Student : NINH DO
SID # 25949105

CS 189 Fall 2017 — Introduction to Machine Learning — HW1

This homework is due **Friday, September 1 at 12pm noon.**

1 Getting Started

You may typeset your homework in latex or submit neatly handwritten and scanned solutions. Please make sure to start each question on a new page, as grading (with Gradescope) is much easier that way! Deliverables:

1. Submit a PDF of your writeup to assignment on Gradescope, "HW1 Write-Up"
2. Submit all code needed to reproduce your results, "HW1 Code".
3. Submit your test set evaluation results, "HW1 Test Set".

After you've submitted your homework, be sure to watch out for the self-grade form.

(a) Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. In case of course events, just describe the group. How did you work on this homework? Any comments about the homework?

None . I work alone
Comments : The homework is totally inappropriate for the first week and that short amount of time.

(b) Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions.
I have credited all external sources in this write up.

Problem # 2

(a) We denote cards $C_1, C_2, \ldots, C_{50}$. E_i is the expected # of days to take card i-th that is different from previous cards. we have:

First card: $E_1 = \frac{50}{50}$

Second card: $P_2 = \frac{49}{50} \Rightarrow E_2 = \frac{50}{49}$

Third card $P_3 = \frac{48}{50} \Rightarrow E_3 = \frac{50}{48}$

50-th card: $P_{50} = \frac{1}{50} \Rightarrow E_{50} = \frac{50}{1}$

$\Rightarrow$ The total expected # of days to win:

$$E = \sum_{i=1}^{50} E_i = 50\left(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{50}\right)$$

(b) Assume we have x cards:

$$P(\text{win}) = \frac{x}{d} = 1 - \delta_w$$

$$\Rightarrow x = d(1-\delta_w)$$

(c) At the end of αd days, he has x random distinct cards.

If $x = d \rightarrow$ he wins

$x < d \rightarrow$ he loses

That he loses means he can never pick up at least 1 specific card during αd days, we call this card is C.

For every day, the probability that he fails to pick C is $\frac{d-1}{d}$

$\Rightarrow$ for αd days, $P(\text{lose}) = \left(\frac{d-1}{d}\right)^{\alpha d}$

$$\Rightarrow P(\text{win}) = 1 - \left(\frac{d-1}{d}\right)^{\alpha d}$$

(d) $d \uparrow \Rightarrow \left(\frac{d-1}{d}\right)^{\alpha d} \nearrow \Rightarrow 1-\left(\frac{d-1}{d}\right)^{\alpha d} \downarrow$ more likely to lose

(e) Let the size of training set be x. Domain size d

$$\Rightarrow \quad P(\text{successful estimation}) = \frac{x}{d} = 1-\delta$$

$$\Rightarrow \quad x = d(1-\delta)$$

Problem # 3

(a) We know that there is no sample X_i within the interval (T_{max}, θ)

$\rightarrow$ The probability of failing to sample in this interval is $1-\varepsilon$

For n samples $\rightarrow$ $P(T_{max} - \theta > \varepsilon) = (1-\varepsilon)^n$

(b) Similar to question (a), $P(\theta - T_{min} > \varepsilon) = (1-\varepsilon)^n$

(c) $P(T_{max} - \theta > \varepsilon \text{ and } \theta - T_{min} > \varepsilon)$

$$= P(T_{max} - \theta > \varepsilon) * P(\theta - T_{min} > \varepsilon)$$

$$= (1-\varepsilon)^{2n} \geq 1 - \delta$$

$$\Rightarrow n \geq \frac{1}{2} \log_{(1-\varepsilon)} (1-\delta)$$

(d) Method to estimate θ

1. We choose n samples so that they are uniformly distributed on the interval $(\frac{1}{4}, \frac{3}{4})$. Call $X_1 < X_2 < \dots < X_n$

2. Evaluate $f(X_1), f(X_2), \dots, f(X_n)$

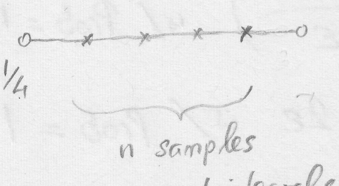
3. Loop through $f(x_1) \rightarrow f(X_n)$, if it changes value from $0 \rightarrow 1$, say $f(X_{i-1}) = 0$, $f(X_i) = 1$, then

$$X_{i-1} \leq \theta < X_i$$

and we know that $T_{min} = X_{i-1}$, $T_{max} = X_i$

The samples are equal-distant from each other

$$\Rightarrow X_i - X_{i-1} = 2\varepsilon$$



$$\frac{3/4 - 1/4}{n+1} = 2\varepsilon$$

$$\Rightarrow n = \frac{0.5}{2\varepsilon} - 1 \quad \text{rounded up}$$

If we want some probability, we can introduce $P = 1-\delta$ into the nominator, i.e. $n = \dfrac{0.5(1-\delta)}{2\varepsilon} - 1$

