Problem Set 4: Due Thursday Week 11

(1) Consider a continuous function $f : \overline{B}_1 \mapsto \mathbb{R}$ which is differentiable in B_1 and such that
$$\int_{B_1} \nabla f \cdot \nabla \phi \, dx = 0 \ \ \forall\, \phi \in C_c^\infty(B_1).$$
Let $K : \mathbb{R}^d \mapsto \mathbb{R}$ be a C^∞, spherically symmetric function such that
$$K(y) \equiv 0 \text{ if } |y| \geq 1, \ \ K(y) \geq 0 \ \ \forall\, y, \int_{\mathbb{R}^d} K(y)\, dy = 1.$$
For $\delta \in (0,1)$ let $K_\delta(y) := \delta^{-d} K(\delta^{-1} y)$, and
$$f_\delta(x) = f * K_\delta(x) \quad \text{in } \{x \in D \mid d(x, \partial D) > \delta\}.$$
Let $\phi \in C^\infty$ be compactly supported in $B_{1-\delta}$, show that
$$\int_{B_1} \nabla f_\delta \cdot \nabla \phi \, dx = \int_{B_1} \nabla f \cdot \nabla \phi_\delta \, dx = 0.$$

(2) Using the conclusion from the previous problem, show that for each $\delta \in (0,1)$ we have f_δ is C^∞ in $B_{1-\delta}$ and harmonic. Furthermore, show f is also C^∞ and harmonic in B_1.

(3) Simulate a two-dimensional Gaussian mixture which is made in equal parts out of 4 different Gaussian distributions whose means are respectivly $(1,1), (1,-1), (-1,1)$, and $(-1,-1)$, the first three of these having common covariance matrix
$$\begin{pmatrix} 0.5 & 0 \\ 0 & 0.5 \end{pmatrix}$$
while the last one has covariance matrix
$$\begin{pmatrix} 0.4 & 0.3 \\ 0.3 & 0.4 \end{pmatrix}.$$
Then

(a) Sample 200 i.i.d. points from this mixture, and record them as $G = \{x_1, \ldots, x_{200}\}$ along with the labels $\{y_1, \ldots, y_{200}\}$ representing which of the Gaussians each point came from.

(b) Create a 200×200 matrix W whose i,j entry is given by (σ TBD)
$$w_{ij} = e^{-\frac{1}{\sigma}|x_i - x_j|^2}$$

(c) Use a linear algebra package to compute the first 10 eigenvalues of the (combinatorial) Laplacian given by W, for four different values of σ of your choosing.

(d) For $m = 1,2,3,4$ define $f : G \mapsto \mathbb{R}$ be defined by
$$f_m(x_i) = \begin{cases} 1 \text{ if } y_i = m \\ 0 \text{ otherwise} \end{cases}$$
For each of the σ's used above, compute the Dirichlet energy of f_4, that is
$$J(f_4) = \sum_{i,j=1}^{200} w_{ij} |f_4(x_i) - f_4(x_j)|^2.$$

(e) Let $D = \{x_{51}, \ldots, x_{200}\}$ so that $G \setminus D = \{x_1, \ldots, x_{50}\}$, use a linear algebra package to compute the unique function $\hat{f}_4 : G \mapsto \mathbb{R}$ which solves
$$\begin{cases} \Delta \hat{f}_4 & = 0 \text{ in } D, \\ \hat{f}_4 & = f_4 \text{ in } G \setminus D. \end{cases}$$
For a point in D, guess its label as 4 if at this point $\hat{f}_4 > 1/2$, and as "unknown" otherwise. Compute the total number of mislabeled entries in D for each of the four σ's chosen above.