

## 1 Κάτω φράγμα για τον Perceptron [1 μονάδα]

Δεδομένου ενός αριθμού  $m$ , βρείτε μια ακολουθία από labeled σημεία  $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^3 \times \{-1, +1\}$  για την οποία ο αλγόριθμος Perceptron να κάνει λάθος πρόβλεψη για κάθε σημείο ενώ το άνω φράγμα για τα λάθη που δείξαμε στην τάξη να είναι  $m$ .

**Υπόδειξη:** Ορίστε κάθε  $x_i$  να είναι της μορφής  $(a, b, y_i)$ , όπου  $a^2 + b^2 = m^2 - 1$ . Έστω  $w^*$  το διάνυσμα  $(0, 0, 1)$ . Τώρα, εξετάστε την απόδειξη του άνω φράγματος του Perceptron, βρείτε πού χρησιμοποιήθηκαν ανισότητες ( $\leq$ ) αντί για ισότητες ( $=$ ), και βρείτε σενάρια όπου η ανισότητα ισχύει πραγματικά με ισότητα.

## 2 SGD για Truncated Statistics [2 μονάδες]

Έστω μια μονοδιάστατη κανονική κατανομή  $N(\mu^*, 1)$  με άγνωστη μέση τιμή  $\mu^*$  και διακύμανση 1. Παρατηρούμε δείγματα  $x \sim N(\mu^*, 1)$  όπου μας δίνονται μόνο εφόσον  $x \geq 0$ . Έστω  $N_{\geq 0}(\mu^*, 1)$  η προκύπτουσα κατανομή. Αν είναι γνωστό ότι  $|\mu^*| \leq B$ , δείξτε ότι αν κάνουμε SGD στην συνάρτηση  $f(\mu) = \mathbf{E}_{x \sim N_{\geq 0}(\mu^*, 1)} \left[ -\ln \frac{e^{-\frac{1}{2}(x-\mu)^2}}{\int_0^\infty e^{-\frac{1}{2}(z-\mu)^2} dz} \right]$  με projection στο σύνολο  $[-B, B]$  μπορούμε να εκτιμήσουμε το  $\mu^*$  σε απόσταση  $\varepsilon$ . Πόσα δείγματα χρειαζόμαστε;