Otherwise $n = \dfrac{0.5}{2\varepsilon} - 1$ for accuracy of 2ε w/ Prob $= 1$

(e) Method to estimate θ: binary search

1. We take the first sample at the middle of the interval $(\frac{1}{4}, \frac{3}{4})$. Say $X_1 = 0.5$

2. Evaluate $f(X_1)$. If $f(X_1) = 0$, reduce the interval by half into $(X_1, \frac{3}{4})$; otherwise, reduce the interval by half into $(\frac{1}{4}, X_1)$

3. Repeat step 1 & 2 until the interval, say (X_{i-1}, X_i), reduced into size 2ε, then:

$$X_{i-1} \leq \theta < X_i$$

and we know that $T_{min} = X_{i-1}$ $T_{max} = X_i$

The samples are logarithmically distributed, so

$$\frac{\frac{3}{4} - \frac{1}{4}}{2^n} = 2\varepsilon$$

$$\Rightarrow 2^n = \frac{0.5}{2\varepsilon}$$

$$\Rightarrow n = \log_2\left(\frac{0.5}{2\varepsilon}\right) \text{ rounded up}$$

If we want some probability $\rightarrow$ $n = \log_2\left(\dfrac{0.5(1-\delta)}{2\varepsilon}\right)$ w/ Prob $= 1-\delta$

otherwise $n = \log_2\left(\dfrac{0.5}{2\varepsilon}\right)$ for accuracy of 2ε w/ Prob $= 1$

(f) Comparison :

	n	accur.	prob
random	$\frac{1}{2}\log_{(1-\varepsilon)}(1-\delta)$	ε	$1-\delta$
deterministic	$\frac{0.5}{2\varepsilon}-1$	ε	1
adaptive	$\log_2\left(\frac{0.5}{2\varepsilon}\right)$	ε	1

If we fix ε & δ :

	n	accur.	prob
random	$\frac{1}{2}\log_{1-\varepsilon}(1-\delta)$	ε	$1-\delta$
deterministic	$\frac{0.5(1-\delta)}{2\varepsilon}-1$	ε	$1-\delta$
adaptive	$\log_2\left(\frac{0.5(1-\delta)}{2\varepsilon}\right)$	ε	$1-\delta$

Thus, adaptive is the most effective, then deterministic, random at last.

Problem #4 : Eigenvalue & Eigenvector Review

a.) Compute the right & left eigenvalues & eigenvectors

i.) $A = \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}$

Right

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

$$(3-\lambda)^2 - 2 = 0$$

$$(3-\lambda-\sqrt{2})(3-\lambda+\sqrt{2}) = 0$$

$$\Rightarrow \lambda_1 = 3-\sqrt{2} \qquad \lambda_2 = 3+\sqrt{2}$$

$$\lambda_1 : \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (3-\sqrt{2})\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 3x_1 + 2x_2 = (3-\sqrt{2})x_1 \\ x_1 + 3x_2 = (3-\sqrt{2})x_2 \end{cases}$$

$$x_1 = -\sqrt{2}\, x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -\sqrt{2}/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$\lambda_2 : \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (3+\sqrt{2})\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 3x_1 + 2x_2 = (3+\sqrt{2})x_1 \\ x_1 + 3x_2 = (3+\sqrt{2})x_2 \end{cases}$$

$$x_1 = \sqrt{2}\, x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} \sqrt{2}/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Left

$$\begin{vmatrix} 3+\lambda & 2 \\ 1 & 3+\lambda \end{vmatrix} = 0$$

$$(3+\lambda)^2 - 2 = 0$$

$$(3+\lambda-\sqrt{2})(3+\lambda+\sqrt{2}) = 0$$

$$\Rightarrow \lambda_3 = \sqrt{2}-3 \qquad \lambda_4 = -\sqrt{2}-3$$

$$\lambda_3 : \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -(\sqrt{2}-3)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 3x_1 + 2x_2 = (3-\sqrt{2})x_1 \\ x_1 + 3x_2 = (3-\sqrt{2})x_2 \end{cases}$$

$$x_1 = -\sqrt{2}\, x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix} -\sqrt{2} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -\sqrt{2}/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$\lambda_4 : \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -(-\sqrt{2}-3)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 3x_1 + 2x_2 = (\sqrt{2}+3)x_1 \\ x_1 + 3x_2 = (\sqrt{2}+3)x_2 \end{cases}$$

$$x_1 = \sqrt{2}\, x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} \sqrt{2}/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

ii) $B = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$

Right

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0$$

$$(5-\lambda)^2 - 4 = 0$$

$$(5-\lambda-2)(5-\lambda+2) = 0$$

$$\lambda_1 = 3 \quad \text{or} \quad \lambda_2 = 7$$

$$\lambda_1 : \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 5x_1 + 2x_2 = 3x_1 \\ 2x_1 + 5x_2 = 3x_2 \end{cases}$$

$$2x_1 + 2x_2 = 0$$

$$\Rightarrow \underline{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\lambda_2 : \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 7\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 5x_1 + 2x_2 = 7x_1 \\ 2x_1 + 5x_2 = 7x_2 \end{cases}$$

$$2x_2 = 2x_1$$

$$\Rightarrow \underline{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Left

