# Assignment 3

# MATH 7502 - Semsester 2, 2018

## Mathematics for Data Science 2

**Created by Zhihao Qiao, Maria Kleshnina and Yoni Nazarathy**

# Question 1

Consider the following recurision,

$$x_{t+1} = A_1 x_t + A_2 x_{t-1}, \quad , t = 2, 3, \ldots$$

where $x_t$ is n-vector and $A_1$ and $A_2$ are $n \times n$ matrices. Define $z_t = (x_t, x_{t-1})$.

Show that $z_t$ satisfies the linear dynamical system equation $z_{t+1} = Bz_t$, for $t = 2, 3, \ldots$, where $B$ is a $(2n) \times (2n)$ matrix.

# Solution

$z_t$ is a stack vector with size (2n), which is

$$z_t = \begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix}$$

Now Let

$$B = \begin{bmatrix} A_1 & 0_n \\ 0_n & A_2 \end{bmatrix}$$

$0_n$ denotes that an $n \times n$ matrix with all zeros.

Therefore, we have

$$Bz_t = \begin{bmatrix} A_1 & 0_n \\ 0_n & A_2 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix} = A_1 x_t + A_2 x_{t-1}$$

# Question 2

Consider the Fibonacci sequence $y_0, y_1, y_2, \ldots$ with $y_0 = 0, y_1 = 1, y_2 = 1, y_3 = 2, \ldots$, and for $t = 2, 3, \ldots, y_t$ is the sum of the previous two terms $y_{t-1}$ and $y_{t-2}$.

(a) Express the Fibonacci sequence as a time-invaraint dynamical system wit state $x_t = (y_t, y_{t-1})$ and output $y_t$ for $t = 1, 2, 3 \ldots.$ as

$$x_{t+1} = Ax_t$$

(b )For the matrix, A, compute the eigenvalues and describe the Fibonnaci sequence interms of eigevnalues and eigenvectors. How does the golden ratio play a role?

# Solution

(a) we know that $y_{t+1} = y_t + y_{t-1}$

therefore

$$\begin{bmatrix} y_t \\ y_{t-1} \end{bmatrix} = \begin{bmatrix} y_{t-1} + y_{t-2} \\ y_{t-1} \end{bmatrix} = A \begin{bmatrix} y_{t-1} \\ y_{t-2} \end{bmatrix}$$
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

.

(b) The eigenvalues of $A$ are $(A - \lambda I) = 0$ we have the quadratic
$$\lambda^2 - \lambda - 1 = 0$$

Therefore, the eigenvalues are $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$, $\lambda_1$ is the golden ratio.

The eigenvectors of $\lambda_1$ is

$$v_1 = \begin{bmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{bmatrix}$$

and the eigenvectors of $\lambda_2$ is

$$v_1 = \begin{bmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{bmatrix}$$

# Question 3

In the sepcial case $n = 1$, the general least square problem reduces to finding a scalar $x$ that minimizes $||ax - b||^2$, where $a$ and $b$ are m-vectors. Assuming $a$ and $b$ are nonzero, show that $||a\hat{x} - b||^2 = ||b||^2 \sin^2(\theta)$, where $\theta = \angle(a, b)$.

# Solution

correspond least square problem of this one is

$$f(x) = (a_1x - b_1)^2 + (a_2x - b_2)^2 + \ldots + (a_mx - b_m)^2$$

set the derivative of $f(x)$ to be zero , which is

$$f'(x) = 2a_1(a_1x - b_1) + 2a_2(a_2x - b_2) + \ldots + 2a_m(a_mx - b_m) = 0$$
$$2[(a_1)^2 + (a_2)^2 + \ldots + (a_m)^2]x - 2(a_1b_1 + a_2b_2 + \ldots a_mb_m) = 0$$
$$2x(a^Ta) - 2a^Tb = 0$$
$$\hat{x} = \frac{a^Tb}{a^Ta}$$

