

Εξαγωγή Συντελεστών Παλινδρόμησης με Άλγεβρα Πινάκων

πραγματικοί συντελεστές

πίνακας τιμών
ανεξαρτημένης μεταβλητής
(X_i)

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}$$

πίνακας τιμών
εξαρτημένης μεταβλητής
(Ψ_i)

πολλαπλασιασμός
πινάκων

$$\mathbf{x}\beta = \begin{bmatrix} \beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots \\ \beta_0 + \beta_1 x_n \end{bmatrix}$$

πραγματικές σχέσεις που
συνδέουν τις τιμές της
ανεξάρτητης μεταβλητής
(X_i) με τις τιμές της
εξαρτημένης (Y_i)

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$
$$Y = X\beta + \varepsilon$$

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ συντελεστές
παλινδρόμησης

		country	tot_edu c	mil_ex p	BOM	fem_em p
1	2017	Austria	11	0,8	42	52
2	2017	Belgium	12	0,9	40	45
3	2017	Denmark	15	1,1	54	54
4	2017	Finland	12	1,4	45	51
5	2017	France	10	1,9	60	46
6	2017	Germany	11	1,2	35	53
7	2017	Ireland	13	0,3	38	52
8	2017	Luxembourg	9	0,6	41	51
9	2017	Netherlands	12	1,2	51	55
10	2017	Norway	16	1,6	56	58
11	2017	Sweden	16	1,0	55	57
12	2017	United Kingdom	14	1,8	48	55

πρωταρχικά δεδομένα: πηγή Παγκόσμια Τράπεζα

ΠΑΡΑΔΕΙΓΜΑ υπολογισμού των συντελεστών της παλινδρόμησης στο απλό γραμμικό μοντέλο

Ψ : εξαρτημένη
("γυναικεία εργασία")

X_1 : ανεξάρτητη
("κρατικές δαπάνες για την εκπαίδευση")

		country	tot_edu c	mil_ex p	BOM	fem_em p
1	2017	Austria	11	0,8	42	52
2	2017	Belgium	12	0,9	40	45
3	2017	Denmark	15	1,1	54	54
4	2017	Finland	12	1,4	45	51
5	2017	France	10	1,9	60	46
6	2017	Germany	11	1,2	35	53
7	2017	Ireland	13	0,3	38	52
8	2017	Luxembourg	9	0,6	41	51
9	2017	Netherlands	12	1,2	51	55
10	2017	Norway	16	1,6	56	58
11	2017	Sweden	16	1,0	55	57
12	2017	United Kingdom	14	1,8	48	55

πραγματικά δεδομένα: πηγή Παγκόσμια Τράπεζα

ΠΑΡΑΔΕΙΓΜΑ υπολογισμού των συντελεστών της παλινδρόμησης στο απλό γραμμικό μοντέλο

$$\Psi_1 = \alpha + \beta * X_1 + \varepsilon_1$$

$$\Psi_2 = \alpha + \beta * X_2 + \varepsilon_2$$

$$\Psi_3 = \alpha + \beta * X_3 + \varepsilon_3$$

...

$$\Psi_{10} = \alpha + \beta * X_{10} + \varepsilon_{10}$$

$$\Psi_{11} = \alpha + \beta * X_{11} + \varepsilon_{11}$$

$$\Psi_{12} = \alpha + \beta * X_{12} + \varepsilon_{12}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$Y = X\beta + \varepsilon$

		country	tot_edu c	mil_ex p	BOM	fem_em p
1	2017	Austria	11	0,8	42	52
2	2017	Belgium	12	0,9	40	45
3	2017	Denmark	15	1,1	54	54
4	2017	Finland	12	1,4	45	51
5	2017	France	10	1,9	60	46
6	2017	Germany	11	1,2	35	53
7	2017	Ireland	13	0,3	38	52
8	2017	Luxembourg	9	0,6	41	51
9	2017	Netherlands	12	1,2	51	55
10	2017	Norway	16	1,6	56	58
11	2017	Sweden	16	1,0	55	57
		United				
12	2017	Kingdom	14	1,8	48	55

πραγματικά δεδομένα: πηγή: Παγκόσμια Τράπεζα

52		1	11		ε1
45		1	12		ε2
54		1	15		ε3
51		1	12		ε4
...	=	...	...	*	...
...		...	...		...
...		...	...		...
58		1	16		ε10
57		1	16		ε11
55		1	14		ε12