$$\begin{vmatrix} 5+\lambda & 2 \\ 2 & 5+\lambda \end{vmatrix} = 0$$

$$(5+\lambda)^2 - 4 = 0$$

$$(5+\lambda-2)(5+\lambda+2) = 0$$

$$\lambda_3 = -3 \qquad \lambda_4 = -7$$

$$\lambda_3 : \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = +3\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 5x_1 + 2x_2 = +3x_1 \\ 2x_1 + 5x_2 = +3x_2 \end{cases}$$

$$2x_1 + 2x_2 = 0$$

$$\Rightarrow \underline{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\lambda_4 : \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = +7\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 5x_1 + 2x_2 = 7x_1 \\ 2x_1 + 5x_2 = 7x_2 \end{cases}$$

$$2x_1 = 2x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

iii) $A^2 = \begin{bmatrix}3&2\\1&3\end{bmatrix}\begin{bmatrix}3&2\\1&3\end{bmatrix} = \begin{bmatrix}11&12\\6&11\end{bmatrix}$

Right

$$\begin{vmatrix}11-\lambda & 12\\ 6 & 11-\lambda\end{vmatrix} = 0$$

$$(11-\lambda)^2 - 72 = 0$$

$$(11-\lambda-6\sqrt{2})(11-\lambda+6\sqrt{2}) = 0$$

$$\Rightarrow \lambda_1 = 11-6\sqrt{2} \qquad \lambda_2 = 11+6\sqrt{2}$$

λ_1 : $\begin{bmatrix}11&12\\6&11\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = (11-6\sqrt{2})\begin{bmatrix}x_1\\x_2\end{bmatrix}$

$$\begin{cases}11x_1 + 12x_2 = (11-6\sqrt{2})x_1\\ 6x_1 + 11x_2 = (11-6\sqrt{2})x_2\end{cases}$$

$$6x_1 = -6\sqrt{2}\,x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix}-\sqrt{2}\\1\end{bmatrix} \rightarrow \begin{bmatrix}-\sqrt{2}/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

λ_2 : $\begin{bmatrix}11&12\\6&11\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = (11+6\sqrt{2})\begin{bmatrix}x_1\\x_2\end{bmatrix}$

$$\begin{cases}11x_1 + 12x_2 = (11+6\sqrt{2})x_1\\ 6x_1 + 11x_2 = (11+6\sqrt{2})x_2\end{cases}$$

$$6x_1 = 6\sqrt{2}\,x_2$$

$$\Rightarrow \underline{x} = \begin{bmatrix}\sqrt{2}\\1\end{bmatrix} \rightarrow \begin{bmatrix}\sqrt{2}/3\\1/3\end{bmatrix}$$

Left :

$$\begin{vmatrix}11+\lambda & 12\\ 6 & 11+\lambda\end{vmatrix} = 0$$

$$(11+\lambda)^2 - 72 = 0$$

$$(11+\lambda-6\sqrt{2})(11+\lambda+6\sqrt{2}) = 0$$

$$\Rightarrow \lambda_3 = -11+6\sqrt{2}$$

$$\lambda_4 = -11-6\sqrt{2}$$

λ_3 : $\underline{x} = \begin{bmatrix}-\sqrt{2}/3\\1/\sqrt{3}\end{bmatrix}$

λ_4 : $\underline{x} = \begin{bmatrix}\sqrt{2}/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$

iv) $B^2 = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix}$

Right

$$\begin{vmatrix} 29-\lambda & 20 \\ 20 & 29-\lambda \end{vmatrix} = 0$$

$(29-\lambda)^2 - 20^2 = 0$

$(29-\lambda-20)(29-\lambda+20) = 0$

$\lambda_1 = 9 \qquad \lambda_2 = 49$

$\lambda_1 : \quad \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 9 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\begin{cases} 29x_1 + 20x_2 = 9x_1 \\ 20x_1 + 29x_2 = 9x_2 \end{cases}$$

$20x_1 + 20x_2 = 0$

$\Rightarrow \quad \underline{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$\lambda_2 : \quad \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 49 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$$\begin{cases} 29x_1 + 20x_2 = 49x_1 \\ 20x_1 + 29x_2 = 49x_2 \end{cases}$$

$20x_1 = 20x_2$

$\Rightarrow \quad \underline{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

Left :

$$\begin{vmatrix} 29+\lambda & 20 \\ 20 & 29+\lambda \end{vmatrix} = 0$$

$(29+\lambda)^2 - 20^2 = 0$

$(29+\lambda-20)(29+\lambda+20) = 0$

$\lambda_3 = -9 \qquad \lambda_4 = -49$

$\lambda_3 : \quad \underline{x} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

$\lambda_4 : \quad \underline{x} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

v)

$$AB = \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 19 & 16 \\ 11 & 17 \end{bmatrix}$$

