20226074 서재민 Writing Assignment (I wrote using MS Word for readability)

What is the size of vector w and y?

the size of w is $d+1$ and the size of y is n because $w = \begin{bmatrix} w_0 \\ \vdots \\ w_d \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

What is the size of matrix A? Write A.

$A = \begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^d \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^d \\ 1 & x_3 & x_3^2 & x_3^3 & \cdots & x_3^d \\ 1 & x_4 & x_4^2 & x_4^3 & \cdots & x_4^d \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^d \end{bmatrix}$ so that the size of matrix A is $n \times (d+1)$

Let $d+1=n$, then, A becomes a square matrix. Compute the determinant of A

If there is $a,\ b$ such that $a \neq b$ and $x_a = x_b$, then A has two same rows so that $\det A = 0$

Assume that there is no $a,\ b$ such that $a \neq b$ and $x_a = x_b$

then, $|A| = \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & x_2^3 & \cdots & x_2^{n-1} \\ 1 & x_3 & x_3^2 & x_3^3 & \cdots & x_3^{n-1} \\ 1 & x_4 & x_4^2 & x_4^3 & \cdots & x_4^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & x_n^3 & \cdots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & x_2^3 - x_1^3 & \cdots & x_2^{n-1} - x_1^{n-1} \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 & x_3^3 - x_1^3 & \cdots & x_3^{n-1} - x_1^{n-1} \\ 0 & x_4 - x_1 & x_4^2 - x_1^2 & x_4^3 - x_1^3 & \cdots & x_4^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & x_n^3 - x_1^3 & \cdots & x_n^{n-1} - x_1^{n-1} \end{vmatrix}$

$= \prod_{i=2}^{n}(x_i - x_1) \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 0 & 1 & x_2 + x_1 & x_2^2 + x_2 x_1 + x_1^2 & \cdots & x_2^{n-2} + x_2^{n-3}x_1 + \cdots + x_2 x_1^{n-3} + x_1^{n-2} \\ 0 & 1 & x_3 + x_1 & x_3^2 + x_3 x_1 + x_1^2 & \cdots & x_3^{n-2} + x_3^{n-3}x_1 + \cdots + x_3 x_1^{n-3} + x_1^{n-2} \\ 0 & 1 & x_4 + x_1 & x_4^2 + x_4 x_1 + x_1^2 & \cdots & x_4^{n-2} + x_4^{n-3}x_1 + \cdots + x_4 x_1^{n-3} + x_1^{n-2} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & x_n + x_1 & x_n^2 + x_n x_1 + x_1^2 & \cdots & x_n^{n-2} + x_n^{n-3}x_1 + \cdots + x_n x_1^{n-3} + x_1^{n-2} \end{vmatrix}$

$= \prod_{i=2}^{n}(x_i - x_1) \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 0 & 1 & x_2 + x_1 & x_2^2 + x_2 x_1 + x_1^2 & \cdots & x_2^{n-2} + x_2^{n-3}x_1 + \cdots + x_2 x_1^{n-3} + x_1^{n-2} \\ 0 & 0 & x_3 - x_2 & x_3^2 - x_2^2 + (x_3 - x_2)x_1 & \cdots & x_3^{n-2} - x_2^{n-2} + (x_3^{n-3} - x_2^{n-3})x_1 + \cdots + (x_3 - x_2)^2 x_1^{n-4} + (x_3 - x_2)x_1^{n-3} \\ 0 & 0 & x_4 - x_2 & x_4^2 - x_2^2 + (x_4 - x_2)x_1 & \cdots & x_4^{n-2} - x_2^{n-2} + (x_4^{n-3} - x_2^{n-3})x_1 + \cdots + (x_4 - x_2)^2 x_1^{n-4} + (x_4 - x_2)x_1^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & x_n - x_2 & x_n^2 - x_2^2 + (x_n - x_2)x_1 & \cdots & x_n^{n-2} - x_2^{n-2} + (x_n^{n-3} - x_2^{n-3})x_1 + \cdots + (x_n - x_2)^2 x_1^{n-4} + (x_n - x_2)x_1^{n-3} \end{vmatrix}$

$= \prod_{i=2}^{n}(x_i - x_1) \times \prod_{i=3}^{n}(x_i - x_2) \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 0 & 1 & x_2 + x_1 & x_2^2 + x_2 x_1 + x_1^2 & \cdots & x_2^{n-2} + x_2^{n-3}x_1 + \cdots + x_2 x_1^{n-3} + x_1^{n-2} \\ 0 & 0 & 1 & x_3 + x_2 + x_1 & \cdots & (x_3^{n-3} + x_3^{n-4}x_2 + \cdots + x_3 x_2^{n-4} + x_2^{n-3}) + (x_3^{n-4} + x_3^{n-5}x_2 + \cdots + x_3 x_2^{n-5} + x_2^{n-4})x_1 + \cdots + (x_3 + x_2)x_1^{n-4} + x_1^{n-3} \\ 0 & 0 & 1 & x_4 + x_2 + x_1 & \cdots & (x_4^{n-3} + x_4^{n-4}x_2 + \cdots + x_4 x_2^{n-4} + x_2^{n-3}) + (x_4^{n-4} + x_4^{n-5}x_2 + \cdots + x_4 x_2^{n-5} + x_2^{n-4})x_1 + \cdots + (x_4 + x_2)x_1^{n-4} + x_1^{n-3} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & x_n + x_2 + x_1 & \cdots & (x_n^{n-3} + x_n^{n-4}x_2 + \cdots + x_n x_2^{n-4} + x_2^{n-3}) + (x_n^{n-4} + x_n^{n-5}x_2 + \cdots + x_n x_2^{n-5} + x_2^{n-4})x_1 + \cdots + (x_n + x_2)x_1^{n-4} + x_1^{n-3} \end{vmatrix}$