β_0
 β_1

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ συντελεστές
παλινδρόμησης

$$\left(\begin{array}{|c|} \hline 1 & 11 \\ \hline 1 & 12 \\ 1 & 15 \\ 1 & 12 \\ \dots & \dots \\ \dots & \dots \\ \dots & \dots \\ 1 & 16 \\ 1 & 16 \\ 1 & 14 \\ \hline \end{array} \right)^{-1} \begin{array}{|c|} \hline 1 & 1 & 1 & \dots & 1 & 1 \\ \hline 11 & 12 & 15 & \dots & 16 & 14 \\ \hline \end{array} \begin{array}{|c|} \hline 52 \\ 45 \\ 54 \\ 51 \\ \dots \\ \dots \\ \dots \\ 58 \\ 57 \\ 55 \\ \hline \end{array}$$

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ συντελεστές
παλινδρόμησης

The diagram illustrates the calculation of the regression coefficients $\mathbf{b}$ using the formula $\mathbf{b} = (X^T X)^{-1} X^T Y$. It shows the visual representation of the matrices involved in the calculation.

Matrix X (Design Matrix): A matrix with 6 columns (intercept and 5 predictors) and 10 rows (observations). The first row is $[1, 11, 12, 15, \dots, 16, 14]$. The first column contains all 1s.

Matrix $X^T X$ (Cross-product matrix): A 6x6 symmetric matrix. The top-left element is n (10). The top row contains the sums of the predictors. The bottom row contains the sums of the products of the predictors. The matrix is inverted, indicated by the -1 superscript.

Matrix $X^T Y$ (Cross-product vector): A 6x1 vector. The top element is $\sum y_i$ (52). The bottom element is $\sum x_i y_i$ (55).

Matrix Y (Response vector): A 10x1 vector containing the response values: 52, 45, 54, 51, ..., 58, 57, 55.

The calculation is shown as $(X^T X)^{-1} X^T Y$, resulting in the vector $\mathbf{b}$.

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ *συντελεστές
παλινδρόμησης*

$$\begin{matrix} \alpha & \beta \\ \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} & * & \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} \\ \gamma & \delta \end{matrix}^{-1}$$

The Inverse of a 2x2 Matrix

For any 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

The inverse is: $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

$$\frac{1}{n \sum X^2 - (\sum X)^2} \left| \begin{array}{cc|c} \sum X^2 & -\sum X & \sum \Psi \\ -\sum X & n & \sum X \Psi \end{array} \right|$$

$$= \frac{1}{12 \cdot 1.957 - 22.801^2} \left| \begin{array}{cc|c} 1.957 & -151 & 629 \\ -151 & 12 & 7.982 \end{array} \right|$$

$$= \frac{1}{683} \left| \begin{array}{c|c} 25.671 & 37.58 \\ 805 & 1.17 \end{array} \right|$$

Ψ

κατανάλωση χρέος

 X_1

μισθοί

2026	148	372	69
2025	145	366	66
2024	143	364	63
2023	141	369	60
2022	138	368	56
2021	127	364	53
2020	121	351	49
2019	129	331	52
2018	126	335	50
2017	123	318	48
2016	120	315	48
2015	119	312	47

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ συντελεστές
παλινδρόμησης

$$\left(\begin{array}{c|c} \begin{array}{cccccc} 1 & 1 & 1 & \dots & 1 & 1 \\ \hline 69 & 66 & 63 & \dots & 48 & 47 \end{array} & * & \begin{array}{c} \begin{array}{cc} 1 & 69 \\ 1 & 66 \\ 1 & 63 \\ 1 & 60 \\ \dots & \dots \\ \dots & \dots \\ \dots & \dots \\ 1 & 48 \\ 1 & 48 \\ 1 & 47 \end{array} & \begin{array}{c} -1 \\ * \\ * \end{array} & \begin{array}{c} \begin{array}{cccccc} 1 & 1 & 1 & \dots & 1 & 1 \\ \hline 69 & 66 & 63 & \dots & 48 & 47 \end{array} & * & \begin{array}{c} \begin{array}{c} 148 \\ 145 \\ 143 \\ 141 \\ \dots \\ \dots \\ \dots \\ 123 \\ 120 \\ 119 \end{array} \end{array} \end{array} \right)$$

$$\begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix}^{-1}$$

$$\begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix}$$

$$\mathbf{b} = (X^T X)^{-1} X^T Y$$

→ *συντελεστές
παλινδρόμησης*

$$\begin{matrix} \alpha & \beta \\ \begin{bmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{bmatrix} & \begin{bmatrix} \sum y_i \\ \sum x_i y_i \end{bmatrix} \end{matrix} \begin{matrix} -1 \\ * \end{matrix}$$