Right

$$\begin{vmatrix} 19-\lambda & 16 \\ 11 & 17-\lambda \end{vmatrix} = 0$$

$$(19-\lambda)(17-\lambda) - 16 \times 11 = 0$$

$$\lambda^2 - 36\lambda + 147 = 0$$

$$\Delta' = 18^2 - 147 = 177$$

$$\lambda_1 = 18 - \sqrt{177} \qquad \lambda_2 = 18 + \sqrt{177}$$

λ_1 : $$\begin{bmatrix} 19 & 16 \\ 11 & 17 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (18-\sqrt{177})\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 19x_1 + 16x_2 = (18-\sqrt{177})x_1 \\ 11x_1 + 17x_2 = (18-\sqrt{177})x_2 \end{cases}$$

$$11x_1 = (1-\sqrt{177})x_2$$

$$\Rightarrow \quad \underline{x} = \begin{bmatrix} 1-\sqrt{177} \\ 11 \end{bmatrix} \to \begin{bmatrix} (1-\sqrt{177})/\sqrt{299-2\sqrt{177}} \\ 11/\sqrt{299-2\sqrt{177}} \end{bmatrix}$$

λ_2 : $$\begin{bmatrix} 19 & 16 \\ 11 & 17 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (18+\sqrt{177})\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\begin{cases} 19x_1 + 16x_2 = (18+\sqrt{177})x_1 \\ 11x_1 + 17x_2 = (18+\sqrt{177})x_2 \end{cases}$$

$$11x_1 = (1+\sqrt{177})x_2$$

$$\Rightarrow \quad \underline{x} = \begin{bmatrix} 1+\sqrt{177} \\ 11 \end{bmatrix} \to \begin{bmatrix} (1+\sqrt{177})/\sqrt{299+2\sqrt{177}} \\ 11/\sqrt{299+2\sqrt{177}} \end{bmatrix}$$

left

$$\begin{vmatrix} 19+\lambda & 16 \\ 11 & 17+\lambda \end{vmatrix} = 0$$

$$(19+\lambda)(17+\lambda) + 11 \times 16 = 0$$

$$\lambda^2 + 36\lambda + 147 = 0$$

$$\Delta' = 18^2 - 147 = 177$$

$$\lambda_3 = -18 + \sqrt{177} \qquad \lambda_4 = -18 - \sqrt{177}$$

λ_3 : $\underline{x} =$

λ_4 : $\underline{x} =$

vi) $BA = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 17 & 16 \\ 11 & 19 \end{bmatrix}$

Right

$\begin{vmatrix} 17-\lambda & 16 \\ 11 & 19-\lambda \end{vmatrix} = 0$

$(17-\lambda)(19-\lambda) - 11 \times 16 = 0$

$\lambda^2 - 36\lambda + 147 = 0$

$\Rightarrow \lambda_1 = 18 - \sqrt{177} \quad \lambda_2 = 18 + \sqrt{177}$

$\lambda_1 : \begin{bmatrix} 17 & 16 \\ 11 & 19 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (18-\sqrt{177}) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\begin{cases} 17x_1 + 16x_2 = (18-\sqrt{177})x_1 \\ 11x_1 + 19x_2 = (18-\sqrt{177})x_2 \end{cases}$

$16x_2 = (1-\sqrt{177})x_1$

$\Rightarrow \underline{x} = \begin{bmatrix} 16 \\ 1-\sqrt{177} \end{bmatrix} \to \begin{bmatrix} 16/\sqrt{434-2\sqrt{177}} \\ (1-\sqrt{177})/\sqrt{434-2\sqrt{177}} \end{bmatrix}$

$\lambda_2 : \begin{bmatrix} 17 & 16 \\ 11 & 19 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (18+\sqrt{177}) \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

$\begin{cases} 17x_1 + 16x_2 = (18+\sqrt{177})x_1 \\ 11x_1 + 19x_2 = (18+\sqrt{177})x_2 \end{cases}$

$16x_2 = (1+\sqrt{177})x_1$

$\Rightarrow \underline{x} = \begin{bmatrix} 16 \\ 1+\sqrt{177} \end{bmatrix} \to \begin{bmatrix} 16/\sqrt{434+2\sqrt{177}} \\ (1+\sqrt{177})/\sqrt{434+2\sqrt{177}} \end{bmatrix}$

Left

$\begin{vmatrix} 17+\lambda & 16 \\ 11 & 19+\lambda \end{vmatrix} = 0$

$(17+\lambda)(19+\lambda) - 11 \times 16 = 0$

$\lambda^2 + 36\lambda + 147 = 0$

$\Rightarrow \lambda_3 = -18 + \sqrt{177}$

$\lambda_4 = -18 - \sqrt{177}$

$\lambda_3 : \underline{x} =$

$\lambda_4 : \underline{x} =$

(b) Singular value decompositions (SVD)

(i) $A = \begin{bmatrix} 3 & 2 \\ 1 & 3 \end{bmatrix}$ $\quad AA^T = \begin{bmatrix} 13 & 9 \\ 9 & 10 \end{bmatrix}$ $\quad A^TA = \begin{bmatrix} 10 & 9 \\ 9 & 13 \end{bmatrix}$

AA^T has eigenvectors $\begin{bmatrix} -0.76 \\ -0.65 \end{bmatrix}$, $\begin{bmatrix} -0.65 \\ 0.76 \end{bmatrix}$

