
Somma quadr. residui	0,651952	E.S. della regressione	0,095157
R-quadro	0,887679	R-quadro corretto	0,870519
F(11, 72)	51,72908	P-value(F)	1,04e-29
Log-verosimiglianza	84,87043	Criterio di Akaike	−145,7409
Criterio di Schwarz	−116,5711	Hannan-Quinn	−134,0148
rho	0,673200	Durbin-Watson	0,508911

Serie destagionalizzata



Serie originale



NB: la costante nella stima si elimina in questo caso perché i regressori devono essere linearmente indipendenti tra loro (la matrice X non ha rango pieno). Gretl elimina di default la dodicesima dummy per poter restituire una stima.

Nota: invece che stimare la regressione di y_t su $d_{1,t},\dots,d_{12,t}$ possiamo introdurre l'intercetta ed eliminare una delle 12 dummy – ad esempio, l'ultima. In questo caso, la regressione stimata è

$$\boxed{y_t = \alpha + \sum_{i=1}^{11} \gamma_i d_{i,t} + \varepsilon_t}$$

Esempio di prima sulla variabile p (cambiano i coefficienti ma le rimanenti quantità sono le stesse)

Modello 4: OLS, usando le osservazioni 1983:01-1989:12 (T = 84)
Variabile dipendente: p

	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	4,69972	0,0233112	201,6	<0,0001	***
dm1	0,00708488	0,0329670	0,2149	0,8304	
dm2	−0,0592034	0,0329670	−1,796	0,0767	*
dm3	−0,0577090	0,0329670	−1,751	0,0843	*
dm4	−0,0672629	0,0329670	−2,040	0,0450	**
dm5	−0,0672914	0,0329670	−2,041	0,0449	**
dm6	−0,0816500	0,0329670	−2,477	0,0156	**
dm7	−0,0805175	0,0329670	−2,442	0,0170	**
dm8	−0,0730033	0,0329670	−2,214	0,0300	**
dm9	−0,0736437	0,0329670	−2,234	0,0286	**
dm10	−0,0799221	0,0329670	−2,424	0,0178	**
dm11	−0,0856664	0,0329670	−2,599	0,0113	**

Media var. dipendente	4,639817	SQM var. dipendente	0,064702
Somma quadr. residui	0,273879	E.S. della regressione	0,061676
R-quadro	0,211785	R-quadro corretto	0,091364
F(11, 72)	1,758698	P-value(F)	0,077731
Log-verosimiglianza	121,2964	Criterio di Akaike	−218,5928
Criterio di Schwarz	−189,4230	Hannan-Quinn	−206,8668
rho	0,972812	Durbin-Watson	0,031265

Stimiamo ora un VAR(2) sui nostri dati:

VAR(2)

Vettore delle variabili

$$\boxed{w_t = \begin{pmatrix} y_t \\ x_t \\ p_t \end{pmatrix}}$$

La dimensione è $m \times 1$ con $m = 3$.

Le componenti deterministiche sono

$$d_t = \begin{pmatrix} 1 \\ d_{1,t} \\ \vdots \\ d_{11,t} \\ t \end{pmatrix}$$

Nota: se dovete inserire d_t nel testo scrivetelo così: $d_t = (1, d_{1,t}, \dots, d_{11,t}, t)'$ e non così: $d_t = \begin{pmatrix} 1 \\ d_{1,t} \\ \vdots \\ d_{11,t} \\ t \end{pmatrix}$.¶

Ci sono le dummy perché abbiamo osservato molta stagionalità nelle serie storiche di partenza

Il nostro modello VAR è

Modello VAR(2) [2 ritardi]:¶

$$\underbrace{w_t}_{3\times1} = \underbrace{A_1}_{3\times3}\ \underbrace{w_{t-1}}_{3\times1} + \underbrace{A_2}_{3\times3}\ \underbrace{w_{t-2}}_{3\times1} + \underbrace{\Phi}_{3\times13}\ \underbrace{d_t}_{13\times1} + \underbrace{\varepsilon_t}_{3\times1}$$

$$E(\varepsilon_t|I_{t-1}) = \underbrace{0}_{3\times1},\ V(\varepsilon_t|I_{t-1}) = \underbrace{\Lambda}_{3\times3}$$

Scriviamolo per esteso:

$$\begin{pmatrix} y_t \\ x_t \\ p_t \end{pmatrix} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ a_{21}^{(1)} & a_{22}^{(1)} & a_{23}^{(1)} \\ a_{31}^{(1)} & a_{32}^{(1)} & a_{33}^{(1)} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \\ p_{t-1} \end{pmatrix} + \begin{pmatrix} a_{11}^{(2)} & a_{12}^{(2)} & a_{13}^{(2)} \\ a_{21}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} \\ a_{31}^{(2)} & a_{32}^{(2)} & a_{33}^{(2)} \end{pmatrix} \begin{pmatrix} y_{t-2} \\ x_{t-2} \\ p_{t-2} \end{pmatrix}$$

$$+ \underbrace{\begin{pmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1,12} & \varphi_{1,13} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2,12} & \varphi_{2,13} \\ \varphi_{31} & \varphi_{32} & \cdots & \varphi_{3,12} & \varphi_{3,13} \end{pmatrix}}_{3\times13} \begin{pmatrix} 1 \\ d_{1,t} \\ \vdots \\ d_{11,t} \\ t \end{pmatrix} + \begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{p,t} \end{pmatrix}$$

$$E\left(\begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{p,t} \end{pmatrix} \middle| I_{t-1} \right) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},\ V\left(\begin{pmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{p,t} \end{pmatrix} \middle| I_{t-1} \right) = \begin{pmatrix} \lambda_{yy} & \lambda_{yx} & \lambda_{yp} \\ \square & \lambda_{xx} & \lambda_{xp} \\ \square & \square & \lambda_{pp} \end{pmatrix}$$

