#1

ΑΣΚΗΣΗ: Μπορεί το κριτήριο της απόστασης να χρησιμοποιηθεί για την εύρεση των άγνωστων παραμέτρων $\underline{\mu}_\kappa$ και $\underline{\underline{\Sigma}}_\kappa$ της δεσμευμένης ως προς την κλάση συνάρτησης πυκνότητας πιθανότητας $p(\underline{x}|\omega_\kappa)$ όταν $\forall \underline{x} \in \omega_\kappa$, $\underline{x} \sim N_n(\underline{\mu}_\kappa, \underline{\underline{\Sigma}}_\kappa)$

- Σε αυτό το πλαίσιο μπορούμε να ορίσουμε το μέσο διάνυσμα $\underline{\mu}_\kappa$ ως το διάνυσμα εκείνο που παρουσιάζει την ελάχιστη συνολική απόσταση προς το σύνολο των δεδομένων της κ-οστής κλάσης

  $\downarrow$

  Mahalanobis

  ως προς του πίνακα συνδιακύμανσης $\underline{\underline{\Sigma}}_\kappa$.

- Έστω $\mathcal{X}^{(\kappa)} = \{\underline{x}_1^{(\kappa)}, \underline{x}_2^{(\kappa)}, \dots, \underline{x}_N^{(\kappa)}\}$ με $\underline{x}_j^{(\kappa)} \in \mathbb{R}^n$ το σύνολο των δεδομένων της κ-οστής κλάσης.

- **ΟΡΙΣΜΟΣ ΣΥΝΑΡΤΗΣΙΟΕΙΔΟΥΣ ΚΟΣΤΟΥΣ:** [COST FUNCTION DEFINITION]

$$\boxed{J_\kappa(\underline{\mu}_\kappa, \underline{\underline{\Sigma}}_\kappa) = \sum_{j=1}^{N} \| \underline{x}_j^{(\kappa)} - \underline{\mu}_\kappa \|^2_{\underline{\underline{\Sigma}}_\kappa^{-1}} \quad (1)}$$

$$\boxed{J_\kappa(\underline{\mu}_\kappa, \underline{\underline{\Sigma}}_\kappa) = \sum_{j=1}^{N} (\underline{x}_j^{(\kappa)} - \underline{\mu}_\kappa)^T \underline{\underline{\Sigma}}_\kappa^{-1} (\underline{x}_j^{(\kappa)} - \underline{\mu}_\kappa) \quad (2)}$$

- **Πρόβλημα Βελτιστοποίησης:** [OPTIMIZATION PROBLEM DEFINITION]

$$\boxed{(\underline{\mu}_\kappa^*, \underline{\underline{\Sigma}}_\kappa^*) = \arg\min_{\underline{\mu}_\kappa, \underline{\underline{\Sigma}}_\kappa} J_\kappa(\underline{\mu}_\kappa, \underline{\underline{\Sigma}}_\kappa) \quad (3)}$$

ΣΥΝΘΗΚΕΣ ΠΡΩΤΗΣ ΤΑΞΗΣ (FIRST-ORDER CONDITIONS)

$$\begin{cases} \dfrac{\partial J_k}{\partial \boldsymbol{\mu}_k} = \mathbf{0} \quad [n \times 1] \\[2ex] \dfrac{\partial J_k}{\partial \boldsymbol{\Sigma}_k} = \mathbf{0} \quad [n \times n] \end{cases} \qquad (3)$$

Επαναδιατύπωση / Επέκταση Συναρτησιοειδούς Κόστους

$$J_k(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \sum_{j=1}^{N} (\mathbf{x}_j^{(k)})^T \boldsymbol{\Sigma}_k^{-1} \mathbf{x}_j^{(k)} - 2 \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}_k^{-1} \mathbf{x}_j^{(k)} + \boldsymbol{\mu}_k^T \boldsymbol{\Sigma}_k^{-1} \boldsymbol{\mu}_k \qquad (4)$$

(VECTOR/MATRIX DERIVATIVES)

Χρειαζόμαστε τις παρακάτω διανυσματικές παραγώγους

(i): $$\frac{\partial \mathbf{a}^T \mathbf{X}^{-1} \mathbf{b}}{\partial \mathbf{X}} = -\mathbf{X}^{-T} \mathbf{a} \mathbf{b}^T \mathbf{X}^{-T} \qquad (5)$$ (Παράγωγος 1ης Τάξης)

(ii) $$\frac{\partial \mathbf{x}^T \mathbf{B} \mathbf{x}}{\partial \mathbf{x}} = (\mathbf{B} + \mathbf{B}^T) \cdot \mathbf{x} \qquad (6)$$ (Παράγωγος 2ης Τάξης)