A^TA has eigenvectors $\begin{bmatrix} -0.65 & -0.76 \\ -0.76 & 0.65 \end{bmatrix}$

They have eigenvalues : 20.62 & 2.38

$$\Rightarrow \quad A = \begin{bmatrix} -0.76 & -0.65 \\ -0.65 & 0.76 \end{bmatrix} \begin{bmatrix} 4.54 & 0 \\ 0 & 1.54 \end{bmatrix} \begin{bmatrix} -0.65 & -0.76 \\ -0.76 & 0.65 \end{bmatrix}$$

(ii) $B = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ $\quad BB^T = \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix}$ $\quad B^TB = \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix}$

BB^T and B^TB have eigenvectors $\begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix}$

and eigenvalue 49 & 9

$$\Rightarrow \quad B = \begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix} \begin{bmatrix} 7 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix}$$

(iii) $A^2 = \begin{bmatrix} 11 & 12 \\ 6 & 11 \end{bmatrix}$ $\quad A^2 A^{2^T} = \begin{bmatrix} 265 & 198 \\ 198 & 157 \end{bmatrix}$ $\quad A^{2^T} A^T = \begin{bmatrix} 157 & 198 \\ 198 & 265 \end{bmatrix}$

$A^2 A^{2^T}$ has eigenvectors $\begin{bmatrix} -0.79 & -0.61 \\ -0.61 & 0.79 \end{bmatrix}$

$A^{2^T} A^2$ has eigenvectors $\begin{bmatrix} -0.61 & -0.79 \\ -0.79 & 0.61 \end{bmatrix}$

They have eigenvalues 416.2 and 5.8

$$\Rightarrow A^2 = \begin{bmatrix} -0.79 & -0.61 \\ -0.61 & 0.79 \end{bmatrix} \begin{bmatrix} 20.4 & 0 \\ 0 & 2.4 \end{bmatrix} \begin{bmatrix} -0.61 & -0.79 \\ -0.79 & 0.61 \end{bmatrix}$$

(iv) $B^2 = \begin{bmatrix} 29 & 20 \\ 20 & 29 \end{bmatrix}$ $B^2 B^{2^T} = \begin{bmatrix} 1241 & 1160 \\ 1160 & 1241 \end{bmatrix}$ $B^{2^T} B^2 = \begin{bmatrix} 1241 & 1160 \\ 1160 & 1241 \end{bmatrix}$

$B^2 B^{2^T}$ and $B^{2^T} B^2$ have eigenvectors $\begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix}$

and eigenvalues 2401 & 81

$$\Rightarrow B^2 = \begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix} \begin{bmatrix} 49 & 0 \\ 0 & 9 \end{bmatrix} \begin{bmatrix} -0.71 & -0.71 \\ -0.71 & 0.71 \end{bmatrix}$$

(v) $AB = \begin{bmatrix} 19 & 16 \\ 11 & 17 \end{bmatrix}$ $AB(AB)^T = \begin{bmatrix} 617 & 481 \\ 481 & 410 \end{bmatrix}$ $(AB)^T AB = \begin{bmatrix} 482 & 491 \\ 491 & 450 \end{bmatrix}$

$AB(AB)^T$ has eigenvectors $\begin{bmatrix} -0.78 & -0.63 \\ -0.63 & 0.78 \end{bmatrix}$

$(AB)^T AB$ has eigenvector $\begin{bmatrix} -0.68 & -0.73 \\ -0.73 & 0.68 \end{bmatrix}$

they have eigenvalues 1005.5 & 21.5

$$\Rightarrow AB = \begin{bmatrix} -0.78 & -0.63 \\ -0.63 & 0.78 \end{bmatrix} \begin{bmatrix} 31.72 & 0 \\ 0 & 4.63 \end{bmatrix} \begin{bmatrix} -0.68 & -0.73 \\ -0.73 & 0.68 \end{bmatrix}$$

(vi) $BA = \begin{bmatrix} 17 & 16 \\ 11 & 19 \end{bmatrix}$ $BA(BA^T) = \begin{bmatrix} 545 & 491 \\ 491 & 482 \end{bmatrix}$ $(BA)^T BA = \begin{bmatrix} 410 & 481 \\ 481 & 617 \end{bmatrix}$

$BA(BA^T)$ has eigenvectors $\begin{bmatrix} -0.73 & -0.68 \\ -0.68 & 0.73 \end{bmatrix}$

$(BA)^T BA$ has eigenvectors $\begin{bmatrix} -0.63 & -0.78 \\ -0.78 & 0.63 \end{bmatrix}$

They have eigenvalues 1005.5 & 21.5

$$\Rightarrow \quad BA = \begin{bmatrix} -0.73 & -0.68 \\ -0.68 & 0.73 \end{bmatrix} \begin{bmatrix} 31.71 & 0 \\ 0 & 4.64 \end{bmatrix} \begin{bmatrix} -0.63 & -0.78 \\ -0.78 & 0.63 \end{bmatrix}$$