The Inverse of a 2x2 Matrix

For any 2x2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

The inverse is: $A^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

$$\begin{array}{c} \begin{array}{c|c|c} 1/n \sum X^2 - (\sum X)^2 & \sum X^2 & -\sum X \\ \hline -\sum X & n & \sum X \Psi \end{array} & \begin{array}{c} \sum \Psi \\ \hline \sum X \Psi \end{array} \\ \hline \end{array} = \frac{1}{12} \begin{array}{c|c|c} 37.114 & -662 & 1.580 \\ \hline -662 & 12 & 87.983 \end{array}$$

$$= \frac{1}{7.521} \begin{array}{c|c} 395.374 & 52.56 \\ \hline 9.836 & 1.30 \end{array}$$

Προβλέψεις

	Coeff	SE
b	56.927544	5.50293
μισθοί	1.356837	0.0990314

$$\Psi_i = a + \beta_1 * X_{1i} + \varepsilon_i$$

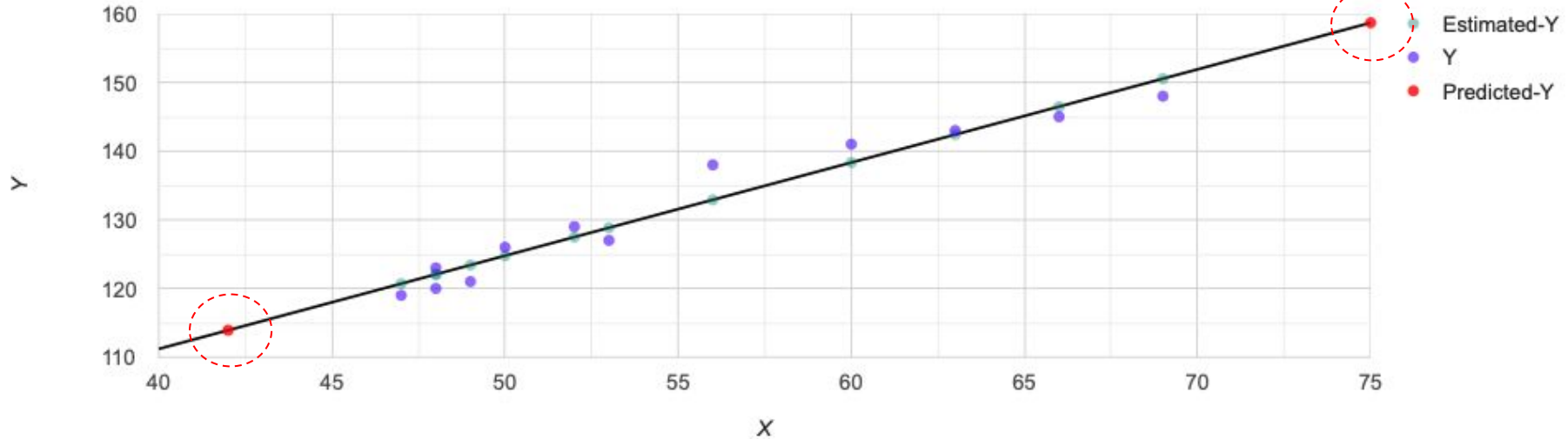
$$\widehat{\Psi}_i = \widehat{a} + \widehat{\beta}_1 * X_{1i}$$

$$\widehat{\Psi}_i = 56,92 + 1,3568 * X_{1i}$$

Αν μισθοί = 75 δισ. $\rightarrow \Psi_o = 56,92 + 1,3568*(75) = 158,68$ δισ.

Αν μισθοί = 42 δισ. $\rightarrow \Psi_o = 56,92 + 1,3568*(42) = 113,90$ δισ.

Line Fit Plot



Αν μισθοί = 75 δισ. $\rightarrow \Psi_0 = 56,92 + 1,3568 \cdot (75) = 158,68$ δισ.

Αν μισθοί = 42 δισ. $\rightarrow \Psi_0 = 56,92 + 1,3568 \cdot (42) = 113,90$ δισ.

	Coeff	SE
b	118.072081	12.507486
ανεργία	-0.699445	0.33654
επενδύσεις	0.936022	0.273198

$$\Psi_i = a + \beta_1 * X_{1i} + \beta_2 * X_{2i} + \varepsilon_i$$

$$\widehat{\Psi}_i = \widehat{a} + \widehat{\beta}_1 * X_{1i} + \widehat{\beta}_2 * X_{2i}$$

$$\widehat{\Psi}_i = 118,07 - 0,6994 * X_{1i} + 0,9360 * X_{2i}$$

$$\widehat{\Psi}_i = 118,07 - 0,6994 * X_{1i} + 0,9360 * X_{2i}$$

Προβλεπτικά Σενάρια:

αισιόδοξο: ανεργία ($X_{1o} = 7\%$), επενδύσεις ($X_{2o} = 43$ δισ.)

