First-Stage PhD Comprehensive Examination

in

CONTINUOUS OPTIMIZATION

Department of Combinatorics and Optimization, Faculty of Mathematics,
University of Waterloo

MC 6486, Monday, June 16, 2014, 1:00p.m. – 4p.m. **(3 hours)**

Examiners: T. Coleman and L. Tunçel

1. (a) Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed convex set and let $p(C, u)$ denote the closest point to $u \in \mathbb{R}^n$, in C (with respect to the Euclidean norm). Based on the above definition, set up a convex optimization problem whose unique solution is $p(C, u)$. Then utilizing a suitable theorem (characterizing minimizers of a convex function over a convex set), prove that for every $u \in \mathbb{R}^n \setminus C$,

$$[u - p(C, u)]^\top [x - p(C, u)] \leq 0, \forall x \in C.$$

   (b) Let $C \subseteq \mathbb{R}^n$ be a nonempty, closed convex set and $p(C, u)$ be as above. Prove that for every $u, v \in \mathbb{R}^n$,

$$\|p(C, u) - p(C, v)\|_2 \leq \|u - v\|_2.$$

   (c) Let $C \subset \mathbb{R}^n$ be a nonempty, compact convex set. Considering (and utilizing) the *farthest point problem* (the problem of finding a point in C with maximum distance from the origin), prove that C has at least one extreme point.

2. (a) State the Farkas' Lemma for the system

$$\text{(I)} \qquad Ax = b, x \geq 0,$$

   where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$.

   (b) Using the statement from part (a), prove that for every triple (A, B, c),with $A \in \mathbb{R}^{p \times n}$, $B \in \mathbb{R}^{q \times n}$, $c \in \mathbb{R}^n$, exactly one of the following systems has a solution:

$$\text{(I)} \qquad \exists d \in \mathbb{R}^n \text{ such that } Ad \leq 0, Bd = 0, c^\top d > 0;$$
$$\text{(II)} \qquad \exists \lambda \in \mathbb{R}^p_+, \mu \in \mathbb{R}^q \text{ such that } A^\top \lambda + B^\top \mu = c.$$

   (c) State the *Hyperplane Separation Theorem* for a closed convex set S in $\mathbb{R}^n$ and a point $u \in \mathbb{R}^n \setminus S$.

   (d) Using the statement in part (c), re-prove the statement in part (b).

3. (a) Given $f : \mathbb{R}^n \to [-\infty, +\infty]$, define what is meant by the *Legendre-Fenchel conjugate* f^* of function f.

   (b) Compute the Legendre-Fenchel conjugate of $\lambda_{\max}(X)$, where X is an n-by-n symmetric matrix with real entries, and $\lambda_{\max}$ denotes the largest eigenvalue function. Prove all your claims.

   (c) Compute the subdifferential of $\lambda_{\max}(X)$. Prove all your claims.

   (d) Compute the Legendre-Fenchel conjugate of

$$-\ln(\det(X)) : \mathbb{S}^n_{++} \to \mathbb{R},$$

   where $\mathbb{S}^n_{++}$ denotes the cone of n-by-n symmetric positive definite matrices, with real entries.

   (e) Compute the Lagrangian dual of the following problem

$$(P) \quad \inf \left\{ \mathrm{Tr}(CX) : \mathcal{A}(X) = b, X \in \mathbb{S}^n_+ \right\},$$

   where $C \in \mathbb{S}^n$, $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$, $b \in \mathbb{R}^m$ are given, $\mathbb{S}^n$ denotes the space of n-by-n symmetric matrices with real entries, $\mathbb{S}^n_+$ denotes the set of positive semidefinite matrices in $\mathbb{S}^n$, $\mathrm{Tr}(X)$ is the trace of the matrix X.

4. Let $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^p$, $h : \mathbb{R}^n \to \mathbb{R}^q$ be given. Consider

$$(P) \quad \begin{array}{rl} \inf & f(x) \\ \text{subject to:} & g(x) \leq 0 \\ & h(x) = 0. \end{array}$$

   (a) State the Karush-Kuhn-Tucker (KKT) theorem for (P) (including all the necessary assumptions on f, g and h).

   (b) Recall that $\mathbb{S}^n$ denotes the space of n-by-n symmetric matrices with real entries and $\mathbb{S}^n_{++}$ denotes the set of positive definite matrices in $\mathbb{S}^n$. Given $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ a linear transformation satisfying $\mathcal{A}(I) = 0$, consider the following optimization problem:

$$(P_0) \quad \begin{array}{rl} \inf & -\ln(\det(X)) \\ \text{subject to:} & \mathcal{A}(X) = 0 \\ & \mathrm{Tr}(X) = n \\ & X \in \mathbb{S}^n_{++}. \end{array}$$

   (c) Prove that (P_0) has a unique optimal solution.

   (d) State the strongest version of KKT Theorem you can for (P_0).

   (e) What is the unique optimal solution of (P_0)? Prove your claim using the KKT theorem from part (d).

5. Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a continuously differentiable system. Assume $B \in \mathbb{R}^{n\times n}$ is the current approximation to the Jacobian matrix (the matrix of first derivatives) and we move $x \to x^+$. The Broyden update to B is $B^+ = B + E$, where E solves

$$\min_E \{\|E\|_F : Es = y\} \tag{1}$$

and $s = x^+ - x$, $y = F(x^+) - F(x)$. (Note: The Frobenius norm of any matrix M is denoted $\|M\|_F = \sqrt{\sum_{i,j} m_{ij}^2}$.)

   (a) Why is (1) a sensible way to define an update to the Jacobian approximation?

   (b) The solution to (1) is

$$E = \frac{ys^\top}{s^\top s}.$$

   Use an optimization argument to derive this solution to (1).

   (c) Suppose S is a set of index pairs such that if $(i, j) \in S$ then element (i, j) of the Jacobian is a known constant value. Show how to modify the Broyden update to incorporate this information in this case. **Hint:** Problem (1) can be solved in a row-by-row fashion.

6. Assume $f : \mathbb{R}^n \to \mathbb{R}$ is a twice continuously differentiable function. The trust region subproblem, defined at point x, is:

$$\min_s \left\{ q(s) \triangleq s^\top g + \frac{1}{2} s^\top H s : \|s\|_2 \le \Delta \right\}, \tag{2}$$

where $g = \nabla f(x)$, $H = \nabla^2 f(x)$, $\Delta > 0$. The solution to (2), s_*, is a trial step and is accepted, i.e., $x^+ \leftarrow x + s_*$, where s_* solves (2) if and only if $f(x + s_*) < f(x)$. The parameter Δ is adjusted for the next iteration depending on the value of
$\mathit{ratio} = [f(x + s_*) - f(x)] / q(s_*)$.

   (a) True or False: Assuming x does not satisfy 2^{nd}-order necessary conditions to be a local minimizer of f, trial step $s_*(\Delta)$ is accepted for Δ sufficiently small. Explain.

   (b) True or False: If $\nabla f(x) = 0$ then $s_* = 0$. Explain.

   (c) True or False: If matrix H has a negative eigenvalue, then $\|s_*(\Delta)\|_2 = \Delta$. Explain.

   (d) True or False: If H is positive definite and $\Delta > \|H^{-1}g\|_2$ then $\|s_*(\Delta)\|_2 < \Delta$. Explain.

   (e) True or False: If $f(x + s_*) > f(x)$ and $q(s_*) \ne 0$ then $\mathit{ratio} < 0$. Explain.