1. (10 points) Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite. The set $\varepsilon = \{ x \in \mathbb{R}^n : x^T A x \leq 1 \}$ is called an ellipsoid. Show that $\varepsilon$ is convex.

Proof. $A$ is symmetric positive definite. By Cholesky decomposition, there exists matrix $R$, such that $A = R^T R$. By definition, to show set $\varepsilon$ is convex, we need to prove, for any $x, y \in \varepsilon$, and $\forall \alpha \in [0, 1]$, we have

$$\alpha x + (1 - \alpha) y \in \varepsilon.$$ 

As $x, y \in \varepsilon$, then $x^T A x = \| Rx \|_2^2 \leq 1$, and $y^T A y = \| Ry \|_2^2 \leq 1$. Let 

$$z = \alpha x + (1 - \alpha) y$$ 

$$z^T A z = \| Rz \|_2^2 = \| \alpha Rx + (1 - \alpha) Ry \|_2^2 \leq (\alpha \| Rx \|_2 + (1 - \alpha) \| Ry \|_2)^2$$ 

So $z^T A z \leq (\alpha + 1 - \alpha)^2 = 1$ and $z \in \varepsilon$, which proves set $\varepsilon$ is convex.

Another method:
We can check directly by estimating on $x^T A y$. Let 

$$z = \alpha x + (1 - \alpha) y$$ 

So $z^T A z = \alpha^2 x^T A x + (1 - \alpha)^2 y^T A y + 2\alpha(1 - \alpha) x^T A y$.

Because $A$ is positive definite, we have that 

$$(x - y)^T A (x - y) = x^T A x + y^T A y - 2x^T A y \geq 0$$

the $=$ holds iff $x = y$. Therefore $2x^T A y \leq x^T A x + y^T A y \leq 2$. So 

$$z^T A z = \alpha^2 x^T A x + (1 - \alpha)^2 y^T A y + 2\alpha(1 - \alpha) x^T A y \leq \alpha^2 + (1 - \alpha)^2 + 2\alpha(1 - \alpha) = 1$$

\[ \square \]

2. (10 points) Let $\{ x_k \} \subset \mathbb{R}$ be a sequence such that $x_k \to x^* > 0$, with convergence order $r \geq 1$. Find the convergence order of the sequence $\{ x_k^2 \}$. 

---

Math 171B Homework Assignment 2 Solutions

Instructor: Jiawang Nie
Proof. By definition of convergence order, sequence $x_k \in \mathbb{R}$,

$$\lim_{k \to \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|^r} = a > 0.$$ 

$f(x) = x^2$ is a continuous function, so as $x_k \to x^*$, we have $x_k^2 \to (x^*)^2$. Next we consider the convergence order of this sequence. As $x^* > 0$ and $r \geq 1$.

$$\lim_{k \to \infty} \frac{|x_{k+1}^2 - (x^*)^2|}{|x_k^2 - (x^*)^2|^r} = \lim_{k \to \infty} \frac{|x_{k+1}^2 + x^*| \cdot |x_{k+1}^2 - x^*|}{|x_k^2 + x^*|^r \cdot |x_k^2 - x^*|^r}.$$ 

So the order of convergence is $r$. \qed

3. (5 points) Suppose $f(x) \in C^1[a, b]$, $f'(x) \neq 0$ for all $x \in (a, b)$, and $c \in [a, b]$ is a minimizer of $f(x)$ on $[a, b]$. Show that $c = a$ or $c = b$.

Proof. By extremal value theorem, we know this $c$ exists.

Suppose $c \in (a, b)$, then $f(a) > f(c)$ and $f(b) > f(c)$.

By mean value theorem, there exists $\xi_1 \in (a, c)$, such that

$$f(a) - f(c) = f'(\xi_1)(a - c) > 0,$$

then $f'(\xi_1) < 0$

Similarly, there exists $\xi_2 \in (c, b)$, such that

$$f(b) - f(c) = f'(\xi_2)(b - c) > 0,$$

then $f'(\xi_2) > 0$.

