Homework 4 Yu Gan

1. (a) Denote A as: $\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p1} & A_{p2} & \cdots & A_{pm} \end{bmatrix}$ X as: $\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix}$ B as: $\begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1n} \\ B_{21} & B_{22} & \cdots & B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B_{q1} & B_{q2} & \cdots & B_{qn} \end{bmatrix}$

$$B^T = \begin{bmatrix} B_{11} & B_{21} & \cdots & B_{q1} \\ B_{12} & B_{22} & \cdots & B_{q2} \\ \vdots & \vdots & \ddots & \vdots \\ B_{1n} & B_{2n} & \cdots & B_{qn} \end{bmatrix} = \begin{bmatrix} B_{1:}, & B_{2:}, & \cdots & B_{q:} \end{bmatrix}$$

Therefore $AX = \begin{bmatrix} A_{1,:}X_{:,1} & A_{1,:}X_{:,2} & \cdots & A_{1,:}X_{:,n} \\ A_{2,:}X_{:,1} & A_{2,:}X_{:,2} & \cdots & A_{2,:}X_{:,n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{p,:}X_{:,1} & A_{p,:}X_{:,2} & \cdots & A_{p,:}X_{:,n} \end{bmatrix} = \begin{bmatrix} (Ax)_{1,:} \\ (Ax)_{2,:} \\ \vdots \\ (Ax)_{p,:} \end{bmatrix}$

$$AxB^T = \begin{bmatrix} (Ax)_{1,:}B_{1,:} & (Ax)_{1,:}B_{2,:} & \cdots & (Ax)_{1,:}B_{q,:} \\ (Ax)_{2,:}B_{1,:} & (Ax)_{2,:}B_{2,:} & \cdots & (Ax)_{2,:}B_{q,:} \\ \vdots & \vdots & \ddots & \vdots \\ (Ax)_{p,:}B_{1,:} & (Ax)_{p,:}B_{2,:} & \cdots & (Ax)_{p,:}B_{q,:} \end{bmatrix}$$

Therefore $vec(AxB^T) = \begin{bmatrix} (Ax)_{1,:}B_{1,:} \\ (Ax)_{2,:}B_{1,:} \\ \vdots \\ (Ax)_{p,:}B_{1,:} \\ (Ax)_{1,:}B_{2,:} \\ \vdots \\ (Ax)_{p,:}B_{q,:} \end{bmatrix}$ a $qp \times 1$ matrix

Also $B \otimes A = \begin{bmatrix} B_{1,:}A \\ B_{2,:}A \\ \vdots \\ B_{q,:}A \end{bmatrix} = \begin{bmatrix} B_{11}A & \cdots & B_{1n}A \\ B_{21}A & \cdots & B_{2n}A \\ \vdots & \ddots & \vdots \\ B_{q1}A & \cdots & B_{qn}A \end{bmatrix} = \begin{bmatrix} B_{11}A_{11} & \cdots & B_{11}A_{1m} & \cdots & B_{1n}A_{11} & \cdots & B_{1n}A_{1m} \\ B_{11}A_{21} & \cdots & B_{11}A_{2m} & \cdots & B_{1n}A_{21} & \cdots & B_{1n}A_{2m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ B_{11}A_{p1} & \cdots & B_{11}A_{pm} & \cdots & B_{1n}A_{p1} & \cdots & B_{1n}A_{pm} \\ \vdots & & \vdots & & \vdots & & \vdots \\ B_{q1}A_{11} & \cdots & B_{q1}A_{1m} & \cdots & B_{qn}A_{11} & \cdots & B_{qn}A_{1m} \\ B_{q1}A_{21} & \cdots & B_{q1}A_{2m} & \cdots & B_{qn}A_{21} & \cdots & B_{qn}A_{2m} \\ B_{q1}A_{p1} & \cdots & B_{q1}A_{pm} & \cdots & B_{qn}A_{p1} & \cdots & B_{qn}A_{pm} \end{bmatrix}$

a $qp \times mn$ matrix

$Vec(x) = \begin{bmatrix} x_{:,1} \\ x_{:,2} \\ \vdots \\ x_{:,n} \end{bmatrix}$ (a $mn \times 1$ matrix) $= \begin{bmatrix} x_{11} \\ x_{21} \\ x_{31} \\ \vdots \\ x_{m1} \\ \vdots \\ x_{mn} \end{bmatrix}$

