

Μηχανική Μάθηση

Δεύτερη προγραμματιστικές ασκήση

Διδάσκων: Μ. Τίτσας

Παράδοση 14 Ιανουαρίου

Η άσκηση αφορά την υλοποίηση του αλγορίθμου EM για την μεγιστοποίηση της πιθανοφάνειας σε μία μίξη Gaussian κατανομών της παρακάτω μορφής

$$p(\mathbf{x}) = \sum_{k=1}^K \pi_k \prod_{d=1}^D \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{1}{2\sigma_k^2}(x_d - \mu_{kd})^2}.$$

Οι εξισώσεις και επαναλήψεις του EM αλγορίθμου περιγράφονται στην «Διάλεξη 7ή». Η μόνη διαφορά είναι ότι, λόγω της μορφής των παραπάνω Gaussian συνιστωσών κατανομών που έχουν σφαιρικό πίνακα συμμεταβλητότητας, οι εξισώσεις ενημέρωσης των πινάκων Σ_k που υπάρχουν στις διαφάνειες απλοποιούνται στις ακόλουθες εξισώσεις ενημέρωσης

$$(\sigma_k^2)^{new} = \frac{\sum_{n=1}^N \sum_{d=1}^D \gamma(z_{nk})(x_{nd} - \mu_{kd}^{new})^2}{D \sum_{n=1}^N \gamma(z_{nk})}, \quad k = 1, \dots, K.$$

Ο υλοποιημένος αλγόριθμος στην συνέχεια θα χρησιμοποιηθεί για την κατάτμηση και συμπίεση εικόνων (δες υποενότητα 9.1.1 του βιβλίου). Ο αλγόριθμος θα εφαρμοστεί για διάφορες τιμές του K (π.χ. $K = 1, 2, 4, 8, 16, 32, 64$) στην εικόνα `im.jpg` που βρίσκεται στο e-class. Για κάθε τιμή του K θα οπτικοποιηθεί η κατατμημένη εικόνα (ομοίως με το Σχήμα 9.3 του βιβλίου) και θα υπολογιστεί το σφάλμα ανακατασκευής της εικόνας (reconstruction error) βάσει της σχέσης

$$error = \frac{1}{N} \sum_{i=1}^N \|\mathbf{x}_{true,i} - \mathbf{x}_{r,i}\|^2,$$

όπου $\mathbf{x}_{r,i}$ είναι η τιμή του pixel που προβλέπεται από την μίξη, δηλ. η τιμή του μέσου μ_k για το οποίο η αντίστοιχη εκ των υστέρων πιθανότητα $\gamma(z_k)$ είναι η μέγιστη, ενώ $\mathbf{x}_{true,i}$ είναι η πραγματική τιμή του pixel.