Scriviamo le singole equazioni del VAR:

Scriviamo la prima equazione, quella per y_t:

$$y_t = a_{11}^{(1)}y_{t-1} + a_{12}^{(1)}x_{t-1} + a_{13}^{(1)}p_{t-1} + a_{11}^{(2)}y_{t-2} + a_{12}^{(2)}x_{t-2} + a_{13}^{(2)}p_{t-2}$$
$$+\varphi_{11} + \varphi_{12}d_{1,t} + \ldots \varphi_{1,12}d_{11,t} + \varphi_{1,13}t + \varepsilon_{y,t}$$

Scriviamo la seconda equazione, quella per x_t:

$$x_t = a_{21}^{(1)}y_{t-1} + a_{22}^{(1)}x_{t-1} + a_{23}^{(1)}p_{t-1} + a_{21}^{(2)}y_{t-2} + a_{22}^{(2)}x_{t-2} + a_{23}^{(2)}p_{t-2}$$
$$+\varphi_{21} + \varphi_{22}d_{1,t} + \ldots \varphi_{2,12}d_{11,t} + \varphi_{2,13}t + \varepsilon_{x,t}$$

Scriviamo la terza equazione, quella per p_t:

$$p_t = a_{31}^{(1)}y_{t-1} + a_{32}^{(1)}x_{t-1} + a_{33}^{(1)}p_{t-1} + a_{31}^{(2)}y_{t-2} + a_{32}^{(2)}x_{t-2} + a_{33}^{(2)}p_{t-2}$$
$$+\varphi_{31} + \varphi_{32}d_{1,t} + \ldots \varphi_{3,12}d_{11,t} + \varphi_{3,13}t + \varepsilon_{pt}$$

Aggiungiamo un trend temporale, cioè il nostro t.

LA STIMA DEL VAR(2)

Modello 5: OLS, usando le osservazioni **1983:03**-1989:12 (T = 82)
Variabile dipendente: y

	Coefficiente	*Errore Std.*	*rapporto t*	*p-value*	
const	9,05452	2,19277	4,129	0,0001	***
x_1	0,0300570	0,0189589	1,585	0,1179	
x_2	−0,0107513	0,0190379	−0,5647	0,5743	
p_1	−0,301110	0,724914	−0,4154	0,6793	
p_2	−0,565734	0,746340	−0,7580	0,4513	
dm1	−0,641377	0,113538	−5,649	<0,0001	***
dm2	−0,834247	0,0960879	−8,682	<0,0001	***
dm3	−0,589297	0,0650067	−9,065	<0,0001	***
dm4	−0,767363	0,0453077	−16,94	<0,0001	***
dm5	−0,698367	0,0414234	−16,86	<0,0001	***
dm6	−0,837259	0,0456950	−18,32	<0,0001	***
dm7	−0,814104	0,0513796	−15,84	<0,0001	***
dm8	−0,955709	0,0542223	−17,63	<0,0001	***
dm9	−0,734585	0,0517324	−14,20	<0,0001	***
dm10	−0,699752	0,0500737	−13,97	<0,0001	***
dm11	−0,798368	0,0367339	−21,73	<0,0001	***
time	−0,00127818	0,000543969	−2,350	0,0219	**
y_1	0,306374	0,129016	2,375	0,0206	**
y_2	0,0773214	0,109500	0,7061	0,4827	

Media var. dipendente	7,199270	SQM var. dipendente	0,266166
Somma quadr. residui	0,180842	E.S. della regressione	0,053577
R-quadro	0,968486	R-quadro corretto	0,959482
F(18, 63)	107,5605	P-value(F)	2,71e-40
Log-verosimiglianza	134,4379	Criterio di Akaike	−230,8759
Criterio di Schwarz	−185,1482	Hannan-Quinn	−212,5169
rho	−0,003584	Durbin-Watson	2,002339

Parte da 03:1983 perché abbiamo introdotto due ritardi nel modello. Osservazioni complessive 82. Prima colonna stima dei coefficienti. Seconda colonna Standard Error che dà una misura della precisione dello stimatore (se grande stimatore impreciso, se piccolo stimatore più preciso. Quindi deve essere più piccolo). Lo SE dipende dall'ordine di grandezza.

tecnicamente, nel modello lineare classico stimato con OLS:

$$\hat{\beta} = (X'X)^{-1}X'y$$

$$V(\hat{\beta}) = \sigma^2(X'X)^{-1}$$

$$s(\hat{\beta}_i) = \sqrt{\hat{\sigma}^2 a_{ii}},\ i = 1,\ldots,q$$

$$\hat{\sigma}^2 = \frac{(y - X\hat{\beta})'(y - X\hat{\beta})}{T-q}$$

dove a_{ii} è l'elemento i,i di $(X'X)^{-1}$

Nota: $t(\hat{\beta}_i) := \hat{\beta}_i / s(\hat{\beta}_i)$. Se $\beta_i = 0$, allora (sotto opportune condizioni) allora $t(\hat{\beta}_i) \sim t_{T-q}$ (t di Student con $T-q$ gradi di libertà). I valori critici della t di Student, quando $T-q$ è grande, sono – al livello di significatività del 5%, pari a ± 1.96. Quindi, se $|t(\hat{\beta}_i)| \geq 1.96$ allora si rifiuta l'ipotesi $H_0 : \beta_i = 0$

Alternativamente alla statistica t, possiamo considerare il p-value:

$$p(\hat{\beta}_i) = \Pr\{|t_{T-q}| \geq |t(\hat{\beta}_i)|\}$$

Se p-value è minore o uguale a 0.05 si rifiuta l'ipotesi nulla. (livello di significatività del 5%)