(iii) $$\frac{\partial \mathbf{x}^T \mathbf{a}}{\partial \mathbf{x}} = \frac{\partial \mathbf{a}^T \mathbf{x}}{\partial \mathbf{x}} = \mathbf{a} \qquad (7)$$ (Παράγωγος 1ης Τάξης)

Αποσύνθεση Συνάρτησης Κόστους

$$J_k(\underline{\mu}_k, \underline{\underline{\Sigma}}_k) = \sum_{j=1}^{n} a_j(\underline{\underline{\Sigma}}_k) + \beta_j(\underline{\mu}_k, \underline{\underline{\Sigma}}_k) + \gamma_j(\underline{\mu}_k, \underline{\underline{\Sigma}}_k) \quad (8)$$

όπου

$$\begin{cases} a_j = (\underline{x}_j^{(k)})^T \underline{\underline{\Sigma}}_k^{-1} (x_j^{(k)}) & (9) \\ \beta_j = -2\underline{\mu}_k^T \underline{\underline{\Sigma}}_k^{-1} (x_j^{(k)}) & (10) \\ \gamma_j = \underline{\mu}_k^T \underline{\underline{\Sigma}}_k^{-1} \underline{\mu}_k & (11) \end{cases}$$

(*) We could write that:

$$\boxed{a_j + \beta_j + \gamma_j = \delta_j = (\underline{x}_j^{(k)} - \underline{\mu}_k)^T \underline{\underline{\Sigma}}_k^{-1} (\underline{x}_j^{(k)} - \underline{\mu}_k) \quad (12)}$$

▶ Υπολογισμός Μερικών Παραγώγων των Ποσοτήτων a_j, β_j, γ_j ως προς τις μεταβλητές $\underline{\mu}_k$ και $\underline{\underline{\Sigma}}_k$.

[A]:

$$(\alpha): \boxed{\frac{\partial a_j}{\partial \underline{\mu}_k} = \underline{0} \quad (13) \quad \{\text{ΔΕΝ ΕΙΝΑΙ ΣΥΝΑΡΤΗΣΗ ΤΟΥ } \underline{\mu}_k\}}$$

$$(\beta): \boxed{\frac{\partial \beta_j}{\partial \underline{\mu}_k} = -2\,\frac{\partial\, \underline{\mu}_k^T \underline{\underline{\Sigma}}_k^{-1} x_j^{(k)}}{\partial \underline{\mu}_k} \xrightarrow[{[n\times n]\,[n\times 1]}]{\underline{\underline{\Sigma}}_k^{-1} \cdot \underline{x}_j^{(k)} = [n\times 1] \text{ vector}} -2\,\underline{\underline{\Sigma}}_k^{-1} \underline{x}_j^{(k)} \quad (13)}$$

$$(\gamma): \boxed{\frac{\partial \gamma_j}{\partial \underline{\mu}_k} = \frac{\partial\, \underline{\mu}_k^T \underline{\underline{\Sigma}}_k^{-1} \underline{\mu}_k}{\partial \underline{\mu}_k} = \left(\underline{\underline{\Sigma}}_k^{-1} + (\underline{\underline{\Sigma}}_k^{-1})^T\right) \cdot \underline{\mu}_k \xrightarrow{(\underline{\underline{\Sigma}}_k^{-1})^T = \underline{\underline{\Sigma}}_k^{-1}} 2\,\underline{\underline{\Sigma}}_k^{-1} \underline{\mu}_k \quad (14)}$$

[B]:

$$(a):\ \frac{\partial \delta_j}{\partial \Sigma_k} = \frac{\partial (x_j^{(k)}-\mu_k)^T \Sigma_k^{-1} (x_j^{(k)}-\mu_k)}{\partial \Sigma_k} = -(\Sigma_k)^{-T}(x_j^{(k)}-\mu_k)(x_j^{(k)}-\mu_k)^T(\Sigma_k)^{-T} \Rightarrow$$

$$\boxed{\frac{\partial \delta_j}{\partial \Sigma_k} = -\Sigma_k^{-1}(x_j^{(k)}-\mu_k)(x_j^{(k)}-\mu_k)^T\Sigma_k^{-1} \quad (15)}$$

ΣΥΝΘΗΚΕΣ ΠΡΩΤΗΣ ΤΑΞΗΣ

$$[A]:\ \frac{\partial J_k}{\partial \mu_k} = 0 \Rightarrow \sum_{j=1}^{N} \cancelto{0}{\frac{\partial a_j}{\partial \mu_k}} + \frac{\partial \beta_j}{\partial \mu_k} + \frac{\partial \gamma_j}{\partial \mu_k} = 0 \rightarrow$$

$$\sum_{j=1}^{N} -\cancel{2}\Sigma_k^{-1}x_j^{(k)} + \cancel{2}\Sigma_k^{-1}\mu_k = 0 \rightarrow$$

$$\sum_{j=1}^{N} \Sigma_k^{-1}x_j^{(k)} = \sum_{j=1}^{N} \Sigma_k^{-1}\mu_k \rightarrow \cancel{\Sigma_k^{-1}} \cdot \sum_{j=1}^{N} x_j^{(k)} = N \cdot \cancel{\Sigma_k^{-1}}\mu_k \rightarrow$$

$$\xrightarrow[\ast\ \Sigma_k]{\text{both sides}} \boxed{\mu_k^{*} = \frac{1}{N}\sum_{j=1}^{N} x_j^{(k)} \quad (16)}$$