$f(x) \in C^1[a, b]$, so $f'(x)$ is continuous over $(a, b)$,

$0 \in [f'(\xi_1), f'(\xi_2)]$, so by intermediate value theorem, there exists $\xi \in [\xi_1, \xi_2] \subset [a, b]$, such that $f'(\xi) = 0$, which contradict the assumption that $f'(x) \neq 0$ over $x \in (a, b)$. So $c = a$ or $c = b$. \qed

4. (5 points) Suppose $f(x) \in C[a, b]$. Show that there exists a number $\xi \in [a, b]$ such that

$$f(\xi) = \frac{4f(a) + 3f(b)}{7}.$$ 

Solution: Let $\alpha = \frac{4}{7} \in [0, 1]$, then

$$c = \alpha f(a) + (1 - \alpha) f(b) \in [\min\{f(a), f(b)\}, \max\{f(a), f(b)\}]$$

By Intermediate value theorem, there exists $\xi \in [a, b]$, such that $f(\xi) = c$. 

2
5. (10 points) Find the gradient and hessian of the function

\[ f(x) = \frac{x_1^2 x_2^2}{x_1^4 + x_2^4}, \] where \( x = (x_1, x_2). \)

Solution: Gradient

\[
\nabla f(x) = \begin{bmatrix}
\frac{\partial f(x)}{\partial x_1} \\
\frac{\partial f(x)}{\partial x_2}
\end{bmatrix} = \begin{bmatrix}
\frac{2x_1 x_2^2 (x_1^2 - x_2^2)}{(x_1^4 + x_2^4)^2} \\
\frac{2x_1^2 x_2 (x_1^2 - x_2^2)}{(x_1^4 + x_2^4)^2}
\end{bmatrix}
\]

Hessian

\[
\nabla^2 f(x) = \begin{bmatrix}
\frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} \\
\frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2}
\end{bmatrix} = \begin{bmatrix}
\frac{2x_1^2 (3x_1^4 - 12x_1^2 x_2^2 + x_2^4)}{(x_1^4 + x_2^4)^3} & \frac{-4x_1 x_2 (x_1^4 + x_2^4)}{(x_1^4 + x_2^4)^3} \\
\frac{-4x_1 x_2 (x_1^4 + x_2^4)}{(x_1^4 + x_2^4)^3} & \frac{2x_1^2 (3x_1^4 - 12x_1^2 x_2^2 + x_2^4)}{(x_1^4 + x_2^4)^3}
\end{bmatrix}
\]

6. (10 points) Let \( a \in \mathbb{R}^n, \ H \in \mathbb{R}^{n \times n} \) be a symmetric matrix, \( A \in \mathbb{R}^{n \times n} \), and \( f(x), g(x) \) be the functions on \( \mathbb{R}^n \) defined as

\[ f(x) = a^T x + \frac{1}{2} x^T H x, \quad g(x) = Ax. \]

Find the gradient and Hessian of the composition \( h(x) = f(g(x)) \).

Solution: \( h(x) = f(g(x)) = a^T A x + \frac{1}{2} x^T A^T H A x \), which is still a quadratic function. \( H \) is symmetric.

Gradient \( \nabla h(x) = A^T H A x + A^T a \),

Hessian \( \nabla^2 h(x) = A^T H A \).

**Method 2:** Use chain rule.

Let \( g : \mathbb{R}^k \rightarrow \mathbb{R}^m \) and \( f : \mathbb{R}^m \rightarrow \mathbb{R}^s \), and \( h \) is their composition, i.e.:

\[ h(x) = f(g(x)) : \mathbb{R}^k \rightarrow \mathbb{R}^s, \]

then \( \nabla h(x) = \nabla g(x) \nabla f(g(x)) \).

In this question, \( k = m = n, \ s = 1, \) \( a \). So use the Chain rule:

\[
\nabla h(x) = \nabla g(x) \nabla f(g(x)).
\]

Review: For \( f : \mathbb{R}^n \rightarrow \mathbb{R}^m, \ f(x) = [f_1(x), \cdots, f_m(x)]^T \), then

\[
\nabla f(x) = [\nabla f_1(x), \cdots, \nabla f_m(x)].
\]

By simple calculation, \( \nabla g(x) = A^T \), and \( \nabla f(g(x)) = a + H A x \), so we have:

\[
\nabla f(g(x)) = A^T (a + H A x) = A^T a + A^T H A x,
\]

which is the same as above. And Hessian is:

\[
\nabla^2 h(x) = \nabla (\nabla f(g(x))) = \nabla (A^T a + A^T H A x) = A^T H A.
\]