CIE 6010 / MDS 6118 Midterm Exam (F2018)

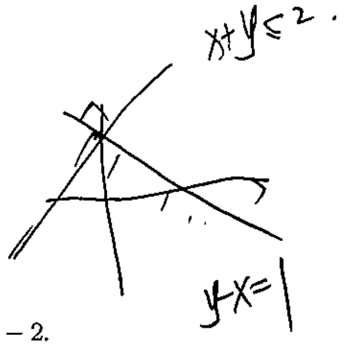
Part 1: In-class. November 7. 10:30am–12:00

— Room F302, Shaw College —

Instructions

- The exam consists of 2 related parts:
  Part 1. Written questions/derivations. (CIE 75 / MDS 65 points)
  Part 2. Problem-solving in Matlab. (CIE 25 / MDS 35 points)
- This Part 1 is closed to books and notes (or any outside resources). One single cheat sheet is allowed. With every answer, please include a concise derivation, unless otherwise specified, (a correct answer without a valid justification might not receive credits).
- Problems designated as "CIE only" are required only for CIE 6010 students, and vice versa for MDS 6118 students. Problems with no designations are for both groups. Please only do required problems.

Part 1: Written Problems (CIE 75 / MDS 65 points)



Problem 1 (15 points)

Consider the nonlinear program

$$\min_{x \in \mathbb{R}^n} \{f(x) : h(x) = 0,\ g(x) \leq 0\}.$$

with $f(x) = (x_1 - 1)^2 + x_2 - 2$, $h(x) = x_2 - x_1 - 1$, and $g(x) = x_1 + x_2 - 2$.

(a) Write down the KKT conditions for this problem (no derivation necessary).

(b) Find a KKT point and the associated multipliers.

(b) Specify the type of the KKT point found (local or global minimum/maximum, saddle point, ... etc) with a clear justification.

$\langle \nabla^T$

Problem 2 (30 points)

Given a complex matrix $C = A + iB \in \mathbb{C}^{m\times n}$ ($m > n$) with real part $A \in \mathbb{R}^{m\times n}$ and imaginary part $B \in \mathbb{R}^{m\times n}$, for a unknown "signal" $\hat{x} \in \mathbb{R}^n$, one can observe noisy data $d \approx |C\hat{x}|^2 \in \mathbb{R}^m$, which is the squared magnitude, taken component-wise, of the complex vector $C\hat{x}$.

To recover the unknown signal $\hat{x}$ from a noisy observation $d > 0$, we consider solving the nonlinear least squares problem:

$$\min_{x\in\mathbb{R}^n} f(x) \triangleq \frac{1}{2}\|r(x)\|^2 \tag{1}$$

with the residual

$$r(x) = |Cx|^2 - d = |(A+iB)x|^2 - d. \tag{2}$$

Recall the general formulas for nonlinear least squares problems

$r^T r$

$$\nabla f(x) = J(x)^T r(x), \quad \nabla^2 f(x) = J(x)^T J(x) + S(x),$$

r

where $J(x)^T = [\nabla r_1(x) \cdots \nabla r_m(x)]$ and $S(x) = \sum_{i=1}^{m} r_i(x)\nabla^2 r_i(x)$. Let

$-J^T J$.

$J^{-1}(J^T)^{-1}J^T r(x)$

$$A^T = [a_1 \cdots a_m] \quad \text{and} \quad B^T = [b_1 \cdots b_m],$$

i.e., a_i^T and b_i^T are the rows of A and B, respectively; and also define the matrices

$$E_i = a_i a_i^T + b_i b_i^T \in \mathbb{R}^{n\times n}, \quad i = 1, 2, \cdots, m. \tag{3}$$

$n\times m$

Under this notation, $r_i(x) = (a_i^T x)^2 + (b_i^T x)^2 - d_i = x^T E_i x - d_i$.

$J = m\times n$

Answer the following questions with brief and clear derivations. In your answers, the formulas should be expressed *in terms of the specified quantities only* (for instance, the uses of $\nabla r_i(x)$, $\nabla^2 r_i(x)$ or $J(x)$ are not allowable). We will assume in general that (a1): $J(x)$ has rank n at any nonzero x, and (a2): A or B has rank n.

1. Derive a formula for $J(x)^T J(x)$ in terms of $\{E_i\}$ and x only.
2. Derive a formula for $S(x)$ in terms of $\{E_i\}$, $\{r_i(x)\}$ and x only.
3. Derive a formula for $\nabla f(x)$ in terms of $S(x)$ and x only?
4. Find an obvious stationary point x^*. What does $\nabla^2 f(x^*)$ tell you about this stationary point and why?
5. At a given x, write down the linear systems for computing the Gauss-Newton direction and the Newton direction, respectively. Is each of these two directions necessarily a descent direction? Briefly explain why or why not.

$x^+ = x - \alpha \nabla f(x)$.

$x^+ = x - \alpha[$

$d_{GN} = -[\nabla^2 f(x)]^{-1} \nabla f(x))$.

Problem 3 (20 points)

(a) (10 points) Prove that $f : \mathbb{R}^n \to \mathbb{R}$ is convex function if and only if its epigraph,

$$\text{epi}(f) = \{(x, t) : t \geq f(x)\},$$

is a convex set in $\mathbb{R}^{n+1}$.

(b) (MDS only. 10 points) A convex optimization problem is to minimize a convex function, or to maximize a concave function, in a convex set. Prove that a convex minimization problem cannot have a local minimum that is not a global minimum.

(c) (CIE only. 10 points) For $f : \mathbb{R}^n \to \mathbb{R}$, the conjugate function of f is $f^* : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f^*(y) = \sup_{x \in \text{dom}(f)} (y^T x - f(x)).$$

The domain of f^*, $\text{dom}(f^*)$, is the set in $\mathbb{R}^n$ on which $f^*(y) < +\infty$.

(2 pts) Is f^* convex for any f? Why or why not.
(4 pts) Let $f(x) = \frac{1}{2}x^T Q x$ where $Q = Q^T \in \mathbb{R}^{n \times n}$ is positive definite. $f^*(y) = ?$
(4 pts) Let $f(x) = e^x$ for $x \in \mathbb{R}$. $f^*(y) = ?$ $\text{dom}(f^*) = ?$

Problem 4 (10 points. CIE only)

Consider the nonlinear program: $f : R^n \to \mathbb{R}$, $h : R^n \to \mathbb{R}^m$, $g : R^n \to \mathbb{R}^r$,

$$p^* = \min_{x \in \mathbb{R}^n} \{f(x) : h(x) = 0,\ g(x) \leq 0\}, \tag{4}$$

where p^* denotes the global minimum. The Lagrangian is $L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$. Define the dual program:

$$d^* = \max_{\lambda, \mu} \{Q(\lambda, \mu) : \mu \geq 0\},$$

where $Q(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu)$. For simplicity, let us assume that both p^* and d^* are finite.

(a) Explain why the dual program is always a convex optimization problem (in this case, to maximize a concave function in a convex set).

(b) Prove the weak duality $d^* \leq p^*$.