(vi) $C = \begin{bmatrix} 5 & 2 \\ 2 & 5 \\ 3 & 2 \\ 1 & 3 \end{bmatrix}$ $CC^T = \begin{bmatrix} 29 & 20 & 19 & 11 \\ 20 & 29 & 16 & 17 \\ 19 & 16 & 13 & 9 \\ 11 & 17 & 9 & 10 \end{bmatrix}$ $C^TC = \begin{bmatrix} 39 & 29 \\ 29 & 42 \end{bmatrix}$

CC^T has eigenvectors $\begin{bmatrix} -0.59 & 0.66 & -0.46 & 0.04 \\ -0.6 & -0.59 & -0.13 & -0.53 \\ -0.42 & 0.26 & 0.88 & -0.003 \\ -0.34 & -0.4 & -0.06 & 0.85 \end{bmatrix}$

C^TC has eigenvectors $\begin{bmatrix} -0.69 & 0.73 \\ -0.73 & -0.69 \end{bmatrix}$

The eigenvalues are 69.54 & 11.46

$$\Rightarrow \quad C = \begin{bmatrix} -0.59 & 0.66 & -0.46 & 0.04 \\ -0.6 & -0.59 & -0.13 & -0.53 \\ -0.42 & 0.26 & 0.88 & -0.003 \\ -0.34 & -0.4 & -0.06 & 0.85 \end{bmatrix} \begin{bmatrix} 8.34 & 0 \\ 0 & 3.39 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0.69 & 0.73 \\ -0.73 & -0.69 \end{bmatrix}$$

(c)

$$A\vec{x} = \lambda x$$

$$A\vec{x} - \lambda x = 0$$

$$A\vec{x} - I\lambda x = 0$$

$$(A - I\lambda)\, x = 0$$

At row i-th : $\sum_{j=1}^{n} (A_{ij} - \lambda I_{ij}) x_j = 0$

$$\sum_{j\neq i}^{n} (A_{ij} - \overset{0}{\cancel{\lambda I_{ij}}}) x_j + (A_{ij} - \lambda) x_i = 0$$

$$(\lambda - A_{ij})\, x_i = \sum_{j\neq i}^{n} A_{ij} x_j$$

(d) From question (c) :

$$(\lambda - A_{ij})\, x_i = \sum_{j\neq i}^{n} A_{ij} x_j$$

since $|x_i| \geq |x_j|$

$$\underbrace{\left| (\lambda - A_{ij})\, x_i \right|} = \left| \sum_{j\neq i}^{n} A_{ij} x_j \right| \leq \sum_{j\neq i}^{n} |A_{ij}|\,|x_j| \leq \sum_{j\neq i}^{n} |A_{ij}|\,|x_i|$$

$$|(\lambda - A_{ij})|\,\cancel{|x_i|} \leq \sum_{j\neq i}^{n} |A_{ij}|\,\cancel{|x_i|}$$

$$|\lambda - A_{ij}| \leq \sum_{j\neq i}^{n} |A_{ij}|$$

Problem # 5

Let $\vec{a} = [a_1\ a_2 \ldots a_n]^T$ indicates measuments over time (n time points)

We have m samples $\rightarrow$ m $\vec{a}$'s. The matrix:

$$A = [\ \vec{a}_1 \quad \vec{a}_2 \quad \cdots \quad \vec{a}_n\ ] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where, a_{ij} is the measurement of i-th sample at time point j

The general problem to solve : $\min \| A\vec{x} - \vec{b} \|$

(a) Now we want to construct a sequence of predictors

First measurement ; $\vec{a}_1 x_{11} \simeq \vec{b} \rightarrow \min \| \vec{a}_1 x_{11} - \vec{b} \|$

First two measurements : $[\vec{a}_1\ \vec{a}_2] \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \simeq \vec{b} \rightarrow \min \| [\vec{a}_1\ \vec{a}_2] \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} - \vec{b} \|$

First three measurements :

$$[\vec{a}_1\ \vec{a}_2\ \vec{a}_3] \begin{bmatrix} x_{31} \\ x_{32} \\ x_{33} \end{bmatrix} \simeq \vec{b} \rightarrow \min \left\| [\vec{a}_1\ \vec{a}_2\ \vec{a}_3] \begin{bmatrix} x_{31} \\ x_{32} \\ x_{33} \end{bmatrix} - \vec{b} \right\|$$

...

Generally, we can stack them up to solve for x_{ij} at the same time.

$$\min \left\| \underbrace{\begin{bmatrix} \vec{a}_1 & & & & & & \\ & \vec{a}_1 & \vec{a}_2 & & & & \\ & & & \vec{a}_1 & \vec{a}_2 & \vec{a}_3 & \cdots \\ & & & & \vdots & & \\ & & & & & & \vec{a}_n \end{bmatrix}}_{(m \times n) \times \frac{n(n-1)}{2}} \underbrace{\begin{bmatrix} x_{11} \\ x_{21} \\ x_{22} \\ x_{31} \\ x_{32} \\ x_{33} \\ \vdots \\ x_{nn} \end{bmatrix}}_{\frac{n(n-1)}{2} \times 1} - \underbrace{\begin{bmatrix} \vec{b} \\ \vec{b} \\ \vec{b} \\ \vdots \\ \vec{b} \end{bmatrix}}_{(m \times n) \times 1} \right\|$$

