APM462H1S, Winter 2014 , Assignment 1,

due: Monday February 3

Exercise 1. To approximate a function $g : [0, 1] \to R$ by a nth order polynomial, one can minimize the function f defined by

$$f(a) = \int_0^1 (g(x) - p_a(x))^2 \, dx$$

where, for $a = (a_0, \ldots, a_n) \in E^{n+1}$, we use the notation

$$p_a(x) = a_0 + a_1 x + \cdots + a_n x^n = \sum_{k=0}^{n} a_k x^k.$$

a. Show that f can be written in the form

$$f(a) = a^T Q a - 2b^T a + c$$

for a $(n+1) \times (n+1)$ matrix Q, a vector $b \in E^{n+1}$, and a number c. Find formulas for Q, b and c. It should be clear from your formula that Q is symmetric.

b. Find the first-order necessary condition for a point $a^* \in E^{n+1}$ to be a minimum point for f.

c. Find *all* minimizing points $a^* \in E^{n+1}$ when g is the constant function $g(x) \equiv 0$. Prove that your answer is correct.

Hint: keep in mind that you have two formulas for $f(a)$ – the formula used to define f, and the one you have found in part (a) above. Take a moment to think about which formula is more useful here.

d. Is Q positive semidefinite? positive definite? justify your answer.

Discussion: one way to do this is to find all the eigenvalues of Q and see whether they are all nonnegative/positive. If you try to do this, you will see that it is basically impossible. So I suggest that you try to think about it more abstractly, taking into account anything you know about the function f that might be useful. For example, the definition of f implies that $f(a) \geq 0$ for all a.....

If you do not completely solve this problem, don't worry – it is meant to be a little harder and in any case is only worth a few marks.

Exercise 2. Find a global minimum point for the function

$$f(x, y, z) = 2x^2 + xy + y^2 + yz + z^2 - 6x - 8y - 8z + 9,$$

and *prove* that your solution really is a global minimum for f.

Exercise 3. Assume that g is a convex function on E^n, that f is a convex function of a single variable, and in addition that f is a nondecreasing function (which means that $f(r) \geq f(s)$ whenever $r \geq s$).

a. Show that $F := f \circ g$ is convex by directly verifying the convexity inequality

$$F(\theta x + (1 - \theta)y) \leq \theta F(x) + (1 - \theta)F(y).$$

Explain where each hypothesis (convexity of g, convexity of f, and the fact that f is nondecreasing) is used in your reasoning. (The notation $F = f \circ g$ means that $F(x) = f(g(x))$.)

b. Now assume that f and g are both C^2. Express the matrix of second derivatives $\nabla^2 F(x)$ in terms of f and g, and prove directly (without using part (a)) that $\nabla^2 F$ is positive semidefinite at every x.

discussion: The half of part (b) – that is, express $\nabla^2 F$ in terms of f and g – is basically an exercise in using the chain rule for functions of several variables. If you find it at all difficult, then **review the chain rule** until you have completely mastered it!

For the second half of part (b) – showing that $\nabla^2 F$ is positive semidefinite – please explain again, as you did in part (a), where each hypothesis is used in your reasoning.

Exercise 4. Verify that if $f_1, \ldots, , f_k$ are convex functions on E^n, then

$$g(x) := \max(f_1(x), \ldots, f_k(x))$$

is also convex.