$$\begin{aligned} \|a(\frac{a^Tb}{a^Ta}) - b\|^2 &= \|\frac{a^Tb}{a^Ta}\|^2\|a\|^2 + \|b\|^2 - 2\|\frac{a^Tb}{a^Ta}\|\|a^Tb\| \\ &= \frac{\|a^Tb\|^2}{\|a\|^2} + \|b\|^2 - 2\frac{\|a^Tb\|^2}{\|a\|^2} \\ &= \|b\|^2(1 - \frac{\|a^Tb\|^2}{\|a\|^2\|b\|^2}) \\ &= \|b\|^2(1 - \cos^2(\theta)) \\ &= \|b\|^2\sin^2(\theta). \end{aligned}$$

# Question 4

Consider a time-invaraint linear dynamical system with offset

$$x_{t+1} = Ax_t + c$$

where $x_t$ is the state n-vector. We say that a vector $z$ is an equilibrium point of the linear dynamical system if $x_1 = z$ implies $x_2 = z, x_3 = z, \ldots$

(a)Find a matrix $F$ and a vector $g$ for which the set of linear equations $Fz = g$ characterizes equilibrium points. (This means: If $z$ is an equilibrium point, then $Fz = g$; conversely if $Fz = g$, then $z$ is an equilibrium point.)

Express $F$ and $g$ interms of $A, c$ any standard matrices or vectors and matrix and vector operations.

# Solution

if $z$ is an equilibrium point, we have

$$\begin{aligned} z &= Az + c \\ A(z) &= A(Az) + A(c) \\ A(z) - A(Az) &= A(c) \\ A(I - A)z &= A(c) \end{aligned}$$

Therefore $F = A(I - A)$ and $g = A(c)$.

Now if

$$\begin{aligned} Fz &= g \\ A(I - A)z &= A(c) \\ (I - A)z &= c \\ z - Az &= c \\ z &= Az + c \end{aligned}$$

# Question 5

Suppose that $m \times n$ matrix $Q$ has orthonormal columns and $b$ is an m-vector. Show that $\hat{x} = Q^T b$ is the vecotr that minimizes $||Qx - b||^2$.

Comment on the complexity of find $\hat{x}$ given $Q$ and $b$ in this case. Compare the complexity with it is a general least squares problem and $Q$ is a coefficient matrix.

# Solution

if $\hat{x} = Q^T b$, and if $Q$ has orthonomarl columns, then

$$\begin{aligned} QQ^T &= I \\ ||Q\hat{x} - b||^2 &= ||QQ^T b - b||^2 \\ &= ||b - b||^2 \\ &= 0 \end{aligned}$$

# Question 6

Suppose $m \times n$ matrix $A$ has linearly independent columns, and $b$ is a m-vector. let $\hat{x} = A^\dagger b$ denote the least squares approximate solution of $Ax = b$.

(a) Show that for any n-vector $x$, $(Ax)^T b = (Ax)^T (A\hat{x})$. Hint: Use $(Ax)^T b = x^T (A^T b)$, and $(A^T A)\hat{x} = A^T b$.

(b) Show that when $A\hat{x}$ and $b$ are both nonzero, we have

$$\frac{(A\hat{x})^T b}{||A\hat{x}||||b||} = \frac{||A\hat{x}||}{||b||}$$

(c) The choice of $x = \hat{x}$ minimizes the distance between $Ax$ and $b$, Show that $x = \hat{x}$ also minimizes the angle between $Ax$ and $b$.