$B \otimes A \cdot Vec(x) = \begin{bmatrix} B_{11}A_{11}x_{11} + B_{11}A_{12}x_{21} + \cdots + B_{1n}A_{1m}x_{mn} \\ B_{11}A_{21}x_{11} + B_{11}A_{22}x_{21} + \cdots + B_{1n}A_{2n}x_{mn} \\ \\ \vdots \\ \\ B_{q1}A_{p1}x_{11} + B_{q1}A_{p2}x_{21} + \cdots + B_{1n}A_{2n}x_{mn} \end{bmatrix}$ a $qp \times 1$ matrix

Then Look back at the representation of $vec(AxB^T)$

$$(Ax)_{1,:} B_{1,:} = A_{1,:} x_{:,1} B_{11} + A_{1,:} X_{:,2} B_{12} + \cdots + A_{1,:} X_{:,n} B_{1n}$$

$$= A_{11} x_{11} B_{11} + A_{12} X_{21} B_{11} + \cdots + A_{11} X_{12} B_{12} + A_{12} B_{22} B_{12} + \cdots + A_{1m} X_{mn} B_{1n}$$

which equals to the first row of $B \otimes A \cdot vec(x)$

Simlarly $(Ax)_{p,:} B_{q,:}$ is the pq^{th} row of $B \otimes A \cdot vec(x)$

That is to say: $Vec(AxB^T) = B \otimes A \cdot vec(x)$

2. (a) To find the eigenvalues of B: $\det(B - I\lambda) = \det(A + xx' - I\lambda) = 0$

$\lambda \neq A_{ii}$ for any i, then $A - \lambda I$ is invertible

$\det(A + xx' - \lambda I) = \det((A - \lambda I)(I + \underbrace{(A-\lambda I)^{-1}}_{w} xx'))$

Based on the property that $\det(I + wx') = 1 + w'x$, the representation can be simplified to:

$\det(A - \lambda I) \cdot (1 + x'(A - \lambda I)^{-1} x)$

Therefore $x'(A - \lambda I)^{-1} x = -1$ must be satisfied.

The eigenvalues of B must satisfy: $x'(\lambda I - A)^{-1} x = 1$

Assume that the eigenvalues of A are: $a_{11}, a_{22}, a_{33} \cdots a_{nn}$

$$\frac{x_1^2}{\lambda - a_{11}} + \frac{x_2^2}{\lambda - a_{22}} + \cdots + \frac{x_n^2}{\lambda - a_{nn}} = 1$$

(b) The eigenvalues of A are: $a_{11} = 1 \quad a_{22} = 2 \quad a_{33} = 3$

To satisfy $x'(\lambda I - A)^{-1} x = 1$

$$\begin{bmatrix} 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\lambda-1} & 0 & 0 \\ 0 & \frac{1}{\lambda-2} & 0 \\ 0 & 0 & \frac{1}{\lambda-3} \end{bmatrix} \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = 1$$

$$\lambda^3 - 9\lambda^2 + 23\lambda - 17 = 0$$

Sovling it by Julia: $\lambda_1 = 5.21 \quad \lambda_2 = 2.46 \quad \lambda_3 = 1.3$

3. (a) Denote $A = [q_1 \ q_2] \quad x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$

The goal is to find $\min \| Ax - b \|_2$.

According to LLS solution, $\hat{x} = A^+ b$

The SVD of A can be written as $A = \sum_{i=1}^{2} 1 \cdot q_i \cdot e_i^T$

Therefore $A^+ = \sum_{i=1}^{2} 1 \cdot e_i \cdot q_i^T$

$$\hat{x} = \sum_{i=1}^{2} 1 \cdot e_i \cdot q_i^T \cdot b = A^T \cdot b$$

(b) $r = b - A\hat{x} = b - A\cdot A^{+}\cdot b = b\cdot(I - A\cdot A^{+})$

$$A'r = A'(I - A\cdot A^{+})b$$
$$= [A' - A'AA^{+}]\,b$$

Since $A' = A'AA^{+}$

$$A'(b - A\hat{x}) = 0\cdot b = 0 \Rightarrow \begin{cases} q_1'\cdot r = 0 \\ q_2'\cdot r = 0 \end{cases}$$

r is orthogonal to both q_1 and q_2

4. $A = U_r\Sigma_r V_r' \Rightarrow A^{+} = V_r\Sigma_r^{-1}U_r'$ (compact SVD)

$$(A^{+}b)'(I - A^{+}A)x = b'U_r\Sigma_r^{-1}V_r'(I - V_rV_r')x = 0$$

Therefore $(I - A^{+}A)x$ and $A^{+}b$ are orthogonal vectors.

