AI61003 Linear Algebra for AI & ML
Assignment 02 - Problem 08

(a) $A \in \mathbb{R}^{(100-M)\times M}$, $0 < M < 100$

where $A_{ij} = z_{M+i-j}$, $1 \le i \le 100-M$
$1 \le j \le M$

$\theta \in \mathbb{R}^M$, $\theta = [\theta_1 \; \theta_2 \; \dots \; \theta_M]^T$

$\hat{b} \in \mathbb{R}^{100-M}$, where $\hat{b}_i = \hat{z}_{M+i}$, $1 \le i \le 100-M$

So, given A, θ, $\hat{b}$ as defined above, the system of $(100-M)$ linear eq^n s can be expressed as -

$$A\theta = \hat{b} \quad - (1)$$

Consider $b \in \mathbb{R}^{100-M}$, where $b_i = z_{M+i}$
$1 \le i \le 100-M$

The objective f^n J hence becomes.-

$J = \|b - \hat{b}\|_2^2$

Using (1)

$J(\theta) = \|A\theta - b\|_2^2$

Since $J(\theta)$ needs to be minimized, the LS solution of the eq^n $A\theta = b$ will be estimated parameters for the model.

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\; J(\theta) = \underset{\theta}{\operatorname{argmin}}\; \|A\theta - b\|_2^2$$

$\hat{\theta}$ = LS sol$^{\underline{n}}$ of eq$^{\underline{n}}$ $(A\theta = b)$

$\therefore \quad A^T A \hat{\theta} = A^T b$ (normal eq$^{\underline{n}}$) $-(\star)$

① If columns of A are linearly independent then eq$^{\underline{n}}$ $(\star)$ has a unique sol$^{\underline{n}}$ given by $\hat{\theta} = (A^T A)^{-1} A^T b = A^+ b$

② If columns of A are linearly dependent then we can use Tikhonov's regularized inversion by minimizing $J'(\theta)$ rather than $J(\theta)$, where

$$J'(\theta) = \|A\theta - b\|_2^2 + \lambda \|\theta\|_2^2, \quad \lambda \neq 0$$

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}}\, J'(\theta) = (A^T A + \lambda I)^{-1} A^T b.$$

★ Hence, $\hat{\theta}$ is the estimated parameters of the model.

(b)

$$A \;(\in \mathbb{R}^{(100-M)\times M}) = \begin{array}{c|cccccccc} & & & & j^{th}\text{ col.} & & & \\ & z_M & z_{M-1} & \cdots & z_{M+1-j} & \cdots & z_2 & z_1 \\ & z_{M+1} & z_M & \cdots & z_{M+2-j} & \cdots & z_3 & z_2 \\ & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ i^{th}\text{ row} \rightarrow & z_{M+i-1} & z_{M+i-2} & \cdots & z_{M+i-j} & \cdots & z_{i+1} & z_i \\ & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ & z_{99} & z_{98} & \cdots & z_{100-j} & \cdots & z_{101-M} & z_{100-M} \end{array}$$

$$A_{ij} = z_{M+i-j}\,, \quad 1 \le i \le 100-M\,, \quad 1 \le j \le M$$

DATE
PAGE

$$b = \begin{bmatrix} z_{M+1} \\ z_{M+2} \\ \vdots \\ z_{M+i} \\ \vdots \\ z_{100} \end{bmatrix}_{(100-M)} \quad (\in \mathbb{R}^{100-M})$$

$$b_i = z_{M+i} \; ; \quad 1 \le i \le 100-M$$

(c) ① There are M columns in A. k^{th} column of A from right $(1 \le k \le M)$ is a vector with entries in AP, with first term k, common difference 1 and length $(100-M)$.

② $A_{ij} = A_{(i-1)(j-1)}$ for $2 \le i \le (100-M)$ and $2 \le j \le M$.

(d) A is an $(100-M) \times M$ matrix. This means that if $(100-M) < M$ i.e. $M > 50$ then the columns of A will definitely not be linearly independent.

$\therefore$ If $50 < M < 100$, then rank(A) will be $< M$.

Otherwise if $0 < M \le 50$, then rank(A) may or may not be $< M$; this depends on the observed values of z_i's. For instance, if the observed values are constant then rank(A)

will be $< M$. Similarly for appropriate values of M in $(0, 50]$ and z_i's, A can be ~~xxx~~ a full rank matrix, because for M in $(0, 50]$, A will be a tall / square matrix.