ME5114 Sp23 **Homework Solution # 1**

Assigned Feb 16, 2023 Due: Feb 23 , 2023

---

1. **Python**: Please go over the Python tutorial posted on the course website. Make sure you can write basic python codes.

2. **Lipschitz Continuity**

   (a) Please state the formal definition of continuous functions

   (b) Please state the formal definitions of Lipschitz continuity and locally Lipschitz continuity.

   Solution: (a): Suppose X and Y are metric spaces, $E \in X$, $p \in E$, and f maps E into Y. Then f is said to be continuous at p if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$d_Y(f(x), f(p)) < \varepsilon \tag{1}$$

   for all points $x \in E$ for which $d_X(x, p) < \delta$. d_X denotes the metric on X.
   If f is continuous at every point of E, then f is said to be continuous on E.

   (b): Suppose X and Y are metric spaces, $E \in X$, $p \in E$, and f maps E into Y. Then f is said to be Lipschitz continuous if there exists a real constant $L \geq 0$ such that

$$d_Y(f(x_1), f(x_2)) \leq L d_X(x_1, x_2) \tag{2}$$

   for all x_1 and x_2 in E.
   f is called locally Lipschitz continuous if for every x in E there exists a neighborhood U of x such that f restricted to U is Lipschitz continuous. □

3. **Matrix calculus**

   (a) Let $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ be a scalar function of matrix variable. Please write a tutorial paragraph explaining (in your own words) the meaning of $\frac{\partial}{\partial X} f(X)$.

   (b) Let $A \in \mathbb{R}^{n\times m}, X \in \mathbb{R}^{m\times n}$. Derive an expression for $\frac{\partial}{\partial X} tr(AX)$ (show your derivation steps; your derivation should be directly from the definition of matrix derivatives)

   (c) Derive an expression for $\frac{\partial}{\partial x} f(x)$, where $f(x) = x^T Q x + tr(xx^T)$ and $x \in \mathbb{R}^n$
   Solution:
   (a): Here we use the "broad definition" in [1] for $\frac{\partial}{\partial X} f(X)$.

$$\frac{\partial}{\partial X} f(X) = \begin{bmatrix} \frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1m}} \\ \frac{\partial f}{\partial x_{21}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{2m}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{n1}} & \frac{\partial f}{\partial x_{n2}} & \cdots & \frac{\partial f}{\partial x_{nm}} \end{bmatrix} \tag{3}$$

   or as

$$\frac{\partial}{\partial X} f(X) = \begin{bmatrix} \frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{21}} & \cdots & \frac{\partial f}{\partial x_{n1}} \\ \frac{\partial f}{\partial x_{12}} & \frac{\partial f}{\partial x_{22}} & \cdots & \frac{\partial f}{\partial x_{n2}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{1m}} & \frac{\partial f}{\partial x_{2m}} & \cdots & \frac{\partial f}{\partial x_{nm}} \end{bmatrix} \tag{4}$$

For (3), every column of $\frac{\partial}{\partial X} f(X)$ can be regarded as the gradient of $f(X)$ in the corresponding column of X.

(b):

$$tr(AX) = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} x_{ji} \tag{5}$$

combine (5) and (3) (or (4)), it is obvious that

$$\frac{\partial}{\partial X} tr(AX) = A^T \tag{6}$$

(c):

$$x^T Q x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ q_{21} & q_{22} & \cdots & q_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_3 \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i q_{ij} x_j \tag{7}$$

$$tr(xx^T) = \sum_{i=1}^{n} x_i^2 \tag{8}$$

So

$$\frac{\partial}{\partial x} f(x) = \begin{bmatrix} \sum_{i=1}^{n} q_{1i} x_i + \sum_{i=1}^{n} q_{i1} x_i + 2x_1 \\ \sum_{i=1}^{n} q_{2i} x_i + \sum_{i=1}^{n} q_{i2} x_i + 2x_2 \\ \vdots \\ \sum_{i=1}^{n} q_{ni} x_i + \sum_{i=1}^{n} q_{in} x_i + 2x_n \end{bmatrix} = Qx + Q^T x + 2x \tag{9}$$

□

4. **Inner product**

   (a) Describe the way to calculate the angle between two vectors $x, y \in \mathbb{R}^n$ using inner product

   (b) Trace can be used to define inner products for matrices. Let $A, B \in \mathbb{R}^{m \times n}$, then $\langle A, B \rangle = tr(A^T B)$. Compute the angle between the following two matrices

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} -1 & 2 & 1 \\ -1 & 0 & 1 \end{bmatrix}$$

Solution:

(a):

$$\alpha = \arccos\left(\frac{\langle x, y \rangle}{||x||||y||}\right) \tag{10}$$

(b):

$$\alpha = \arccos\left(\frac{\langle A, B\rangle}{\sqrt{\langle A, A\rangle}\sqrt{\langle B, B\rangle}}\right) = \arccos\left(\frac{tr(A^T B)}{tr(A^T A)tr(B^T B)}\right) = \frac{\pi}{2} \tag{11}$$

$\square$

5. **Some linear algebra**

   (a) State the condition on A such that $Ax = b$ has at least one solution.

   (b) Let $A = [a_1, a_2, a_3, a_4]$, where $a_i \in \mathbb{R}^n$ are columns of A. Suppose a_1, a_2 are linearly independent, and $a_3 + a_1 = a_2$ and $a_4 - a_3 = a_1$. Compute $rank(A)$ and $Null(A)$.

