

ΤΑΞΙΝΟΜΗΤΕΣ ΕΛΑΧΙΣΤΗΣ ΑΠΟΣΤΑΣΗΣ

#1

Θεωρώντας το γενικότερο πλαίσιο ενός M -σταθμικού προβλήματος ταξινόμησης, υποθέτουμε την ισχύ των παρακάτω συνθηκών:

(i): οι εκ των προτέρων πιθανότητες εμφάνισης των κλάσεων είναι ίδιες:

$$P(\omega_i) = P(\omega_j), \quad \forall i \neq j.$$

(ii): οι πίνακες συνδιακύμανσης των διαφόρων κλάσεων είναι ίδιοι:

$$\underline{\underline{\Sigma}}_i = \underline{\underline{\Sigma}}_j = \underline{\underline{\Sigma}}, \quad \forall i \neq j.$$

Σύμφωνα με τα παραπάνω, η συνάρτηση διάκρισης της i -οσής κλάσης μπορεί να αποδοποιηθεί στην παρακάτω μορφή:

$$g_i(\underline{x}) = -\frac{1}{2} (\underline{x} - \underline{\mu}_i)^T \underline{\underline{\Sigma}}^{-1} (\underline{x} - \underline{\mu}_i)$$

(α): Στην περίπτωση κατά την οποία ο πίνακας $\underline{\underline{\Sigma}}$ είναι διαγώνιος, δηλαδή είναι της μορφής:

$$\underline{\underline{\Sigma}} = \sigma^2 \underline{\underline{I}},$$

η συνάρτηση διάκρισης της i -οσής κλάσης λαμβάνει την μορφή:

$$g_i(\underline{x}) = -\frac{1}{2\sigma^2} (\underline{x} - \underline{\mu}_i)^T (\underline{x} - \underline{\mu}_i) = -\frac{1}{2} \|\underline{x} - \underline{\mu}_i\|^2$$

Αυτό σημαίνει πως ο κανόνας απόφασης που αντιστοιχεί ένα άγνωστο πρότυπο στην κλάση με την μεγαλύτερη εικόνη της αντιστοιχίας συνάρτησης διάκρισης, μετασχηματίζεται στην ανάθεση του άγνωστου διαυδμένου χαρακτηριστικού σε εκείνη των κλάσεων με την μικρότερη ευκαμψία απόστασης από το αντιστοιχικό κέντρο.

ΚΑΝΟΝΑΣ ΑΠΟΦΑΣΗΣ:

Ανάθεση ενός άγνωστου προτύπου \underline{x} στην κλάση (ω_i) όταν: $\|\underline{x} - \underline{\mu}_i\| < \|\underline{x} - \underline{\mu}_j\|, \quad \forall i \neq j.$

Σύμφωνα με τα παραπάνω, σε περιοχές της απόστασης από το κέντρο ως κάθε κλάση αντιστοιχούν σε υπερσφαιρές ακτίνας c που δίνονται από την σχέση:

$$d_e(\underline{x}, \mu_i) = c \Leftrightarrow \|\underline{x} - \mu_i\| = c \Leftrightarrow \|\underline{x} - \mu_i\|^2 = c^2 \Leftrightarrow$$

$$\sum_{k=1}^p (x_k - \mu_{ik})^2 = c^2, \quad \forall i \in [M].$$

(β): Για την περίπτωση όπου ο κοινός πίνακας συνδιακύμανσης δεν είναι διαγώνιος, τότε η διαδικασία μεγιστοποίησης της τιμής της εκάστοτε συνάρτησης διαίρεσης $g_i(\underline{x})$ οδηγεί στην ελαχιστοποίηση της απόστασης από το κάθε κέντρο με την νόρμα ως προς του πίνακα $\underline{\Sigma}^{-1}$.

ΚΑΝΟΝΑΣ ΑΠΟΦΑΣΗΣ:

Αυθόθημη ενός άγνωστου πτυσίου \underline{x} στην κλάση (ω_i) όταν: $\|\underline{x} - \mu_i\|_{\underline{\Sigma}^{-1}} < \|\underline{x} - \mu_j\|_{\underline{\Sigma}^{-1}}$, $\forall i \neq j$.