$= \prod_{i=2}^{n}(x_i - x_1) \times \prod_{i=3}^{n}(x_i - x_2) \begin{vmatrix} 1 & x_1 & x_1^2 & x_1^3 & \cdots & x_1^{n-1} \\ 0 & 1 & x_2 + x_1 & x_2^2 + x_2 x_1 + x_1^2 & \cdots & x_2^{n-2} + x_2^{n-3}x_1 + \cdots + x_2 x_1^{n-3} + x_1^{n-2} \\ 0 & 0 & 1 & x_3 + x_2 + x_1 & \cdots & (x_3^{n-3} + x_3^{n-4}x_2 + \cdots + x_3 x_2^{n-4} + x_2^{n-3}) + (x_3^{n-4} + x_3^{n-5}x_2 + \cdots + x_3 x_2^{n-5} + x_2^{n-4})x_1 + \cdots + (x_3 + x_2)x_1^{n-4} + x_1^{n-3} \\ 0 & 0 & 0 & x_4 - x_3 & \cdots & x_4^{n-3} - x_3^{n-3} + (x_4^{n-4} - x_3^{n-4})(x_2 + x_1) + \cdots + (x_4^2 - x_3^2)(x_2^{n-5} + x_2^{n-6}x_1 + \cdots + x_2 x_1^{n-6} + x_1^{n-5}) + (x_4 - x_3)(x_2^{n-4} + x_2^{n-5}x_1 + \cdots + x_2 x_1^{n-5} + x_1^{n-4}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & x_n - x_3 & \cdots & x_n^{n-3} - x_3^{n-3} + (x_n^{n-4} - x_3^{n-4})(x_2 + x_1) + \cdots + (x_n^2 - x_3^2)(x_2^{n-5} + x_2^{n-6}x_1 + \cdots + x_2 x_1^{n-6} + x_1^{n-5}) + (x_n - x_3)(x_2^{n-4} + x_2^{n-5}x_1 + \cdots + x_2 x_1^{n-5} + x_1^{n-4}) \end{vmatrix} = \cdots$

Through the above equation, I hypothesized that $\det A = \prod_{j=1}^{n-1}\prod_{i=j+1}^{n}(x_i - x_j)$ (for $n \geq 2$) [for $n = 1$, $\det A = 1$]

Let $P(n)$ denote '$\det A = \prod_{j=1}^{n-1}\prod_{i=j+1}^{n}(x_i - x_j)$'

(basic step) $P(2) \equiv \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = \mathrm{x}_2 - \mathrm{x}_1 \equiv \mathrm{T}$

(Induction step) Assume that $P(k) \equiv T$ for $k \in \mathbb{N}$, which means that $\begin{vmatrix} 1 & c_1 & \cdots & c_1^{k-1} \\ 1 & c_2 & \cdots & c_2^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & c_k & \cdots & c_k^{k-1} \end{vmatrix} = \prod_{j=1}^{k-1}\prod_{i=j+1}^{k}(c_i - c_j)$

then, $\begin{vmatrix} 1 & x_1 & \cdots & x_1^{k-1} & x_1^k \\ 1 & x_2 & \cdots & x_2^{k-1} & x_2^k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_k & \cdots & x_k^{k-1} & x_k^k \\ 1 & x_{k+1} & \cdots & x_{k+1}^{k-1} & x_{k+1}^k \end{vmatrix} = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{k-1} & 0 \\ 1 & x_2 & \cdots & x_2^{k-1} & x_2^{k-1}(x_2 - x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_k & \cdots & x_k^{k-1} & x_k^{k-1}(x_k - x_1) \\ 1 & x_{k+1} & \cdots & x_{k+1}^{k-1} & x_{k+1}^{k-1}(x_{k+1} - x_1) \end{vmatrix}$ [$\because$ (n^{th} column) – ($x_1 \times (n-1)^{th}$ column)]

