University of Waterloo CO 367 Midterm 1

Fall 2018

Solutions

Exam details:

Course:	CO 367 – Nonlinear optimization
Instructor:	Laurent Poirrier
Date and time of exam:	Friday October 12, $2018 - 1:30$ pm to $2:20$ pm
Duration of exam:	50 minutes
Location:	PHY 145
Number of exam pages:	3 (includes this cover page)
Exam type:	Closed book – no calculators, no materials allowed

Instructions:

Your answers must be stated and justified in a clear and logical form, and you must show all of your steps in order to receive full marks. You may use any result from class without proof, unless you are being asked to prove this result. You will be graded not only on correctness, but also on clarity of exposition. No collaboration is allowed.

Question 1 [2 marks] Let $A \in \mathbb{R}^{n \times n}$ be a symmetric, positive definite matrix. Prove that all the diagonal entries A_{ii} of A are positive, for i = 1, ..., n.

Solution: Since A is positive definite, we know that $x^T A x > 0$ for all $x \in \mathbb{R}^n$. In particular, for $x = e_i$, $e_i^T A e_i > 0$. Observe that $e_i^T A$ is the first row of A, and $e_i^T A e_i = A_{ii}$.

Question 2 [3 marks] Prove that for any symmetric matrix $A \in \mathbb{R}^{n \times n}$, there exists a finite constant $\beta \in \mathbb{R}$ such that $(A + \beta I)$ is positive semidefinite.

Solution: The eigenvalues of $A + \beta I$ are the values for λ that satisfy $\det(A + \beta I - \lambda I) = 0$. This can be rewritten $\det(A + (\beta - \lambda)I) = 0$. Thus, λ is an eigenvalue of $A + \beta I$ if and only if $\sigma = \lambda - \beta$ is an eigenvalue of A. In other words, σ is an eigenvalue of A if and only if $\lambda = \sigma + \beta$ is an eigenvalue of $A + \beta I$. If the smallest eigenvalue of A is σ_1 , then choosing $\beta = -\sigma_1$ guarantees that all eigenvalues values of $A + \beta I$ satisfy $\lambda_i = \sigma_i + \beta = \sigma_i - \sigma_1 \ge 0$.

Alternative proof: Since A is symmetric, we know that there exist Q orthogonal and $D = \text{diag}(\sigma_1, \ldots, \sigma_n)$ such that $A = QDQ^T$ where $\sigma_1, \ldots, \sigma_n$ are the eigenvalues of A. Consider the matrix $Q(D + \beta I)Q^T$ which has eigenvalues $\sigma_1 + \beta, \ldots, \sigma_n + \beta$. We can show that $Q(D + \beta I)Q^T = QDQ^T + Q\beta IQ^T = A + \beta I$. Therefore, for any $\beta \ge -\sigma_1$, the eigenvalues of $A + \beta I$ are nonnegative.

Question 3 [5 marks] Consider $f : \mathbb{R}^n \to \mathbb{R}$ where $f \in C^1(D)$ and $D \subseteq \mathbb{R}^n$ is a convex set. Note that f is not

necessarily C^2 -smooth. Prove that if

$$(\nabla f(x) - \nabla f(y))^T (x - y) \ge 0, \quad \forall x, y \in D,$$

then f is convex over D.

Solution: We have seen in class that f is convex over D if and only if

$$f(b) \ge f(a) + (b-a)^T \nabla f(a), \text{ for all } a, b \in D,$$

that is, iff

$$f(b) - f(a) - (b - a)^T \nabla f(a) \ge 0$$
 for all $a, b \in D$.

We now need to prove the above inequality. We use the first-order Taylor expansion of f(b) around a:

$$f(b) = f(a) + (b - a)^T \nabla f(a + \lambda(b - a)),$$

for some $0 \leq \lambda \leq 1$. The gradient exists since $f \in C^1(D)$, $a, b \in D$ and D is convex. We get

$$\begin{aligned} f(b) - f(a) - (b-a)^T \nabla f(a) &= f(a) + (b-a)^T \nabla f(a+\lambda(b-a)) - f(a) - (b-a)^T \nabla f(a) \\ &= (b-a)^T \left(\nabla f(a+\lambda(b-a)) - \nabla f(a) \right) \end{aligned}$$

If $\lambda = 0$, the above expression is 0 which satisfies the claim. Otherwise, we get

$$f(b) - f(a) - (b - a)^T \nabla f(a) = \frac{1}{\lambda} \lambda (b - a)^T (\nabla f(a + \lambda (b - a)) - \nabla f(a))$$

$$= \frac{1}{\lambda} (y - x)^T (\nabla f(y) - \nabla f(x))$$

$$\ge 0,$$

where we defined x := a and $y := a + \lambda(b - a)$, then used the hypothesis.

Alternative solution: Let $c = (1 - \alpha)a + \alpha b$, for some $a, b \in D$, thus $c \in D$. Equivalently, we have $a = c - \alpha(b - a)$ and $b = c + (1 - \alpha)(b - a)$. We need to prove that convexity holds, i.e.

$$f(z) \ge (1 - \alpha)f(a) + \alpha f(b),$$

for all such a, b, c in D. Let us write the Taylor expansion of f(a) around c:

$$f(a) = f(c) - \alpha(b-a)^T \nabla f(c - \lambda \alpha(b-a)), \quad \text{for some } 0 \le \lambda \le 1,$$

and the Taylor expansion of f(b) around c:

$$f(b) = f(c) + (1 - \alpha)(b - a)^T \nabla f(c + \sigma(1 - \alpha)(b - a)), \quad \text{for some } 0 \le \sigma \le 1.$$

Taking a linear combination of these two equations with multipliers $(1 - \alpha)$ and α , we get

$$\begin{aligned} (1-\alpha)f(a) + \alpha f(b) &= (1-\alpha)f(c) + \alpha f(c) \\ &- (1-\alpha)\alpha(b-a)^T \nabla f(c-\lambda\alpha(b-a)) \\ &+ \alpha(1-\alpha)(b-a)^T \nabla f(c+\sigma(1-\alpha)(b-a)) \end{aligned}$$
$$= f(c) + \alpha(1-\alpha)(b-a)^T \left(\nabla f(c+\sigma(1-\alpha)(b-a)) - \nabla f(c-\lambda\alpha(b-a))\right) \\ &= f(c) + \alpha(1-\alpha)(b-a)^T \left(\nabla f(y+(\alpha\lambda+(1-\alpha)\sigma)(b-a)) - \nabla f(y)\right),\end{aligned}$$

where we let $y := c - \lambda \alpha (b - a)$. If both σ and λ are zero, then the above expression is zero and the claim is true. Otherwise, $(\alpha \lambda + (1 - \alpha)\sigma) > 0$ and we can write

$$(1-\alpha)f(a) + \alpha f(b) = f(c) + \frac{\alpha(1-\alpha)}{\alpha\lambda + (1-\alpha)\sigma}(\alpha\lambda + (1-\alpha)\sigma)(b-a)^T \left(\nabla f(y + (\alpha\lambda + (1-\alpha)\sigma)(b-a)) - \nabla f(y)\right).$$

We can now let $x := y + (\alpha \lambda + (1 - \alpha)\sigma)(b - a)$ and obtain

$$(1-\alpha)f(a) + \alpha f(b) = f(c) + \frac{\alpha(1-\alpha)}{\alpha\lambda + (1-\alpha)\sigma}(x-y)^T \left(\nabla f(x) - \nabla f(y)\right)$$

$$\geq f(c).$$