Ο ΥΠΟΛΟΓΙΣΜΟΣ ΤΟΥ ΜΕΣΟΥ ΔΙΑΝΥΣΜΑΤΟΣ ΕΙΝΑΙ ΑΝΕΞΑΡΤΗΤΟΣ ΤΟΥ ΜΗΤΡΩΟΥ ΣΥΝΔΙΑΚΥΜΑΝΣΗΣ Σ_k.

ΣΥΝΘΗΚΕΣ ΠΡΩΤΗΣ ΤΑΞΗΣ ΣΥΝΕΧΕΙΑ

[B]:

$$\frac{\partial J_k}{\partial \underline{\underline{\Sigma}}_k} = \underline{\underline{0}} \rightarrow \sum_{j=1}^{N} \frac{\partial J_j}{\partial \underline{\underline{\Sigma}}_k} = \underline{\underline{0}} \rightarrow$$

$$-\sum_{j=1}^{N} \underbrace{\underline{\underline{\Sigma}}_k^{-1}}_{[n\times n]} \underbrace{(\underline{x}_j^{(k)} - \underline{\mu}_k)}_{[n\times 1]} \underbrace{(\underline{x}_j^{(k)} - \underline{\mu}_k)^T}_{[1\times n]} \underbrace{\underline{\underline{\Sigma}}_k^{-1}}_{[n\times n]} = \underline{\underline{0}} \rightarrow$$

$$\boxed{\underbrace{\underline{\underline{\Sigma}}_k^{-1}}_{\underline{\underline{R}}} * \underbrace{\boxed{\sum_{j=1}^{N} (\underline{x}_j^{(k)} - \underline{\mu}_k)(\underline{x}_j^{(k)} - \underline{\mu}_k)^T}}_{\underline{\underline{W}}} * \underbrace{\underline{\underline{\Sigma}}_k^{-1}}_{\underline{\underline{R}}} = \underline{\underline{0}} \quad (17)}$$

⊛ Γνωρίζουμε ότι για το Δειγματικό Μητρώο Συνδιακύμανσης (Sample Covariance Matrix) ισχύει ότι:

$$\hat{\underline{\underline{\Sigma}}}_k = \frac{1}{N-1} * \sum_{j=1}^{N} (\underline{x}_j^{(k)} - \underline{\mu}_k)(\underline{x}_j^{(k)} - \underline{\mu}_k)^T \quad (18) \Rightarrow$$

$$\sum_{j=1}^{N} (\underline{x}_j^{(k)} - \underline{\mu}_k)(\underline{x}_j^{(k)} - \underline{\mu}_k)^T = \boxed{\underline{\underline{W}} = (N-1) * \hat{\underline{\underline{\Sigma}}}_k \quad (18)}$$

⊛ Η εξίσωση (17) μπορεί να γραφεί στην μορφή:

$$\boxed{\underline{\underline{R}} \cdot \underline{\underline{W}} \cdot \underline{\underline{R}} = \underline{\underline{0}} \quad (19)}$$

⊛ Invertibility of matrix $\underline{\underline{R}}$ suggests that it is positive definite.

⊛ $\underline{\underline{W}}$ is the sample covariance matrix which is positive semi-definite.

⊛ Equation (19) suggests that:

$$\underline{\underline{R}}\,\underline{\underline{W}}\,\underline{\underline{R}} = \underline{\underline{0}} \Rightarrow \underline{\underline{R}}^{-1}(\underline{\underline{R}}\,\underline{\underline{W}}\,\underline{\underline{R}}) = \underline{\underline{R}}^{-1}\underline{\underline{0}} \Rightarrow$$

$$\underline{\underline{W}}\,\underline{\underline{R}} = \underline{\underline{0}} \rightarrow \underline{\underline{W}} \cdot \underline{\underline{R}} \cdot \underline{\underline{R}}^{-1} = \underline{\underline{0}}\,\underline{\underline{R}}^{-1} \Rightarrow \boxed{\underline{\underline{W}} = \underline{\underline{0}}}$$

⊛ Eq. (19) holds for the trivial $\underline{\underline{W}} = \underline{\underline{0}}$.

⊛ THEREFORE, the DISTANCE CRITERION cannot provide an estimation for the covariance matrix.