$= \begin{vmatrix} 1 & x_1 & \cdots & 0 & 0 \\ 1 & x_2 & \cdots & x_2^{k-2}(x_2 - x_1) & x_2^{k-1}(x_2 - x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_k & \cdots & x_k^{k-2}(x_k - x_1) & x_k^{k-1}(x_k - x_1) \\ 1 & x_{k+1} & \cdots & x_{k+1}^{k-1}(x_{k+1} - x_1) & x_{k+1}^{k-1}(x_{k+1} - x_1) \end{vmatrix} = \cdots = \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 \\ 1 & x_2 - x_1 & \cdots & x_2^{k-2}(x_2 - x_1) & x_2^{k-1}(x_2 - x_1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_k - x_1 & \cdots & x_k^{k-2}(x_k - x_1) & x_k^{k-1}(x_k - x_1) \\ 1 & x_{k+1} - x_1 & \cdots & x_{k+1}^{k-1}(x_{k+1} - x_1) & x_{k+1}^{k-1}(x_{k+1} - x_1) \end{vmatrix}$

$= \begin{vmatrix} x_2 - x_1 & \cdots & x_2^{k-2}(x_2 - x_1) & x_2^{k-1}(x_2 - x_1) \\ \vdots & \ddots & \vdots & \vdots \\ x_k - x_1 & \cdots & x_k^{k-2}(x_k - x_1) & x_k^{k-1}(x_k - x_1) \\ x_{k+1} - x_1 & \cdots & x_{k+1}^{k-1}(x_{k+1} - x_1) & x_{k+1}^{k-1}(x_{k+1} - x_1) \end{vmatrix}$ [$\because$ cofactor expansion for first row]

$= \prod_{i=2}^{k+1}(x_i - x_1) \begin{vmatrix} 1 & \cdots & x_2^{k-2} & x_2^{k-1} \\ \vdots & \ddots & \vdots & \vdots \\ 1 & \cdots & x_k^{k-2} & x_k^{k-1} \\ 1 & \cdots & x_{k+1}^{k-1} & x_{k+1}^{k-1} \end{vmatrix} = \prod_{i=2}^{k+1}(x_i - x_1) \times \prod_{j=2}^{k}\prod_{i=j+1}^{k+1}(x_i - x_j)$ [$\because$ $P(k) \equiv T$]

$= \prod_{j=1}^{k}\prod_{i=j+1}^{k+1}(x_i - x_j)$, so that $P(k+1) \equiv T$

$\therefore$ $\det A = \prod_{j=1}^{n-1}\prod_{i=j+1}^{n}(x_i - x_j)$ (for $n \geq 2$) [for $n = 1$, $\det A = 1$]

What is the condition that makes the determinant of A non-zero?

$\det A = \prod_{j=1}^{n-1}\prod_{i=j+1}^{n}(x_i - x_j)$ is non-zero when there is no $a,\ b$ such that $a \neq b$ and $x_a = x_b$

Assume that the determinant of A is non-zero, then, what is the solution of linear equation, $Aw = y$, with respect to w?

'the determinant of A is non-zero' means that A is invertible. so, we can get solution to compute $w = A^{-1}y$

Suppose that $n > d+1$. Then, we cannot comput the inverse of A since A is not a square matrix.

In this case, how can we solve the linear equation $Aw = y$?

with SVD, we can write $A = U\Sigma V^T$ [U : $n \times n$ orthogonal matrix, V : $(d+1) \times (d+1)$ orthogonal matrix, Σ : $n \times (d+1)$ matrix] [Let $rank\, A = r$, then $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$, Σ_r : $r \times r$ diagonal matrix]

notice that $U^T = U^{-1}$ and $V^T = V^{-1}$, and pseudo inverse matrix $\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}$ [$\because$ $\Sigma\Sigma^+\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}\Sigma = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\Sigma = \Sigma$ and $\Sigma^+\Sigma\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}\Sigma^+ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}\Sigma^+ = \Sigma$]

then, $A^+ = V\Sigma^+U^T$ [$\because$ $A^+AA^+ = VE^+U^TUEV^TVE^+U^T = VE^+EE^+U^T = VE^+U^T = A^+$, and $AA^+A = U\Sigma V^T V\Sigma^+ U^T U\Sigma V^T = U\Sigma\Sigma^+\Sigma V^T = U\Sigma V^T = A$]

on the other hand, $A^+A = V\Sigma^+U^TU\Sigma V^T = V\Sigma^+\Sigma V^T = V\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}V^T \neq I_{d+1}$. but, if $rank\, A = d+1$, then we can write $\Sigma = \begin{bmatrix} \Sigma_r \\ 0 \end{bmatrix}$ so that $\Sigma^+\Sigma = [\Sigma_r^{-1} \quad 0]\begin{bmatrix} \Sigma_r \\ 0 \end{bmatrix} = I_{d+1}$

so, if $rank\, A = d+1$, that is, if A has linearly independent columns, than $A^+A = I$ so that $Aw = y \Rightarrow A^+Aw = A^+y \Rightarrow w = A^+y$.

on the other hand, if A has linearly independent columns, A^TA is invertible [$\because$ $Col\, A = (Nul\, A^T)^\perp$ so that $A^TAx = 0 \Rightarrow Ax = 0 \Rightarrow x = 0$]

so, A^+ is existed by $(A^TA)^{-1}A^T$ [$\because$ $A^+A = (A^TA)^{-1}A^TA = I$] $\therefore$ we can solve the linear equation $Aw = y$ by computing $w = A^+y = (A^TA)^{-1}A^Ty$