Continuous Maths HT 2019: Problem Sheet 2

Optimization, Convexity, Langrage Multipliers, and Numerical Integration

2.1 Find and classify the stationary points of these functions:

(a) $g(x,y) = \exp(x^3 - 3x)\exp(y^3 - 3y)$,

(b) $h(x,y,z) = \frac{1}{2}x^2 + y^2 + \frac{27}{2}z^2 - xy + xz + 4yz - x - y - 11z$.

2.2 Let $\boldsymbol{a}$ be a vector with $a_i > 0$, $\sum_i a_i = 1$. Using the facts in the lecture notes (or otherwise), show that that matrix $\boldsymbol{\Delta_a} - \boldsymbol{a}\boldsymbol{a}^T$, where $\boldsymbol{\Delta_a}$ is the diagonal matrix with $\boldsymbol{a}$ on the diagonal, is positive semidefinite.

Deduce that the log-sum-exp function, $l(\boldsymbol{x}) = \ln\left(\sum_{i=1}^n \exp(x_i)\right)$, is convex.

2.3 Here is an application from economics. A factory makes a product out of three infinitely-divisible ingredients X, Y, and Z. The ingredients cost p, q, and r pounds per kg, respectively, with $p, q, r > 0$ all different, and the factory has a budget of $B > 0$ pounds. The value of the product is some function $g(x,y,z)$ of the weight of the ingredients.

(a) Formulate, as constrained optimization in standard form, the problem that the factory wishes to maximize the value of their product, within their budget.

(b) Solve the problem in the case when $g(x,y,z) = xyz$. Be careful not to divide by zero, or assume that a constraint is slack when it might be tight (or vice versa).

(c) Now solve in the case when $g(x,y,z) = \exp(x+y+z)$.

2.4 Here is an example from image retrieval and machine learning. We have objects $O_1, \ldots, O_n$ for which we have computed some measure of pairwise distance (or *weights*), $w_{ij} = w_{ji}$ denoting the distance between O_i and O_j. We wish to assign each O_i a scalar x_i such that objects that have a low distance are assigned scalars that are close. After fixing scale and location, we arrive at the constrained optimization problem

$$\underset{x_1,\ldots,x_n}{\text{minimize}} \; E(\boldsymbol{x}) = \sum_i \sum_j w_{ij}(x_i - x_j)^2 \text{ subject to } \sum_i x_i^2 = 1, \; \sum_i x_i = 0.$$

(a) Show that another way to write the objective function is

$$E(\boldsymbol{x}) = 2\boldsymbol{x}^T(\mathbf{D} - \mathbf{W})\boldsymbol{x},$$

where $\mathbf{W}$ is the symmetric matrix (w_{ij}), and $\mathbf{D}$ is a diagonal matrix that satisfies $\mathbf{1}^T\mathbf{D} = \mathbf{1}^T\mathbf{W}$. Also write the constraints as vector equations.

(b) By considering the original problem, show that $\mathbf{D} - \mathbf{W}$ is positive semidefinite.

(c) Show that the first-order condition, for a stationary point of the Lagrangian, is

$$4(\mathbf{D} - \mathbf{W})\boldsymbol{x} - 2\lambda\boldsymbol{x} - \mu\mathbf{1} = \mathbf{0}. \quad (*)$$

By applying $\mathbf{1}^T$ to $(*)$, show that $\mu = 0$.

(d) Show that every solution $\boldsymbol{x}$ of $(*)$ is an eigenvector of $\mathbf{D} - \mathbf{W}$, and that as long as the constraints are satisfied $E(\boldsymbol{x}) = 2\nu$, where ν is the corresponding eigenvalue. Show also that $\mathbf{1}^T\boldsymbol{x} = 0$ unless $\nu = 0$.

(e) Deduce that the solution to the optimization problem is the unit eigenvector of $\mathbf{D} - \mathbf{W}$ with the *second*-smallest eigenvalue.

2.5 Let $f(x) = \exp(\exp x)$. Using the error bounds in the lecture notes, find the smallest number of strips guaranteed to find $\int_0^1 f(x)\,\mathrm{d}x$ to within 10^{-6} using (a) the Midpoint rule, (b) the Trapezium rule, and (c) Simpson's rule.

2.6 We will derive another method of numerical integration. For one strip $[0, 2l]$, we approximate $f(x)$ by its third-order Taylor polynomial about $x = l$:

$$\hat{f}_3(x) = f(l) + (x-l)\frac{\mathrm{d}f}{\mathrm{d}x}(l) + \frac{(x-l)^2}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(l) + \frac{(x-l)^3}{6}\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}(l).$$

Then we approximate $\int_0^{2l} f(x)\,\mathrm{d}x$ by $A_1[f, 0, 2l] = \int_0^{2l} \hat{f}_3(x)\,\mathrm{d}x$.

(a) Verify that $A_1[f, 0, 2l] = 2lf(l) + \frac{1}{3}l^3\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(l)$.

(b) Use Taylor's theorem to bound $\hat{f}_3(x) - f(x)$ in terms of x, l,

$$\underline{D_4} = \min_{\xi\in(0,2l)}\left|\tfrac{\mathrm{d}^4 f}{\mathrm{d}x}(\xi)\right|, \quad \text{and} \quad \overline{D_4} = \max_{\xi\in(0,2l)}\left|\tfrac{\mathrm{d}^4 f}{\mathrm{d}x}(\xi)\right|.$$

(c) Derive upper and lower bounds on the error of the single-strip integral

$$\mathrm{err}(A_1)[f, 0, 2l] = A_1[f, 0, 2l] - \int_0^{2l} f(x)\,\mathrm{d}x,$$

in terms of l, $\underline{D_4}$, and $\overline{D_4}$.

(d) Derive the corresponding composite method that uses n equal-sized strips to approximate $\int_a^b f(x)\,\mathrm{d}x$, and give error bounds for the composite rule.

(e) How does this method compare with Simpson's rule, in terms of accuracy and speed?

2.7 Compute $\int_0^1 x^2\,\mathrm{d}x$ exactly, and then estimate it using the Midpoint and Trapezium rules, using two strips. Compute $\int_0^1 x^4\,\mathrm{d}x$ exactly, and then estimate it using Simpson's rule and the method of question **2.6**, using two strips.

Compare the error of your approximations with the error bounds derived in the lecture notes and question **2.6**. What do these results tell you?

2.8 Write a program to estimate $\int_0^2 x^{3/2}\,\mathrm{d}x$ using Simpson's rule with n strips, for $n = 2^1, 2^2, 2^3, \ldots, 2^{24}$ (or n up to as large as your computer can handle). Draw a table or plot a graph of the observed error: the results are not quite as predicted by the error bound in the lecture notes. Explain this phenomenon.

`adk@cs.ox.ac.uk`, HT 2019