   (c) Given a vector $y \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{n\times m}$, find an expression of the projection of y onto the column space of A.

Solution:

(a): When $rank(A) = rank(A|b)$, $Ax = b$ has at least one solution. OR $Ax = b$ has at least one solution if and only if b is in the space spanned by column vectors of A.

(b): From $a_3 + a_1 = a_2$ and $a_4 - a_3 = a_1$, we can know that $a_2 = a_4$, and they are linear combination of a_1 and a_3. $rank(A) = 2$

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} = 0, \quad \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix} = 0$$

So

$$Null(A) = span\left\{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}\right\}$$

(c) Given a vector $y \in \mathbb{R}^n$ and a matrix $A \in \mathbb{R}^{n\times m}$, find an expression of the projection of y onto the column space of ALet's denote the projection of y onto the column space of A as $p = Ax$, then the difference between the original vector y and the projection p, which is called the error e can be expressed as

$$e = y - p = y - Ax$$

Since the column space of A and the left nullspace of A are orthogonal, we have

$$A^T(y - Ax) = A^T y - A^T Ax = 0$$

Then we can solve for x using the equation above:

$$x = (A^T A)^{-1} A^T y$$

Finally we can compute the projection p:

$$p = Ax = A(A^T A)^{-1} A^T y$$

□

6. **Ellipsoids:** Ellipsoid in $\mathbb{R}^n$ have two equivalent representations: (i) $E_1(P, x_c) = \{x \in \mathbb{R}^n : (x - x_c)^T P^{-1}(x - x_c) \leq 1\}$ and (ii) $E_2(A, x_c) = \{Au + x_c : \|u\|^2 \leq 1\}$. Given an eillipsoid $E_1(P, x_c)$ with P positive definite, its volume is $\nu_n \sqrt{det(P)}$ where ν_n is the volume of unit ball in $\mathbb{R}^n$, its semi-axes directions are given by the eigenvectors of P and the lengths of semi-axes are $\sqrt{\lambda_i}$, where λ_i are eigenvalues of P.

   (a) Given an Ellipsoid $E_1(P, x_c)$, find the corresponding (A, b) (in terms of P and x_c) such that $E_2(A, b)$ represents the same ellipsoid as $E_1(P, x_c)$

   (b) Draw the ellipse $E_1(P, x_c)$ with $P = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ and $x_c = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ by hand.

   Solution:
   (a): For representation (i), the positive definite P can be decomposed as $P = LL^T$ where L is invertible. So

$$(x - x_c)^T P^{-1}(x - x_c) = (x - x_c)^T {L^{-1}}^T L^{-1}(x - x_c) \leq 1 \tag{12}$$

   which leads

$$||L^{-1}(x - x_c)|| \leq 1 \tag{13}$$

   If we let $L^{-1}(x - x_c) = u$, it is easily to know

$$Lu + x_c = x \tag{14}$$

   where $||u|| \leq 1$ (because of (13)).
   Now we can find that (14) has the same formulation with the representation (ii). So give a ellipsoid $E_1(P, x_c)$, the another representation is $E_2(L, x_c)$ where $P = LL^T$.

   (b): See Figure 1

□

7. **Linear System Solution**: Consider the following linear control system:

$$\dot{x}(t) = Ax(t) + Bu(t), \quad \text{with } x(0) = x_0$$

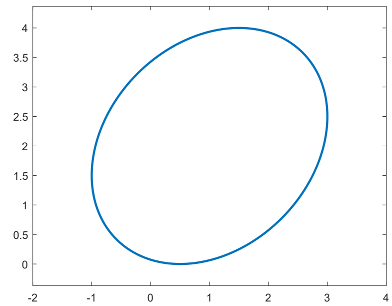

Figure 1: The ellipse $E_1(P, x_c)$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the control input. Show

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

is a solution to the above control system.

Solution:

We have

$$\begin{aligned} \frac{d}{dt}e^{At} &= \frac{d}{dt}(I + At + \frac{A^2}{2!}t^2 + \cdots) \\ &= A + A^2 t + \frac{A^3}{2!}t^2 + \cdots \\ &= A(I + At + \frac{A^2}{2!}t^2 + \cdots) \\ &= Ae^{At} \end{aligned} \tag{15}$$

$$\frac{\partial}{\partial t}\int_{t_0}^t f(t,\tau)d\tau = \int_{t_0}^t \left(\frac{\partial}{\partial t}f(t,\tau)\right)d\tau + f(t,\tau)|_{\tau=t} \tag{16}$$

Now we show

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$$

is a solution to the above control system:

$$\begin{aligned} \frac{d}{dt}x(t) &= Ae^{At}x_0 + \int_0^t Ae^{A(t-\tau)}Bu(\tau)d\tau + e^{A(t-\tau)}Bu(\tau)|_{\tau=t} \\ &= A\left(e^{At}x_0 + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau\right) + Bu(\tau) \\ &= Ax(t) + Bu(\tau) \end{aligned} \tag{17}$$

And $x(0) = e^{\mathbf{0}} x_0 = x_0$ which is consistent with the initial condition.

$\square$

References

[1] Jan R. Magnus. On the concept of matrix derivative. *Journal of Multivariate Analysis*, 101(9):2200–2206, 2010.