Math 4630/6630 - Nonlinear Optimization - Spring 2021

Questions for Topic 1

1A. (a) Formulate the problem of finding the best line $y = a_0 + a_1 x$ through the the points $(1, 1)$, $(2, 3)$ and $(4, 6)$ as an optimization problem using the least squares objective function.

(b) Linear least squares problems, of which (a) is an example, can be solved by taking the gradient of the objective function and setting it equal to 0. This gives a system of linear equations that can be solved to find the optimum values of the coefficients. Solve (a) in this way. (The unknowns here are a_0, a_1, so take the partial derivatives of your objective function with respect to each of these, and set them equal to 0.)

1B. (a) Formulate the problem of finding the best even quadratic (quadratic symmetric about the y-axis) $y = a_0 + a_2 x^2$ through the the points $(0, 7)$, $(1, 6)$ and $(2, 1)$ as an optimization problem using the least squares objective function.

(b) Linear least squares problems, of which (a) is an example, can be solved by taking the gradient of the objective function and setting it equal to 0. This gives a system of linear equations that can be solved to find the optimum values of the coefficients. Solve (a) in this way, giving an *exact* solution. (The unknowns here are a_0, a_2, so take the partial derivatives of your objective function with respect to each of these, and set them equal to 0.)

1C. Find the critical point of the quadratic $f(x, y) = 3x^2 + 10xy + 8y^2 + 4x + 6y$ and use definiteness properties of the Hessian matrix to say as much as you can about whether the critical point is a local minimizer, local maximizer or saddle point.

1D. Find the critical point of the quadratic $f(x, y) = 2x^2 - 8xy + 10y^2 - 4y + 19$ and use definiteness properties of the Hessian matrix to say as much as you can about whether the critical point is a local minimizer, local maximizer or saddle point.

1E. (Characterizations of positive definiteness) The following two properties are different ways of characterizing whether a real symmetric $n \times n$ matrix A is positive definite. Property 1 is usually used as the definition. In this question we will show that Property 2 is equivalent to Property 1, using some fundamental properties of real symmetric matrices.

Property 1: $v^{\mathrm{T}} A v > 0$ for all $v \in \mathbf{R}^n$, $v \neq 0$.

Property 2: All eigenvalues of A are positive (remember that $\lambda \in \mathbf{R}$ is an eigenvalue of A if there is $v \in \mathbf{R}^n$, $v \neq 0$ with $Av = \lambda v$).

(a) (i) Show that Property 1 implies Property 2.

(ii) Use orthogonal diagonalization, below, to show that Property 2 implies Property 1.

Orthogonal diagonalization: If A is a real symmetric matrix, then $A = U^{\mathrm{T}} D U$, where U is a real orthogonal matrix ($U^{-1} = U^{\mathrm{T}}$) and D is a real diagonal matrix with the eigenvalues of A down the diagonal.

(b) There is a third property that characterizes being positive definite, i.e., it is equivalent to Properties 1 and 2.

Property 3: All of $\det A_1, \det A_2, \ldots, \det A_n$ are positive, where A_k is the $k \times k$ matrix obtained by taking just the first k rows and the first k columns of A.

[Comment: Property 3 can be checked efficiently. It is the source of the usual second order condition for a function $f(x, y)$ of two variables to have a local minimizer: $\nabla f = 0$, $f_{xx} = \det A_1 > 0$ and $f_{xx} f_{yy} - f_{xy}^2 = \det A_2 > 0$, where $A = \nabla^2 f$: this shows $A = \nabla^2 f$ is positive definite.]

Use Property 3 to determine whether or not each of the following matrices is positive definite.

$$\text{(i)} \begin{bmatrix} 3 & 2 & 1 \\ 2 & 5 & 4 \\ 1 & 4 & 6 \end{bmatrix}, \qquad \text{(ii)} \begin{bmatrix} 4 & 2 & 0 \\ 2 & 3 & -2 \\ 0 & -2 & 2 \end{bmatrix}.$$

1F. (a) Suppose $A_i, i \in \mathcal{I}$, is a family of convex sets. Prove that the intersection $\cap_{i \in \mathcal{I}} A_i$ is also convex.

(b) Give an example of two convex sets whose union is not convex.

1G. Let $r \geq 0$. Prove that the set $S = \{x \in \mathbf{R}^n \mid \|x\| \leq r\}$ is a convex set, where $\|x\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$ is the usual length of a vector. (Hint: use the properties that $\|\alpha x\| = |\alpha| \, \|x\|$ for a scalar α and vector x, and the triangle inequality $\|x + y\| \leq \|x\| + \|y\|$ for vectors x and y.)

1H. A function $f : \mathbf{R}^n \to \mathbf{R}$ is *affine* if $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$ whenever $x, y \in \mathbf{R}^n$ and $\alpha, \beta \in \mathbf{R}$ with $\alpha + \beta = 1$. (Note that there is no other restriction on the value of α or β; in particular we are not required to have $\alpha \in [0, 1]$ or $\beta \in [0, 1]$.)

Prove that if f is both convex on $\mathbf{R}^n$ and concave on $\mathbf{R}^n$, then f is affine. (Note that you will have to treat the cases $\alpha < 0$ and $\alpha > 1$ differently from $\alpha \in [0, 1]$.)

1I. Suppose f is a convex function on $\mathbf{R}^n$. Prove that the set S of global minimizers of f is a convex set.

1J. Show that each of the following functions is convex on the specified convex set. If you can show that the function is strictly convex on the interior of the given set, mention that also.

(a) $f(x_1, x_2) = 5x_1^2 + 2x_1x_2 + x_2^2 - x_1 + 2x_2 + 3$ on $S = \mathbf{R}^2$.

(b) $f(x_1, x_2) = x_1^2/2 + 3x_2^2/2 + \sqrt{3}x_1x_2$ on $S = \mathbf{R}^2$.

(c) $f(x_1, x_2) = -\sin x_1 \sin x_2$ on $S = \{(x_1, x_2) \in \mathbf{R}^2 \mid \frac{\pi}{4} \leq x_1 \leq \frac{3\pi}{4}, \frac{\pi}{4} \leq x_2 \leq \frac{3\pi}{4}\}$. (It's probably easiest to use the characterization of positive definiteness in terms of determinants of submatrices.)

1K. Use properties of the Hessian matrix to determine whether $f(x_1, x_2) = 4x_1^2 + 12x_1x_2 - x_1 + 9x_2^2 - x_2 + 4$ is convex, concave, both, or neither on $\mathbf{R}^2$. Explain whether or not you can determine strict convexity or strict concavity from the Hessian matrix.

1L. Suppose that $a > 1$. Consider the sequence $x_0, x_1, x_2, \ldots$ defined by

$$x_0 = a, \qquad x_{k+1} = \frac{1}{2}(x_k + \frac{a}{x_k}) \text{ for } k \geq 0.$$

It can be shown that this is a decreasing sequence that converges to a positive limit.

(a) Given that the limit exists and is positive, prove that $\lim_{k \to \infty} x_k = \sqrt{a}$.

(b) Use the limit from (a) and the fact that the sequence is decreasing to prove that the sequence converges (Q-)quadratically.

1M. If $p > 0$, then the sequence $x_k = 1/k^p$ converges to 0. Show that it does *not* converge (Q-)linearly. (This is known as *sublinear convergence*.)