# Solution

(a) we have

$$\begin{aligned}(Ax)^T b &= x^T(A^T b)\\ &= x^T(A^T A)\hat{x}\\ &= (Ax)^T(A\hat{x})\end{aligned}$$

(b) using (a) , we have

$$(A\hat{x})^T b = (A\hat{x})^T(A\hat{x})$$

Therefore

$$\begin{aligned}\frac{(A\hat{x})^T b}{||A\hat{x}||||b||} &= \frac{(A\hat{x})^T(A\hat{x})}{||A\hat{x}||||b||}\\ &= \frac{||A\hat{x}||^2}{||A\hat{x}||||b||}\\ &= \frac{||A\hat{x}||}{||b||}\end{aligned}$$

(c) By definition, the angle between $Ax$ and $b$ is defined as

$$\theta = \cos^{-1}\left(\frac{(Ax)^T b}{||Ax||||b||}\right) = \cos^{-1}\left(\frac{||Ax||}{||b||}\right)$$

from (b), since $x = \hat{x}$ is the least square approximation, therefore $||A\hat{x}||$ is minimized for all $x$, therefore, $\theta$ is also minimized in this expression.

# Question 7

Suppose $A$ is an $m \times n$ matrix with linearly independent columns and $QR$ factorization $A = QR$, and $b$ is the m-vector. The vector $A\hat{x}$ is the linear combination of the columns of $A$ that is closet to the vector $b$, i.e., it is the projection of $b$ onto the set of linear combinations of the columns of $A$.

(a) Show that $A\hat{x} = QQ^T b$.

(b) Show that $||A\hat{x} - b||^2 = ||b||^2 - ||Q^T b||^2$.

# Solution

(a) If $\hat{x}$ is a solution of the least quare problem, then

$$\begin{aligned}\hat{x} &= (A^TA)^{-1}A^Tb \\ \hat{x} &= A^{-1}(A^T)^{-1)}A^Tb \\ A\hat{x} &= ((QR)^T)^{-1}(QR)^Tb \\ &= (R^TQ^T)^{-1}(R^TQ^T)b \\ &= (Q^T)^{-1}(R^T)^{-1}R^TQ^Tb \\ &= (Q^{-1})^{-1}Q^Tb \\ &= QQ^Tb\end{aligned}$$

(b)

First consider the norm

$$\|QQ^Tb\|^2 = (QQ^Tb, QQ^Tb) = (Q^TQQ^Tb, Q^Tb) = (Q^Tb, Q^Tb) = \|Q^Tb\|^2$$

and the norm

$$\begin{aligned}(QQ^Tb, b) &= (Q^TQQ^Tb, Q^Tb) = (Q^Tb, Q^Tb) = \|Q^Tb\|^2 \\ \|A\hat{x} - b\|^2 &= (QQ^Tb, QQ^Tb) - 2(Q^TQQ^Tb, Q^Tb) + \|b\|^2 \\ &= \|Q^Tb\|^2 - 2\|Q^Tb\|^2 + \|b\|^2 \\ &= \|b\|^2 - \|Q^Tb\|^2\end{aligned}$$

# Question 8

A generalization of the least squares problem adds an affine function to the least sqres objective

$$\text{minimize} \qquad \|Ax - b\|^2 + c^Tx + d$$

where $x$ is an n-vector as a variable to be chosen, and the data are the $m \times n$ matrix $A$, the m-vector $b$, the n-vector $c$. and the number $d$. The columns of $A$ are linearly indeppendent.

Show that that objective of the problem above can be expressed in the form

$$\|Ax - b\|^2 + c^Tx + d = \|Ax - b + f\|^2 + g$$

,

for some m-vecotr $f$ and some constant $g$. It follows that we can solve the generalized least squares problem by minimizing $\|Ax - (b - f)\|$, an ordinary least squares problem with solution $\hat{x} = A^{\dagger}(b - f)$.

Hint: Express the norm squared term on the right-hand side as $\|(Ax - b) + f\|^2$ and expand it.

# Solution

If we expand the term

$$\begin{aligned}\|(Ax - b) + f\|^2 &= \|Ax - b\|^2 + 2f^T(Ax - b) + \|f\|^2 \\ &= \|Ax - b\|^2 + 2f^TAx - 2f^Tb + \|f\|^2 \\ &= \|Ax - b\|^2 + 2(A^Tf)^Tx + (\|f\|^2 - 2f^Tb)\end{aligned}$$

Therefore, $c = A^Tf$ , and $d = \|f\|^2 - 2f^Tb$