

Λυμένα προβλήματα στην πρόβλεψη

Πρόβλημα 1

Το μοντέλο ζήτησης κάποιου προϊόντος για τον μήνα t είναι $\hat{z}_t = at^2 + bt$. Στους μήνες $t=1, \dots, 4$ η ζήτηση ήταν $z_1=5, z_2=20, z_3=43, z_4=71$. Εύρετε τους συντελεστές a, b που ελαχιστοποιούν το άθροισμα τετραγώνων των σφαλμάτων και μετά κάνετε πρόβλεψη για την Z_7 .

Λύση

Εδώ $t_1=1, \dots, t_4=4$. Κριτήριο, $\varepsilon^2(1)+\dots+\varepsilon^2(4)$ άρα $q_i=1$. $\hat{z}_t = \alpha^T F(t) = [a \ b] \begin{bmatrix} t^2 \\ t \end{bmatrix}$

Διάνυσμα βέλτιστων τιμών (τυπολόγιο): $\alpha = [a \ b]^T = (AQA^T)^{-1}AQX$, όπου

$$A = \begin{bmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 1 & 2 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \end{bmatrix} \quad q_i=1 \Rightarrow Q = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad X = \begin{bmatrix} 5 \\ 20 \\ 43 \\ 71 \end{bmatrix}$$

Απλοποίηση $\alpha = (AA^T)^{-1}AX$. Πράξεις:

$$AA^T = \begin{bmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 4 & 2 \\ 9 & 3 \\ 16 & 4 \end{bmatrix} = \begin{bmatrix} 354 & 100 \\ 100 & 30 \end{bmatrix} \quad AX = \begin{bmatrix} 1 & 4 & 9 & 16 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 20 \\ 43 \\ 71 \end{bmatrix} = \begin{bmatrix} 1608 \\ 458 \end{bmatrix}$$

! Αντιστροφή πίνακα (2×2) και (3×3)

$$(AA^T)^{-1} = \frac{1}{354 \cdot 30 - 100 \cdot 100} \begin{bmatrix} 30 & -100 \\ -100 & 354 \end{bmatrix} \quad \begin{bmatrix} a \\ b \end{bmatrix} = (AA^T)^{-1}AX = \frac{1}{620} \begin{bmatrix} 2440 \\ 1332 \end{bmatrix} \approx \begin{bmatrix} 3.94 \\ 2.15 \end{bmatrix}$$

$$\text{Τελικά } \hat{z}_7 = 3.94(7^2) + 2.15(7) = 208.11 \quad 2$$

Πρόβλημα 2

Έχει βρεθεί εμπειρικά ότι η ζήτηση d_t κάποιου προϊόντος κατά τον μήνα t προσεγγίζεται ικανοποιητικά από την εκτιμήτρια $\hat{z}_t = at + br_t$ όπου r_t είναι η τιμή του προϊόντος. Δεδομένα:

Εύρετε τα a και b που ελαχιστοποιούν το άθροισμα τετραγώνων των σφαλμάτων.

μήνας t :	1	2	3	4
τιμή p_t :	25	25	26	24
ζήτηση d_t :	0	5	1	52

Λύση

Όμοιο με το Πρόβλημα 1. Η διαφορά είναι στο ότι η p_t **δεν είναι συνάρτηση**, αλλά γνωστή χρονοσειρά. Όμοια επίλυση:

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 25 & 25 & 26 & 24 \end{bmatrix} \quad X = \begin{bmatrix} 0 \\ 5 \\ 1 \\ 52 \end{bmatrix}$$

Απλοποίηση $\alpha = (AA^T)^{-1}AX$. Πράξεις:

$$AA^T = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 25 & 25 & 26 & 24 \end{bmatrix} \begin{bmatrix} 1 & 25 \\ 2 & 25 \\ 3 & 26 \\ 4 & 24 \end{bmatrix} = \begin{bmatrix} 30 & 249 \\ 249 & 2502 \end{bmatrix} \quad AX = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 25 & 25 & 26 & 24 \end{bmatrix} \begin{bmatrix} 0 \\ 5 \\ 1 \\ 52 \end{bmatrix} = \begin{bmatrix} 221 \\ 1399 \end{bmatrix}$$

$$(AA^T)^{-1} = \frac{1}{13059} \begin{bmatrix} 2502 & -249 \\ -249 & 30 \end{bmatrix} = \begin{bmatrix} 0.191592 & -0.0191 \\ -0.019067 & 0.0023 \end{bmatrix} \quad \begin{bmatrix} a \\ b \end{bmatrix} = (AA^T)^{-1}AX = \begin{bmatrix} 15.6667 \\ -1 \end{bmatrix}$$

Είναι αναμενόμενο ο συντελεστής b να είναι αρνητικός: όταν αυξάνει η τιμή, μειώνεται η ζήτηση.