Αντίστοιχα, σε περιοχές της απόστασης από το κέντρο της κάθε κλάσης θα δίνονται την γενική περίπτωση από υπερελλειψοειδείς καμπύλες της μορφής:

$$d_m(\underline{x}, \mu_i) = \sqrt{(\underline{x} - \mu_i)^T \underline{\Sigma}^{-1} (\underline{x} - \mu_i)} = c \Leftrightarrow$$

$$(\underline{x} - \mu_i)^T \underline{\Sigma}^{-1} (\underline{x} - \mu_i) = c^2, \quad \forall i \in [M]$$

↓ Mahalanobis Distance

Η συμπεριφορά του πίνακα συδιακύμανσης $\underline{\underline{\Sigma}}$ παρέχει πληροφορίες των δυνατοτήτων διαγωνιοποίησης του μέσω του μοναδιαίου μετασχηματισμού:

[Ορθογώνια Matrix]

$$\underline{\underline{\Sigma}} = \underline{\underline{\Phi}} \underline{\underline{\Lambda}} \underline{\underline{\Phi}}^T$$

όπου $\underline{\underline{\Phi}}^T = \underline{\underline{\Phi}}^{-1}$ και $\underline{\underline{\Lambda}}$ ένας διαγώνιος πίνακας τα στοιχεία της κύριας διαγωνίου του οποίου είναι οι ιδιοτιμές του πίνακα $\underline{\underline{\Sigma}}$. Αντίστοιχα, οι στήλες του πίνακα $\underline{\underline{\Phi}}$ δίνονται από τα ορθοκανονικά ιδιοδιανύσματα του πίνακα $\underline{\underline{\Sigma}}$ έτσι ώστε:

$$\underline{\underline{\Phi}} = [\underline{u}_1, \underline{u}_2, \dots, \underline{u}_e]$$

* Οι στήλες του πίνακα $\underline{\underline{\Phi}}$ ικανοποιούν την σχέση $\underline{\underline{\Sigma}} \underline{u}_j = \lambda_j \cdot \underline{u}_j$, $\forall j \in \{e\}$ με $\|\underline{u}_j\| = 1$, $\forall j \in \{e\}$.

$$\text{και } \underline{u}_r^T \underline{u}_s = \delta_{rs}, \forall r \neq s.$$

• Συνδυάζοντας τις προηγούμενες σχέσεις έχουμε ότι:

$$\underline{\underline{\Sigma}}^{-1} = (\underline{\underline{\Phi}} \underline{\underline{\Lambda}} \underline{\underline{\Phi}}^T)^{-1} \Leftrightarrow \underline{\underline{\Sigma}}^{-1} = \underline{\underline{\Phi}}^{-T} \underline{\underline{\Lambda}}^{-1} \underline{\underline{\Phi}}^{-1} \Leftrightarrow$$

$$\underline{\underline{\Sigma}}^{-1} = (\underline{\underline{\Phi}}^T)^T \underline{\underline{\Lambda}}^{-1} \underline{\underline{\Phi}}^T \Leftrightarrow \underline{\underline{\Sigma}}^{-1} = \underline{\underline{\Phi}} \underline{\underline{\Lambda}}^{-1} \underline{\underline{\Phi}}^T$$

Επείδη, η εξίσωση που θα περιγράψει τον γεωμετρικό τόπο των σημείων που ικανοποιούν υπό την έννοια της απόστασης Mahalanobis από το μέτρο $\underline{\mu}$, θα δίνεται από την σχέση:

$$(\underline{x} - \underline{\mu})^T \underline{\underline{\Phi}} \underline{\underline{\Lambda}}^{-1} \underline{\underline{\Phi}}^T (\underline{x} - \underline{\mu}) = c^2 \Leftrightarrow$$

$$(\underline{x} - \underline{\mu})^T \underbrace{[\underline{u}_1 \ \underline{u}_2 \ \dots \ \underline{u}_e]}_{(\underline{\underline{\Phi}})} \begin{bmatrix} \lambda_1^{-1} & & & \\ & \lambda_2^{-1} & & \\ & & \ddots & \\ & & & \lambda_e^{-1} \end{bmatrix} \begin{bmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \vdots \\ \underline{u}_e^T \end{bmatrix} (\underline{x} - \underline{\mu}) = c^2$$

[1x0]

[ex0]

[ex0]

[ex0]

[ex1]

[1x1]

$(\underline{\underline{\Lambda}}^{-1})$

$(\underline{\underline{\Phi}}^T)$

Συν προσημαστικότητα ο πίνακας $\underline{\Phi}$ ορίζει έναν γραμμικό μετασχηματισμό του εκάστου διανύσματος $\underline{x} \in \mathbb{R}^e$ έτσι ώστε το μετασχηματισμένο διάνυσμα $\underline{x}' \in \mathbb{R}^e$ να δίνει από την σχέση:

$$\underline{x}' = \underline{\Phi}^T \cdot \underline{x} \Leftrightarrow \underline{\text{LINEAR TRANSFORMATION}}$$

$$\underline{x}' = \begin{bmatrix} \underline{u}_1^T \\ \underline{u}_2^T \\ \vdots \\ \underline{u}_e^T \end{bmatrix} \cdot \underline{x} \Leftrightarrow [x'_1 \ x'_2 \ \dots \ x'_e] = [u_1^T \underline{x} \ u_2^T \underline{x} \ \dots \ u_e^T \underline{x}]$$

