ECES 811 Homework 5 & 6

1. Consider the function[1]
$$f(x, y) = (y - x^2)(y - 3x^2) \tag{1}$$
   (a) Calculate the gradient and Hessian matrix of this function.
   (b) Evaluate the gradient and the Hessian at $(x, y) = (0, 0)$.
   (c) Prove that, when restricted to any line $y = mx$ through the origin, the function has a strict local minimum at $x = 0$.
   (d) Plot this function in a small neighborhood around $(0, 0)$. Does this function have a local minimum at $(0, 0)$? Prove your answer by considering instead restrictions of the function to curves of the form $y = mx^2$ for various m.

2. Consider the Monkey Saddle
$$f(x, y) = x^3 - 3xy^2 \tag{2}$$
   (a) Calculate the gradient and Hessian matrix of this function.
   (b) Evaluate the gradient and the Hessian at $(x, y) = (0, 0)$.
   (c) Plot this function near $(x, y) = (0, 0)$. What is significantly different about this type of saddle point and a saddle point associated with a Hessian matrix with one eigenvalue positive and one eigenvalue negative?

3. Consider set of probability distributions p_X for a random variable X taking values in a finite set $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$, parameterized via the vector $\mathbf{p} = [p_X(x_1), p_X(x_2), \ldots, p_X(x_N)]^T$.
   (a) Show that the set of such parameter vectors $\mathbf{p}$ is a polyhedron $\mathcal{P}$ and determine this polyhedron.
   (b) The Shannon entropy of the random variable X, which measures the amount of randomness in the random variable, is defined as
$$H(X) = -\sum_{x \in \mathcal{X}} p_X(x) \log(p_X(x)) \tag{3}$$
   with the convention that $0 \log(0)$ is 0, and through a standard abuse of notation the entropy is written as a function of the random variable even though it is a function of only its distribution. Show that the Shannon entropy is a concave function of $\mathbf{p}$.
   (c) Consider the problem of maximizing the Shannon entropy of the random variable $\mathcal{X}$ subject to the constraints
$$\mathbb{E}[F_1(X)] = f_1, \ldots, \mathbb{E}[F_k(X)] = f_k \tag{4}$$
      i. Show that this may be formulated via a convex optimization problem (i.e. convex objective, with any equality constraints linear, and any inequality constraints convex).
      ii. Using the KKT conditions, what form must the p_X having the maximum entropy subject to these constraints take?
      iii. Use an extension of this conclusion to deduce a special property of exponentially distributed random variables and normally distributed random variables. For this, it is useful to note the definition of differential entropy of a continuous random variable with PDF p_X is
$$h(X) = -\int_{\mathcal{X}} p_X(x) \log(p_X(x)) dx \tag{5}$$

4. A certain power distribution company can buy electric power from N different power generation companies. The cost of x_n units of power bought from company n is $f_n(x_n)$ dollars. Due to legislation regarding alternative energy requirements and the abilities of the generation equipment, company n can only produce between a_n and b_n units of power, i.e. $a_n \leq x_n \leq b_n$. At a particular time, there is demand for a total of D units of power.
   (a) Write the optimization problem for finding the minimum operating cost for the distribution company at this demand.

[1]Example taken from *Counterexamples in Analysis*, Gelbaum & Olmsted.

(b) Determine the analytical form of the Lagrangian dual function for this problem by defining the domain $\mathcal{X} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_N$ with $\mathcal{X}_i = [a_i, b_i]$ and dualizing the demand constraint. Express the resulting Lagrangian dual is the sum of N functions of a single Lagrange multiplier variable μ.

(c) Consider the special case that $f_i(x_i) = \sqrt{x_i} + .2\sin(x_i)$ and $a_i = 0,\ b_i = 2\pi$. Use the first order necessary conditions together with numerical minimization over the stationary points to find the exact solution to this problem for $N = 10, 100, 500$ as a function of the demand D.
(plot f^* as a function of $D/(2\pi N)$ in Matlab for $N = 10, 100, 500$)

(d) Evaluate and plot the relative duality gap $(f^* - g^*)/f^*$ for the previous problem for $N = 10, 100, 500$ as a function of $D/(2\pi N)$ in Matlab.

5. Consider the problem of maximizing the rate of communication over N parallel additive white Gaussian noise channels with a total power budget of P.

$$\max_{\mathbf{p}=(p_1,\ldots,p_N)|\mathbf{p}\geq 0, \mathbf{1}^T\mathbf{p}\leq P} \sum_{k=1}^{N} \log_2(1 + \frac{p_k}{N_k}) \tag{6}$$

(a) Find the Lagrangian dual function for this problem. Is there a duality gap?

(b) Analytically solve this optimization problem.