Πρόβλημα 3

Σε κάποιο πρατήριο βενζίνης, οι συνολικές πωλήσεις X_t (σε λίτρα) την ημέρα t έχουν άγνωστη μέση τιμή μ και περιγράφονται ικανοποιητικά από το μοντέλο κινητού μέσου όρου (MA), $X_t - \mu = a_t - 0.9a_{t-3}$, όπου a_t είναι λευκός θόρυβος με μέση τιμή 0 και διασπορά 10000. Να υπολογίσετε τις αυτοσυμμεταβλητότητες $\gamma_0, \gamma_1, \gamma_2$ και γ_3 , όπου $\gamma_k = E[(X_t - \mu)(X_{t-k} - \mu)]$. Ποια είναι η τάξη του μοντέλου MA;

Λύση

$$\begin{aligned}\gamma_k &= E[(a_t - 0.9a_{t-3})(a_{t-k} - 0.9a_{t-3-k})] \\ &= E(a_t a_{t-k}) - 0.9E(a_{t-3} a_{t-k}) - 0.9E(a_t a_{t-3-k}) + 0.9^2 E(a_{t-3} a_{t-3-k})\end{aligned}$$

Αντικαθιστούμε $k = 0, \dots, 3$ και θέτουμε $E(a_{t-i} a_{t-j}) = 0$ για $i \neq j$, αφού ο λευκός θόρυβος είναι ασυσχέτιστος. Προκύπτουν:

$$\begin{aligned}\gamma_0 &= E(a_t a_t) - 0.9E(a_{t-3} a_t) - 0.9E(a_t a_{t-3}) + 0.9^2 E(a_{t-3} a_{t-3}) \\ &= \sigma_a^2 - 0.9(0) - 0.9(0) + 0.9^2 \sigma_a^2 = 1.81(10000) = 18100\end{aligned}$$

$$\gamma_1 = E(a_t a_{t-1}) - 0.9E(a_{t-3} a_{t-1}) - 0.9E(a_t a_{t-3-1}) + 0.9^2 E(a_{t-3} a_{t-3-1}) = 0$$

$$\gamma_2 = E(a_t a_{t-2}) - 0.9E(a_{t-3} a_{t-2}) - 0.9E(a_t a_{t-3-2}) + 0.9^2 E(a_{t-3} a_{t-3-2}) = 0$$

$$\begin{aligned}\gamma_3 &= E(a_t a_{t-3}) - 0.9E(a_{t-3} a_{t-3}) - 0.9E(a_t a_{t-3-3}) + 0.9^2 E(a_{t-3} a_{t-3-3}) \\ &= -0.9(10000) = -9000\end{aligned}$$

Πρόκειται για μοντέλο MA(3). *Εύκολα αποδεικνύεται ότι $\gamma_k = 0$ για $k > 3$.*

4

Πρόβλημα 4

Η ζήτηση ενός προϊόντος τους προηγούμενους δέκα μήνες ήταν:

t	1	2	3	4	5	6	7	8	9	10
z_t	30.3	29.8	29.6	30	29.4	30.5	28.7	29.8	28.8	30.1

Προσαρμόστε στα δεδομένα το πλέον κατάλληλο μοντέλο AR(2) και από αυτό εύρετε μία εκτίμηση της ζήτησης για τον μήνα 11.

Λύση

Μέσος όρος $\bar{z} = (30.3 + \dots + 30.1)/10 = 29.7$. Υπολογισμοί $y_t = z_t - \bar{z}$, γ_k , ρ_k της Y_t .

t	1	2	3	4	5	6	7	8	9	10	$\bar{z} = 29.7$
z_t	30.3	29.8	29.6	30	29.4	30.5	28.7	29.8	28.8	30.1	
$y_t = z_t - \bar{z}$	0.6	0.1	-0.1	0.3	-0.3	0.8	-1	0.1	-0.9	0.4	
y_t^2	0.36	0.01	0.01	0.09	0.09	0.64	1	0.01	0.81	0.16	
$y_t y_{t-1}$	-	0.060	-0.010	-0.030	-0.090	-0.240	-0.800	-0.100	-0.090	-0.36	
$y_t y_{t-2}$	-	-	-0.060	0.030	0.030	0.240	0.300	0.080	0.900	0.04	

