Theoretical task 1

due January 19.

Remark: all solutions should be short, mathematically precise and contain proof unless qualitative explanation / intuition is needed. Solutions should be sent electronically to *v.v.kitov@yandex.ru* and can be written in any clear and understandable format - latex, handwritten/scanned or other. Late submissions (by no more than 3 days) will be penalized by 50%, identical solutions will not be graded. The title of your e-mail should be "ICL homework <homework number> - <your first name and last name> "

1. Consider real numbers $z_1, z_2, ... z_N$. Find such constant approximation μ of these numbers, so that

   (a) the sum of square deviations from these points to μ $\sum_{n=1}^{N}(z-\mu)^2$ is minimized.

   (b) the sum of absolute deviations from these points to μ $\sum_{n=1}^{N}|z_n-\mu|$ is minimized.

   Hint: will the functions by convex? why? you may look at the derivative of the minimized criterion.

2. Consider C class classification in D dimensional feature space. Prove that the decision boundary will be piecewise linear for the nearest centroids method.

3. Suppose $x \in \mathbb{R}^D$ is a feature vector. Consider transformation $f = \Sigma^{-1/2}(x-\mu)$, where $\mu = \mathbb{E}x$, $\Sigma = cov[x,x]$, $\Sigma^{1/2} \in \mathbb{R}^{DxD}$ is such matrix that $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$ and $\Sigma^{-1/2} = (\Sigma^{1/2})^{-1}$. Prove that this transformation will give new feature vector f with:

   (a) $\mathbb{E}f = \mathbf{0}$ (all zeroes vector)

   (b) $cov[f,f] = I$ (identity matrix)

4. Consider training set $x_1, ... x_N$ and some linear subspace L_K with lower dimensionality $K \leq D$. Let $x_i = p_i + h_i$ where p_i are projections of x_i onto L_K and h_i are orthogonal complements. Suppose we perform optimization over different K-dimensional subspaces L_K. Prove equivalence of the following two optimization tasks:

   (a) $\sum_{i=1}^{N}\|h_i\|^2 \to \min_{L_K}$

   (b) $\sum_{i=1}^{N}\|p_i\|^2 \to \max_{L_K}$

   Comment: $\|z\|$ is L_2 norm.