or $\min \| A_{new} \vec{x}_{new} - \vec{b}_{new} \|$. Solve for $\vec{x}_{new}$

$$\vec{x}_{new} = (A_{new}^T A_{new})^{-1} A_{new}^T \vec{b}_{new}$$

(b) Part 1 : predict the next measurements based on the measurements so far

First measurements is $\vec{a}_1$. Second measurement is $\vec{a}_2$

$$\vec{a}_1 x_{11} \simeq \vec{a}_2 \Rightarrow \quad \min \| \vec{a}_1 x_{11} - \vec{a}_2 \| \text{ solve for } x_{11}$$

then $\hat{a}_2 = \vec{a}_1 x_{11} \rightarrow$ innovation $\vec{a}_2 - \hat{a}_2$

First & second measurements are $\vec{a}_1, \vec{a}_2$. Third measurement is $\vec{a}_3$

$$\begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \simeq \vec{a}_3$$

$\Rightarrow$ solve problem $\min \left\| \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} - \vec{a}_3 \right\|$ for $\begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}$

then $\hat{a}_3 = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 \end{bmatrix} \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix} \rightarrow$ innovation $\vec{a}_3 - \hat{a}_3$

Very similar to question (a), we can form the big matrix equation to solve for $\vec{x}_{new}$ based on $\min \| A_{new} \vec{x}_{new} - \vec{a}_{new} \|$

After obtaining $\vec{x}_{new}$ we predict $\hat{a}_i$ based on previous $\vec{a}_{i-1}, \vec{a}_{i-2} \cdots \vec{a}_1$

$$\hat{a} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \vdots \\ \hat{a}_n \end{bmatrix} = \vec{A}_{new} \vec{x}_{new}$$

Part 2: innovation & prediction for $\vec{b}$

Calculate the innovation :

$$\hat{A} = \begin{bmatrix} | & | & | & & | \\ \vec{a}_1 & (\vec{a}_2 - \hat{a}_2) & (\vec{a}_3 - \hat{a}_3) & \cdots & (\vec{a}_n - \hat{a}_n) \\ | & | & | & & | \end{bmatrix}$$

Then we use OLS to learn the weights used to update the prediction for $\vec{b}$ by solving :

$$\min \| \hat{A} \vec{x}_{update} - \vec{b} \|$$

again $\vec{x}_{update} = (\hat{A}^T A)^{-1} \hat{A}^T \vec{b}$

(c) They are equivalent. The matrix $\hat{A}$ (innovation) is actually the orthogonalization of the matrix A, i.e. $\hat{a}_i \cdot \hat{a}_j = 0 \;\forall i \neq j;\; \hat{a}_i, \hat{a}_j \in \hat{A}$

Reasoning: The prediction $\hat{a}_i$ of $\vec{a}_i$ base on $\vec{a}_1, \vec{a}_2, \dots \vec{a}_{i-1}$ is the projection of $\hat{a}_i$ on the plane spanned by $\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_{i-1}\}$

Thus, the innovation $a_i - \hat{a}_i \perp$ the plane.

Since $\hat{A}$ is orthogonal, solving $\vec{x}_{update}$ may be faster:

$$\vec{x}_{update} = \underbrace{(\hat{A}^T \hat{A})^{-1}}_{\text{diagonal matrix}} \hat{A}^T \vec{b}$$

(d) (BONUS) - skip

(e) It does make sense because each component of a vector is mathematically independent w/ the other components, i.e each component can be considered a dimension of n dimensions of n-component vector. Thus the computation of the best linear prediction of a vector is merely the <u>parallel</u> computation of a set of the best predictors of each scalar component.

Problem # 6

(a) $x[t+1] = Ax[t] + Bu[t] + w \qquad t = 1 \to n$

or $x[t+1] \simeq Ax[t] + Bu[t]$

matrix form : $\vec{b} = [\vec{x} \ \ \vec{u}] \begin{bmatrix} A \\ B \end{bmatrix}$

where : $\vec{b} = [x_2 \dots x_n]^T$

$\vec{x} = [x_1 \dots x_{n-1}]^T$

$\vec{u} = [u_1 \dots u_{n-1}]^T$

$\to$ use OLS : $\min \left\| [\vec{x} \ \ \vec{u}] \begin{bmatrix} A \\ B \end{bmatrix} - \vec{b} \right\|$ solve for A, B

See code submitted.