5. (a) Based on the definition of convex function, if $f(x) = \|Ax - b\|_2$ is a convex function:

$$\|A(ax_1 + (1-a)x_2) - b\|_2 \leq a\|Ax_1 - b\|_2 + (1-a)\|Ax_2 - b\|_2$$

To prove the inequality equation:

As for left side: $\|A(ax_1 + (1-a)x_2) - b\|_2 = \|A(ax_1 + (1-a)x_2) - b(a + 1 - a)\|_2$

$$= \|(Aax_1 - ab) + [A(1-a)x_2 - b(1-a)]\|_2$$

Given the property that $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$, the inequality:

$$\|(Aax_1 - ab) + [A(1-a)x_2 - b(1-a)]\|_2 \leq \|a(Ax_1 - b)\|_2 + \|(1-a)(Ax_2 - b)\|_2$$

$$\leq a\|Ax_1 - b\|_2 + (1-a)\|Ax_2 - b\|_2$$

(b) Set two matrix M, N

We have to prove that:

$$\sigma_1(aM + (1-a)N) \leq a\,\sigma_1(M) + (1-a)\,\sigma_1(N)$$

Based on the hint: $\sigma_1(aM + (1-a)N) = \max\limits_{\|u\|_2=1} \|(aM+(1-a)N)\cdot u\|_2$

$$\leq \max_{\|u\|_2=1}\left(\|aMu\|_2 + \|(1-a)Nu\|_2\right) = \max_{\|u\|_2=1}\left(a\|Mu\|_2 + (1-a)\|Nu\|_2\right)$$

$$= \max_{\|u\|_2=1} a\|Mu\|_2 + \max_{\|u\|_2=1} (1-a)\|Nu\|_2 = a\,\sigma_1(M) + (1-a)\,\sigma_1(N)$$

Therefore, $\sigma_1(\cdot)$ is a convex function.

6.(a) $x = \begin{bmatrix} 5 \\ 15 \\ 25 \\ 35 \\ 45 \\ 55 \\ 65 \\ 75 \end{bmatrix}$ income

math $= y = [457 \quad 465 \quad 474 \quad 488 \quad 501 \quad 509 \quad 515 \quad 521 \quad 534 \quad 564]$

$\min \|y - \beta_0\|_2$

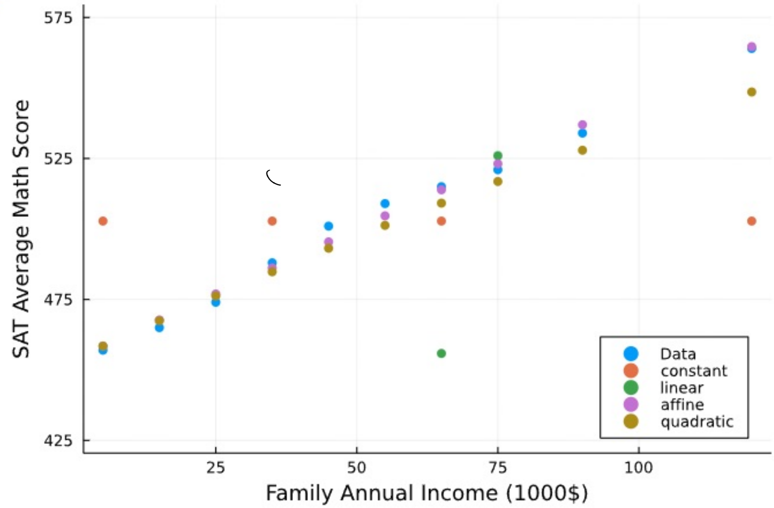
$\min \|y - \beta_1 x\|_2$

$\min \|y - \beta_1 x - \beta_0\|_2$

$\min \|y - \beta_2 x^2 - \beta_1 - \beta_0\|_2$

```
In  [2]: A0=ones(length(income),3)
         A0[:,2]=zeros(length(income))
         A0[:,3]=zeros(length(income))

         A1=zeros(length(income),3)
         A1[:,2]=income

         A2=ones(length(income),3)
         A2[:,2]=income
         A2[:,3]=zeros(length(income))

         A3=ones(length(income),3)
         A3[:,2]=income
         A3[:,3]=income.^2

         X0=A0\math
         X1=A1\math
         X2=A2\math
         X3=A3\math
         @show X0,X1,X2,X3

(X0, X1, X2, X3) = ([502.79999999999995, 0.0, 0.0], [0.0, 7.013164556962025, 0.0], [453.8597721297108, 0.9234005258545132, 0.0], [451.11444428602186, 1.0584447335413263, -0.0011169658642309185])

Out[2]: ([502.79999999999995, 0.0, 0.0], [0.0, 7.013164556962025, 0.0], [453.8597721297108, 0.9234005258545132, 0.0], [451.11444428602186, 1.0584447335413263, -0.0011169658642309185])
```