Με άλλα λόγια, οι νέες συντεταγμένες του διανύσματος \underline{x} , (\underline{x}'), δίνονται ως οι προβολές του \underline{x} στο εκάστοτε ορθομοναδιαίο ιδιοδιάνυσμα ως:

$$x'_j = \underline{u}_j^T \cdot \underline{x}, \quad \forall j \in \{e\}.$$

Με βάση τα παραπάνω, η ποσότητα q_j μπορεί να επαναδιατυπωθεί ως:

$$q_j = \underline{u}_j^T \cdot \underline{x} - \underline{u}_j^T \cdot \underline{\mu}_j \Rightarrow q_j = x'_j - \mu'_{ij}$$

που με την σειρά της θα δώσει:

$$\sum_{j=1}^e \frac{(x'_j - \mu'_{ij})^2}{\lambda_j} = c^2$$

ΥΠΕΡΕΛΛΙΨΕΙΔΕΣ

Άσκηση: Να ευρεθεί ο γεωμετρικός τόπος των σημείων που έχουν σταθερή απόσταση Mahalanobis από το κέντρο $\underline{\mu} = [0 \ 0]^T$ όταν ο πίνακας συνδιακύμανσης έχει την παρακάτω μορφή:

$$\underline{\Sigma} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \quad \mu \in \rho \in [-1, +1].$$

Η ζητούμενη εξίσωση έχει την μορφή:

$$d_m(\underline{x}; \rho) = \|\underline{x} - \underline{\mu}\|_{\underline{\Sigma}^{-1}} = c \Leftrightarrow (\underline{x} - \underline{\mu})^T \underline{\Sigma}^{-1} (\underline{x} - \underline{\mu}) = c^2.$$

Λαμβάνοντας υπόψη το γεγονός ότι $\underline{\mu} = [0 \ 0]^T$ και $\underline{x} = [x_1 \ x_2]^T$, θα έχουμε ότι: $\underline{x} - \underline{\mu} = [x_1 \ x_2]^T - [0 \ 0]^T = [x_1 \ x_2]^T = \underline{x}$. Επομένως, η παραπάνω σχέση γίνεται:

$$d_m(\underline{x}; \rho) = \underline{x}^T \underline{\Sigma}^{-1} \underline{x} = c^2$$

Αρχικά θα πρέπει να διαγωνιοποιήσουμε τον πίνακα $\underline{\Sigma}$ έτσι ώστε:

$$\underline{\Sigma} = \underline{\Phi} \underline{\Lambda} \underline{\Phi}^T$$

$$\mu \in \left\{ \begin{array}{l} \underline{\Phi} = [\underline{u}_1 \ \underline{u}_2] \\ \underline{\Phi}^T = \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} \end{array} \right\} \text{ και } \underline{\Lambda} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \text{ όπου}$$

\underline{u}_1 και \underline{u}_2 είναι τα ορθοκανονικά διανύσματα του πίνακα $\underline{\Sigma}$ και λ_1, λ_2 οι αντίστοιχες ιδιοτιμές.

Αρχικά υπολογίζουμε τις ιδιοτιμές του πίνακα Σ προσδιορίζοντας τις ρίζες του χαρακτηριστικού πολυωνύμου του Σ ως εξής:

$$p(\lambda) = \left| \Sigma - \lambda I \right| = 0 \Leftrightarrow \left| \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0 \Leftrightarrow$$

$$\Leftrightarrow \left| \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \right| = 0 \Leftrightarrow \begin{vmatrix} 1-\lambda & \rho \\ \rho & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow$$

$$p(\lambda) = (1-\lambda)^2 - \rho^2 = 0 \Leftrightarrow p(\lambda) = [(1-\lambda)-\rho][(1-\lambda)+\rho] = 0 \Leftrightarrow$$

$$p(\lambda) = (1-\lambda-\rho)(1-\lambda+\rho) = 0 \Leftrightarrow$$

$$\begin{cases} \lambda_1 = 1-\rho \\ \lambda_2 = 1+\rho \end{cases}$$

Εν συνεχεία, υπολογίζουμε τα ορθομοναδιαία διανύσματα του πίνακα Σ , \underline{u}_1 και \underline{u}_2 που αντιστοιχούν στις ιδιοτιμές λ_1 και λ_2 .

