GI07/COMP012 – Assignment (due: Noon, November 21, 2014)

Please handle your assignment to CS Reception, Room 5.25, Malet Place Engineering Building.

Note: Presentation, clarity, and synthesis of exposition will be taken into account in the assessment of the following exercises.

This document is available at: `http://www.cs.ucl.ac.uk/staff/M.Pontil/courses/homeworkgi07.pdf`

1. **[30 pts]** (*Vector and matrix operations*). Please provide a short, yet justified, answer to the following questions.

   (a) Let a be a real parameter and consider the linear system of equations:

   $$\begin{aligned} x + y &= 1 \\ ax + y &= 0 \end{aligned}$$

   For which value(s) of a is there a solution to the above equations?
   For which value(s) of a is the solution unique?

   (b) Consider the matrix

   $$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$

   This matrix is
   [ ]: symmetric
   [ ]: invertible
   [ ]: singular

   (c) The product between an $n \times k$ matrix and a $k \times \ell$ matrix is a
   [ ]: $n \times \ell$ matrix
   [ ]: $k \times k$ matrix
   [ ]: this matrix product is not well defined because in general $n \neq \ell$.

   (d) The identity matrix
   [ ]: has all of its elements equal to 1
   [ ]: has determinant equal to -1
   [ ]: is a diagonal matrix, whose diagonal elements are all equal to 1

   (e) Consider the vector $x = (-3, 0, 5)$. Compute the p-norm of x for $p = 1, 2, \infty$. Observe that

   $$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1.$$

   Argue that these inequalities are true for every vector x. Give an example of a vector where these inequalities are all tight.

2. **[20 pts]** (*Projections*).

   (a) Argue that if P is an orthogonal projection, all of its singular values are either 1 or 0.

   (b) Let I be the identity matrix. Show that if P is an orthogonal projection, then $I - 2P$ is an orthogonal matrix.

   (c) Compute the projection P which maps the point $(1, 1)$ to $(0, 0)$ and the point $(0, 1)$ to $(0, 1)$. Is P an orthogonal projection?

3. **[10 pts]** (*Graph Laplacian*).

   Consider a graph with six vertices, consisting of two cliques, $\{v_1, v_2, v_3\}$ and $\{v_4, v_5, v_6\}$, which are connected by a single edge, joining v_3 to v_4. Compute the corresponding graph Laplacian.

4. **[20 pts]** (*Ridge regression*)
Consider the ridge regression problem: $(w_\lambda, b_\lambda) = \text{argmin}_{w,b} \left\{ \sum_{i=1}^m (y_i - b - w^\top x_i)^2 + \lambda w^\top w \right\}$. Show that this problem is equivalent to the problem

$$(\hat{w}_\lambda, \hat{b}_\lambda) = \text{argmin}_{w,b} \left\{ \sum_{i=1}^m \big(y_i - b - w^\top (x_i - \bar{x})\big)^2 + \lambda w^\top w \right\}$$

where $\bar{x}$ is the average of the input data. Give the correspondence between $(\hat{w}_\lambda, \hat{b}_\lambda)$ and (w_λ, b_λ).

5. **[20 pts]** (*convex functions*)
Let $f : \mathbb{R}^d \to \mathbb{R}$ be a *convex* function such that

$$f(w) \geq 0 \qquad \text{for every } w \in \mathbb{R}^d .$$

Show that the function $g : \mathbb{R}^d \to \mathbb{R}$ defined as

$$g(w) = \big(f(w)\big)^2 \qquad \text{for every } w \in \mathbb{R}^d$$

is also convex. As a special case deduce that, for every norm $\|\cdot\|$, the function

$$w \mapsto \|w\|^2$$

is convex.