$$\begin{aligned} \rightarrow \Psi_o &= 118,07 - 0,6994 * 7 + 0,9360 * 43 = \\ &118,07 - 4,8958 + 40,248 = \mathbf{153,42} \text{ (δισ.)} \end{aligned}$$

απαισιόδοξο: ανεργία ($X_{1o} = 28\%$), επενδύσεις ($X_{2o} = 18$ δισ.)

$$\begin{aligned} \rightarrow \Psi_o &= 118,07 - 0,6994 * 28 + 0,9360 * 18 = \\ &118,07 - 19,5832 + 16,848 = \mathbf{115,33} \text{ (δισ.)} \end{aligned}$$

συγκρατημένο: ανεργία ($X_{1o} = 21\%$), επενδύσεις ($X_{2o} = 24$ δισ.)

$$\begin{aligned} \rightarrow \Psi_o &= 118,07 - 0,6994 * 21 + 0,9360 * 24 = \\ &118,07 - 13,988 + 22,464 = \mathbf{126,54} \text{ (δισ.)} \end{aligned}$$

Διάστημα Εμπιστοσύνης για τους πραγματικούς συντελεστές β

	Coeff	SE
b	118.072081	12.507486
ανεργία	-0.699445	0.33654
επενδύσεις	0.936022	0.273198

Ο πίνακας στα αριστερά δίνει τα αποτελέσματα της παλινδρόμησης για το μοντέλο:

$$\Psi_i = \alpha + \beta_1 * X1_i + \beta_2 * X2_i + \varepsilon_i$$

όπου:

$\Psi \rightarrow$ κατανάλωση

$X1 \rightarrow$ ανεργία

$X2 \rightarrow$ επενδύσεις

$n=12$ (παρατηρήσεις)

Να κατασκευαστούν τα διαστήματα εμπιστοσύνης στο επίπεδο $\alpha=5\%$ για τους πραγματικούς συντελεστές β

	Coeff	SE
b	118.072081	12.507486
ανεργία	-0.699445 *	0.33654
επενδύσεις	0.936022 ***	0.273198

$$P(\hat{\beta}_j - t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j) < \beta_j < \hat{\beta}_j + t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j)) = 95\%$$

$$P(-0,6994 - 2,26 * 0,3365 < \beta_1 < -0,6994 + 2,26 * 0,3365)$$

$$P(-1,45 < \beta_1 < 0,06) = 95\%$$

	Coeff	SE
b	118.072081	12.507486
ανεργία	-0.699445 *	0.33654
επενδύσεις	0.936022 ***	0.273198

$$P(\hat{\beta}_j - t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j) < \beta_j < \hat{\beta}_j + t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j)) = 95\%$$

$$P(0,9360 - 2,26 * 0,2731 < \beta_2 < 0,9360 + 2,26 * 0,2731)$$

$$P(0,3187 < \beta_2 < 1,5532) = 95\%$$

	Coeff	SE
b	46.687956	12.989219
μισθοί	1.237617	0.16945
χρέος	0.0484225	0.0554917

Ο πίνακας στα αριστερά δίνει τα αποτελέσματα της παλινδρόμησης για το μοντέλο:

$$\Psi_i = \alpha + \beta_1 * X1_i + \beta_2 * X2_i + \varepsilon_i$$

όπου:

$\Psi \rightarrow$ κατανάλωση

$X1 \rightarrow$ μισθοί

$X2 \rightarrow$ χρέος

$n=12$ (παρατηρήσεις)

Να κατασκευαστούν τα διαστήματα εμπιστοσύνης στο επίπεδο $\alpha=1\%$ για τους πραγματικούς συντελεστές β

	Coeff	SE
b	46.687956	12.989219
μισθοί	1.237617	0.16945
χρέος	0.0484225	0.0554917

$$P(\hat{\beta}_j - t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j) < \beta_j < \hat{\beta}_j + t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j)) = 99\%$$

$$P(1,2376 - 3,169 * 0,1694 < \beta_1 < 1,2376 + 3,169 * 0,1694)$$

$$P(0,7007 < \beta_1 < 1,7744) = 99\%$$

	Coeff	SE
b	46.687956	12.989219
μισθοί	1.237617	0.16945
χρέος	0.0484225	0.0554917

$$P(\hat{\beta}_j - t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j) < \beta_j < \hat{\beta}_j + t_{\alpha/2, n-k-1} * s.e(\hat{\beta}_j)) = 99\%$$

$$P(0,0484 - 3,169 * 0,0554 < \beta_2 < 0,0484 + 3,169 * 0,0554)$$

$$P(-0,1271 < \beta_2 < 0,2239) = 99\%$$