$$(i): \Sigma \cdot \underline{u}_1 = \lambda_1 \cdot \underline{u}_1 \Leftrightarrow \Sigma \cdot \underline{u}_1 - \lambda_1 \underline{u}_1 = \underline{0} \Leftrightarrow (\Sigma - \lambda_1 I) \cdot \underline{u}_1 = \underline{0}$$

Θέτουμε, $\underline{u}_1 = \begin{bmatrix} w \\ z \end{bmatrix}$ και αντικαθιστώντας στην παραπάνω σχέση έχουμε:

$$\left\{ \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - (1-\rho) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\left\{ \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - \begin{bmatrix} 1-\rho & 0 \\ 0 & 1-\rho \end{bmatrix} \right\} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} 1-(1-\rho) & \rho \\ \rho & 1-(1-\rho) \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{bmatrix} \rho & \rho \\ \rho & \rho \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\begin{cases} \rho w + \rho z = 0 \\ \rho w + \rho z = 0 \end{cases} \Leftrightarrow \rho w = -\rho z \xrightarrow{\rho \neq 0} \boxed{w = -z}$$

Επομένως, $\underline{u}_1 = \begin{bmatrix} -z \\ z \end{bmatrix}$ που με $z = \kappa$ έχουμε ότι: $\underline{u}_1 = \begin{bmatrix} -\kappa \\ \kappa \end{bmatrix} \Leftrightarrow$

$\underline{u}_1 = \kappa \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Για να εξασφαλισουμε ότι $\|\underline{u}_1\| = 1$, η τιμή του κ

θα πρέπει να δίνεται ως: $\kappa = \frac{1}{\sqrt{(-1)^2 + 1^2}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$.

Άρα, θα έχουμε ότι:

$$\underline{u}_1 = \begin{bmatrix} -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix}$$

(ii): $\underline{\Sigma} \cdot \underline{u}_2 = \lambda_2 \cdot \underline{u}_2 \Leftrightarrow \underline{\Sigma} \cdot \underline{u}_2 - \lambda_2 \underline{u}_2 = \mathbf{0} \Leftrightarrow (\underline{\Sigma} - \lambda_2 \underline{I}) \cdot \underline{u}_2 = \mathbf{0}$

Θέτουμε, $\underline{u}_2 = \begin{bmatrix} w \\ z \end{bmatrix}$ και αντικαθιστώντας στην παραπάνω σχέση έχουμε:

$$\left\{ \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - (1+\rho) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\left\{ \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} - \begin{bmatrix} 1+\rho & 0 \\ 0 & 1+\rho \end{bmatrix} \right\} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} 1-(1+\rho) & \rho \\ \rho & 1-(1+\rho) \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow$$

$$\begin{bmatrix} -\rho & \rho \\ \rho & -\rho \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow \begin{cases} -\rho w + \rho z = 0 \\ -\rho w + \rho z = 0 \end{cases} \Leftrightarrow -\rho w = -\rho z \xrightarrow{\rho \neq 0} \boxed{w = z}$$

Επομένως, $\underline{u}_2 = \begin{bmatrix} z \\ z \end{bmatrix}$ που με $z = \kappa$ έχουμε ότι: $\underline{u}_2 = \begin{bmatrix} \kappa \\ \kappa \end{bmatrix} \Leftrightarrow$

$\underline{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \kappa$. Για να εξασφαλισουμε ότι $\|\underline{u}_2\| = 1$, η τιμή του κ θα πρέπει να

δίνεται ως: $\kappa = \frac{1}{\sqrt{1^2 + 1^2}} = \frac{1}{\sqrt{2}} = \frac{\sqrt{2}}{2}$. $\underline{u}_2 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}$

Επομένως, ο πίνακας Φ^T θα δίνεται από την σχέση:

$$\Phi^T = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix} \text{ και το διάνυσμα } x' = \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} \text{ θα δίνεται ως:}$$