Out[3]:



(b)

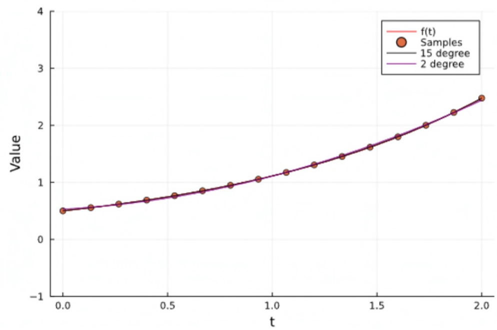
Constant Fit	Linear Fit	Affine Fit	Quadratic Fit
502.79999999999995	0	453.8597721297108	451.11444428602186
0	7.013164556962025	0.9234005258545132	1.0584447335413263
0	0	0	-0.0011169658642309

7. (a)

In [6]:

```
using LinearAlgebra

T = LinRange(0,2,16)

f = 0.5 * exp.(0.8*T)

A15 = [TT^j for TT in T, j=0:length(T)-1]
A2 = A15[:,1:3]
X15 = pinv(A15)*f
X2 = pinv(A2)*f

t = LinRange(0,2,501)
y= 0.5*exp.(0.8*t)

using Plots
plot(t, y; color =:red, label = "f(t)", ylim=(-1,4))
scatter!(T,f;marker=:circle, label="Samples")

a15 = [tt^j for tt in t, j=0:length(T)-1]
a2 = a15[:,1:3]

plot!(t,(a15*X15),color=:black,label="15 degree")
plot!(t,(a2*X2),color=:purple,label="2 degree", xaxis="t", yaxis="Value")
```

Out[6]:

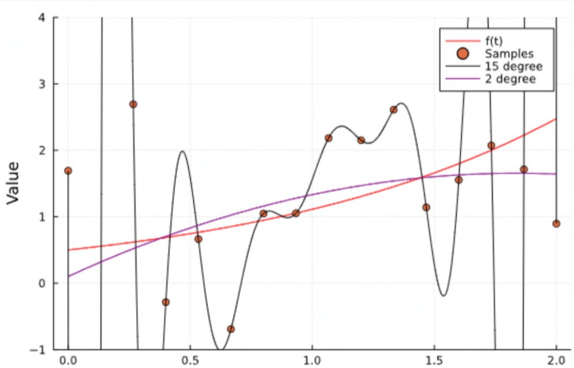


$$y \triangleq \begin{bmatrix} f(t_1) \\ \vdots \\ f(t_{16}) \end{bmatrix} \qquad A \triangleq \begin{bmatrix} 1 & t_1 & t_1^2 & \cdots & t_1^d \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & t_{16} & t_{16}^2 & \cdots & t_{16}^d \end{bmatrix} \in \mathbb{R}^{16\times(d+1)} \qquad x \triangleq \begin{bmatrix} x_1 \\ \vdots \\ x_{d+1} \end{bmatrix} \in \mathbb{R}^{d+1}$$

As the above diagram shows, the least-square coefficients obtained when $d=2$ and $d=15$ are almost the same.

(b)

```
using LinearAlgebra
using Random: seed!

seed!(3); e=randn(length(T))

T = LinRange(0,2,16)

f = 0.5 * exp.(0.8*T) + e

A15 = [TT^j for TT in T, j=0:length(T)-1]
A2 = A15[:,1:3]
X15 = pinv(A15)*f
X2 = pinv(A2)*f

t = LinRange(0,2,501)
y= 0.5*exp.(0.8*t)

using Plots
plot(t, y; color =:red, label = "f(t)", ylim=(-1,4))
scatter!(T,f;marker=:circle, label="Samples")

a15 = [tt^j for tt in t, j=0:length(T)-1]
a2 = a15[:,1:3]

plot!(t,(a15*X15),color=:black,label="15 degree")
plot!(t,(a2*X2),color=:purple,label="2 degree", xaxis="t", yaxis="Value")
```



When noise is added, polynomial for $d=15$ passes each sample point precisely. The error is zero, but large relative to the function.

(c)

polynomial degree:	d = 2	d = 15
Residual norm $\|A\hat{x}(b)-b\|_2$ noiseless (a)	0.3502105199189893	0.0001622183167221819
Residual norm $\|A\hat{x}(y)-y\|_2$ noisy (b)	6.191069208724523	1047.713599472052
Fitting error $\|A\hat{x}(y)-b\|_2$	4.641460505926311	4.641460505926311

(d) because $A_{15} = [A_2 \quad A_{3:15}]$ so $\mathcal{R}(A_2) \subseteq \mathcal{R}(A_{15})$