Εξισώσεις Yule-Walker (Y-W), γραμμικό σύστημα 2x2, επίλυση:

$$\begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} \Leftrightarrow \begin{bmatrix} \phi_1 \\ \phi_2 \end{bmatrix} = \frac{1}{1 - \rho_1^2} \begin{bmatrix} 1 & -\rho_1 \\ -\rho_1 & 1 \end{bmatrix} \begin{bmatrix} \rho_1 \\ \rho_2 \end{bmatrix} = \frac{1}{1 - \rho_1^2} \begin{bmatrix} \rho_1 - \rho_1 \rho_2 \\ \rho_2 - \rho_1^2 \end{bmatrix} = \begin{bmatrix} -0.3672 \\ 0.2980 \end{bmatrix}$$

Πρόβλεψη: $\hat{y}_{11} = \phi_1 y_{10} + \phi_2 y_9 = -0.416$ και $\hat{z}_{11} = y_{11} + \mu_z = 29.386$

Αντί Y-W, γραμμική παλινδρόμηση $\hat{y}_t = \alpha_1 f_1(t) + \alpha_2 f_2(t)$ όπου $f_i(t) = z_{t-i}$. Το κριτήριο εκεί είναι η ελαχιστοποίηση. Τότε $\alpha_i \approx \phi_i$ αλλά όχι ακριβώς ίδια.

5

Πρόβλημα 5

Αναλάβετε το τμήμα πωλήσεων προϊόντος που προσφέρεται σε ακέραιες ποσότητες. Στο γραφείο, ένα σημείωμα από την προκάτοχό σας γράφει:

Η ζήτηση ανά λεπτό έχει γεωμετρική κατανομή με παράμετρο από 0 ως 1.

Το βιβλίο πιθανοτήτων γράφει: *Γεωμετρική, $P(k)=(1-p)p^k$, $k=0,1,\dots$, $0<p<1$.*

"ΟΚ", σκέφτεστε, "p είναι η παράμετρος αλλά ποια τιμή να έχει;"

Στο γραφείο δεν υπάρχουν παλαιότερα δεδομένα της ζήτησης. Όμως στο πρώτο ακριβώς λεπτό έχετε την πρώτη μέτρηση, z_1 . Με ποιον τρόπο ακριβώς θα χρησιμοποιήσετε τη z_1 για να βελτιώσετε την εκτίμηση του p;

Λύση

Εδώ δίνεται συνδυασμός μέτρησης (z_1) και θεωρητικής κατανομής. Τέτοιο δεδομένο συναντήσαμε στις εκτιμήσεις Bayes και ελαχίστων τετραγώνων.

Βοήθειες: • ολοκλήρωμα $\int_0^1 x^n dx = 1/(n+1)$ • p ομοιόμορφη στο (0,1), άρα $f(p)=1$ • ο τύπος του Bayes, αντί $f(x_1, \dots, x_n | \theta)$ έχει $P(z_1 | p) = (1-p)p^{z_1}$.

Από τη θεωρία: Η εκτίμηση ελαχίστων τετραγώνων του Y αν γνωρίζουμε ότι $X=x$ είναι $\hat{y} = E(Y | X=x) = \int y f(y | x) dy$. Εδώ $\hat{p} = E(p | z_1) = \int_0^1 p f(p | z_1) dp$ όπου

$$f(p | z_1) = \frac{P(z_1 | p)f(p)}{\int_0^1 P(z_1 | p)f(p)dp} = \frac{(1-p)p^{z_1}}{\int_0^1 (1-p)p^{z_1} dp} = \frac{(1-p)p^{z_1}}{\int_0^1 p^{z_1} dp - \int_0^1 p^{z_1+1} dp} = \dots = (z_1 + 1)(z_1 + 2)(1-p)p^{z_1}$$

$$\hat{p} = E(p | z_1) = \int_0^1 (z_1 + 1)(z_1 + 2)(1-p)p^{z_1+1} dp = \frac{(z_1 + 1)(z_1 + 2)}{(z_1 + 2)(z_1 + 3)} = \frac{z_1 + 1}{z_1 + 3} \quad (!)$$