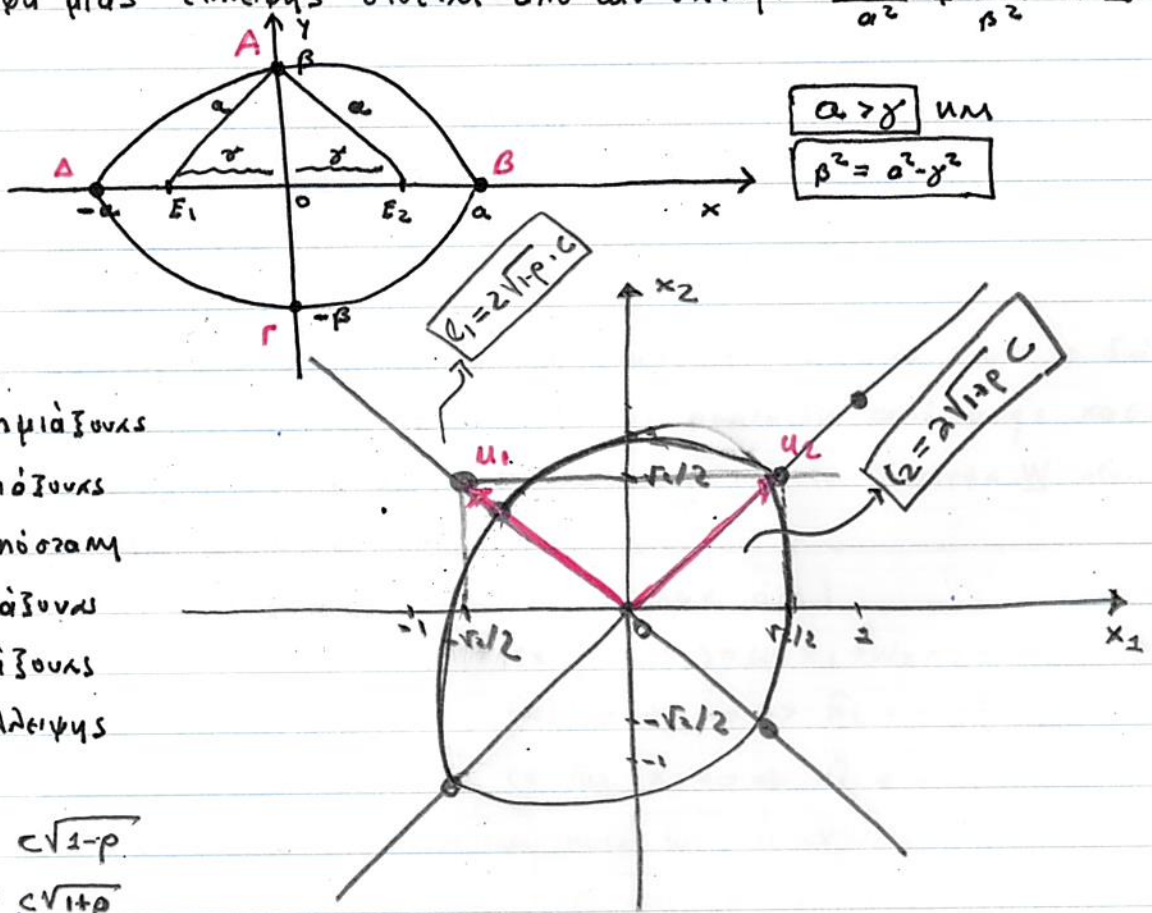
$$x' = \Phi^T \cdot x \Leftrightarrow \begin{bmatrix} x'_1 \\ x'_2 \end{bmatrix} = \begin{bmatrix} -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Leftrightarrow$$

$$\left\{ \begin{array}{l} x'_1 = -\frac{\sqrt{2}}{2} x_1 + \frac{\sqrt{2}}{2} x_2 \\ x'_2 = \frac{\sqrt{2}}{2} x_1 + \frac{\sqrt{2}}{2} x_2 \end{array} \right\}$$

Αυτή σχέση ως η εξίσωση $d_m(x;p) = c \Leftrightarrow \frac{(x'_1)^2}{1-\rho} + \frac{(x'_2)^2}{1+\rho} = c^2 \Leftrightarrow$

$$\frac{(x'_1)^2}{\sqrt{(1-\rho)^2}} + \frac{(x'_2)^2}{\sqrt{(1+\rho)^2}} = c^2 \Leftrightarrow \boxed{\frac{(x'_1)^2}{c^2 \sqrt{(1-\rho)^2}} + \frac{(x'_2)^2}{c^2 \sqrt{(1+\rho)^2}} = 1}$$

⊗ Η γενική μορφή μιας έλλειψης δίνεται από την σχέση: $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$



- a: Μεγάλος ημιάξονας
- b: Μικρός ημιάξονας
- γ: Εστιακή απόσταση
- ΔΒ: Μεγάλος άξονας
- ΑΓ: Μικρός άξονας
- Ο: Κέντρο έλλειψης

$$a = c\sqrt{1-\rho}$$

$$b = c\sqrt{1+\rho}$$

β1 και β2 τα μήκη των κυρίων αξόνων της έλλειψης.