$A = 0.9776 \qquad\qquad B = -0.0878$

(b) Similar to question (a) but we have to stack vectors, matrices

Solve for $\vec{t}$: $\min \| X\vec{t} - \vec{y} \|$

where : $t = \begin{bmatrix} a_1^T \\ b_1^T \\ \vdots \\ a_3^T \\ b_3^T \end{bmatrix}_{18\times 1}$ stack rows of A & B

$y = \begin{bmatrix} x[2] \\ x[3] \\ \vdots \\ x[n] \end{bmatrix}_{3(n-1)\times 1}$ each $x[\cdot]$ has 3 components (vertical)

$$X = \begin{bmatrix} x_1^T \ u_1^T & & & & \\ & x_1^T \ u_1^T & & & \\ & & x_1^T \ u_1^T & & \\ x_2^T \ u_2^T & & & & \\ & x_2^T \ u_2^T & & & \\ & & & \ddots & \\ & & & & x_{n-1}^T \ u_{n-1}^T \end{bmatrix}_{3(n-1)\times 18}$$

$x[i]$ and x_i are the same (3 components)

See code submitted. After solved for t, rearrange to get:

$$A = \begin{bmatrix} 0.1521 & 0.9348 & -0.001 \\ 0.0389 & 0.3096 & 0.8744 \\ -0.5255 & 0.0541 & -0.4702 \end{bmatrix} \qquad B = \begin{bmatrix} 0.0489 & 0.206 & -0.3709 \\ -0.0452 & -0.9786 & 0.1276 \\ 0.911 & -0.4712 & -0.8422 \end{bmatrix}$$

(c)

$$\ddot{x}_i = a x_i + b\dot{x}_i + c x_{i-1} + d\dot{x}_{i-1} + w(t)$$

Let $X_1 = [x_1 \; x_2 \; \dots \; x_n]^T$

$$X_2 = \dot{X}_1 = [\dot{x}_1 \; \dot{x}_2 \; \dots \; \dot{x}_n]^T \Rightarrow \dot{X}_2 = \ddot{X}_1 = [\ddot{x}_1 \; \ddot{x}_2 \; \dots \; \ddot{x}_n]$$

$$X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [x_1 \; x_2 \dots x_n \; \dot{x}_1 \; \dot{x}_2 \dots \dot{x}_n]^T$$

$$\dot{X} = \begin{bmatrix} \dot{X}_1 \\ \dot{X}_2 \end{bmatrix} = \begin{bmatrix} X_2 \\ a x_i + b\dot{x}_i + c x_{i-1} + d\dot{x}_{i-1} + w(t) \end{bmatrix}$$

$$= \begin{bmatrix} X_2 \\ aX_1 + bX_2 + c x_{i-1} + d\dot{x}_{i-1} + w(t) \end{bmatrix}$$

$$\dot{X} = \underbrace{\begin{bmatrix} 0 & 1 \\ a & b \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}}_{X} + \underbrace{\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} x_{i-1} \\ \dot{x}_{i-1} \end{bmatrix}}_{U} + \underbrace{\begin{bmatrix} 0 \\ w(t) \end{bmatrix}}_{W}$$

$$\dot{X} = AX + BU + W$$

(d) $\ddot{x}_i = a x_i + b \dot{x}_i + c x_{i-1} + d \dot{x}_{i-1} + w(t)$

$\Rightarrow \quad \ddot{x}_i \simeq a x_i + b \dot{x}_i + c x_{i-1} + d \dot{x}_{i-1}$

$$\underbrace{\begin{bmatrix} x_i & \dot{x}_i & x_{i-1} & \dot{x}_{i-1} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}}_{\vec{x}} - \underbrace{\ddot{x}_i}_{\vec{b}} = 0$$

each x_i, $\dot{x}_i$, x_{i-1}, $\dot{x}_{i-1}$, $\ddot{x}_i$ is a column of n values

OLS:

$$\min \| A\vec{x} - \vec{b} \|$$

$$\Rightarrow \quad \vec{x} = (A^T A)^{-1} A^T \vec{b}$$

where $\vec{x}$ is the stacked parameter a, b, c, d

(e) See code submitted.

$a = -0.0115$

$b = -0.3147$

$c = 0.0115$

$d = 0.2808$

Description of dynamics:

- When $\dot{x}_i$ is large $\rightarrow \ddot{x}_i \downarrow$ and $\dot{x}_i$ is small, $\ddot{x}_i \uparrow$
- When $\dot{x}_{i-1}$ is large $\rightarrow \ddot{x}_i \uparrow$ and $\dot{x}_{i-1}$ is small, $\ddot{x}_i \downarrow$
- The positions of the car and the previous car have the same effect on $\ddot{x}$ but in much smaller scale.

It does make sense because when the car's speed is large, it tend to slow down and when the previous car's speed is large, it tends to speed up. Thus, the acceleration is positive for the latter case: previous car is fast & far ahead

(f) See code submitted

Problem # 7

(a) Roll two dices, the probability of getting total 8 ?

Space : 36

$8 = (6,2), (5,3), (4,4), (3,5), (2,6)$

$\Rightarrow \quad P(\text{total} = 8) = \frac{5}{36}$

(b) Generalization : Roll two dices (m and n), the probability of getting total k ?

Space : $m \times n$

If $k < 2$ or $k > m+n \rightarrow P = 0$

Assume $m \leq n$

If $2 \leq k \leq m \Rightarrow P = \frac{k-1}{m \times n}$

If $k > m$ and $k \leq n+1 \Rightarrow P = \frac{m}{m \times n} = \frac{1}{n}$

If $n+1 < k \leq m+n \Rightarrow P = \frac{m+n-k+1}{m \times n}$