△ Activity 09 - item 1

a) $Xw = y$, $X = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$, $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$, $y = \begin{bmatrix} 2 \\ -4 \end{bmatrix}$ → $w_1 + w_2 = 2 \rightarrow w_2 = -w_1 + 2$

$-2w_1 - 2w_2 = -4$

$(0, 2)$ → not unique

$(2, 0)$ infinite ans ⇒ $\min \underline{e} = 0$

$\cancel{X\tilde{w} = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} -1+3 \end{bmatrix}}$

b) $\cancel{\min_{w_1} = (-w_1+2)^2 + w_1^2 = w_1^2 - 4w_1 + 4 + w_1^2}$

$\cancel{= 2w_1^2 - 4w_1 + 4}$

$\cancel{(\text{diff}) \Rightarrow 2 \cdot 2w_1 - 4 = 0}$

$\cancel{w_1 = 1}$

$w_2 = 1$

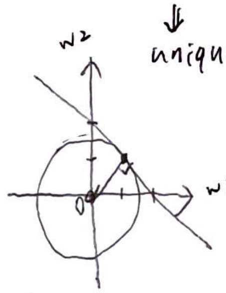
⇓

unique, what!?

$\min \|w\|_2^2$

$X\tilde{w} = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix} = 0$

$\|w\|_2^2 = 2$




c) $\hat{w} = (X^TX + \lambda I)^{-1} X^T y$

$= \left( \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & -2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \end{bmatrix}$

$= \left( \begin{bmatrix} 1+4 & 1+4 \\ 1+4 & 1+4 \end{bmatrix} + \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 2+8 \\ 2+8 \end{bmatrix}$

$= \left( \begin{bmatrix} 5+\lambda & 5 \\ 5 & 5+\lambda \end{bmatrix} \right)^{-1} \begin{bmatrix} 10 \\ 10 \end{bmatrix}$

$= \frac{1}{(5+\lambda)^2 - 25} \begin{bmatrix} 5+\lambda & -5 \\ -5 & 5+\lambda \end{bmatrix} \begin{bmatrix} 10 \\ 10 \end{bmatrix}$

$= \frac{1}{\lambda^2 + 10\lambda} \begin{bmatrix} \cancel{50} + 10\lambda - \cancel{50} \\ -\cancel{50} + \cancel{50} + 10\lambda \end{bmatrix}$

$\tilde{w}_1 + \tilde{w}_2 = 2$

$w_1 + w_2 = 2$

$= \frac{1}{\lambda + 10} \begin{bmatrix} 10 \\ 10 \end{bmatrix}$

$\lambda = 0$, $\frac{1}{\cancel{10}} \begin{bmatrix} \cancel{10} \\ \cancel{10} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \rightarrow 1 + 1 = 2$

$\lambda = 5$, $\frac{1}{\cancel{15}} \begin{bmatrix} \cancel{10} \\ \cancel{10} \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \rightarrow \frac{4}{9} \times 2 = \frac{8}{9}$

$X\hat{w} = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix} \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ -\frac{8}{3} \end{bmatrix} - \begin{bmatrix} 2 \\ -4 \end{bmatrix}$

$\tilde{w}_1 + \tilde{w}_2 = \frac{4}{3}$

$\left(\frac{4}{3} - 2\right)^2 = \frac{4}{9}$

$\frac{2}{3}$

$= \begin{bmatrix} -\frac{2}{3} \\ \frac{4}{3} \end{bmatrix} \Rightarrow \frac{4}{9} + \frac{16}{9} = \frac{20}{9}$

△ Item 2

a) $X = \begin{bmatrix} 1 & r \\ 1 & -r \\ 1 & -r \\ 1 & r \end{bmatrix}$

$x_1 \quad x_2$

$x_1^T x_2 = r - r - r + r = 0$

→ orthogonal to each other

b) $X = U\Sigma$

$4\times2 \quad 2\times2$ → diagonal

orthonormal

$\frac{1}{\cancel{\sqrt{4}}\,2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \cancel{\sqrt{4}}\,2 & 0 \\ 0 & \cancel{\sqrt{4}r}\,2r \end{bmatrix}$

c) $\min \|Xw - y\|_2^2$, $X = U\Sigma$

$w = (X^TX)^{-1} X^T y = (\Sigma^T U^T U \Sigma)^{-1} \Sigma^T U^T y = \Sigma^{-2} \Sigma^T U^T y$

$= \Sigma^{-1} U^T y$

(ortho-normal)

$= \cancel{\left( \begin{bmatrix} 2 & 0 \\ 0 & 2r \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 2r \end{bmatrix} \right)^{-1} \begin{bmatrix} 2 & 0 \\ 0 & 2r \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}}$

d) $\Sigma^{-1} U^T y = w$

$\frac{1}{2\cancel{4}r} \begin{bmatrix} \cancel{2}r & 0 \\ 0 & \cancel{2} \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

$2\times2 \quad 2\times4 \quad 4\times1$

$\frac{1}{2r} \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{\cancel{2}} \begin{bmatrix} \cancel{2}\,1 \\ \cancel{2}\,1 \end{bmatrix} = \frac{1}{2r} \begin{bmatrix} r \\ 1 \end{bmatrix} = w$, as $r \rightarrow 0$

$\cancel{\|w\|_2^2 = \left(\frac{1}{2}\right)^2 + \frac{1}{4}}$

$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2r} \end{bmatrix} = w$ $\quad \|w\|_2^2 = \left(\frac{1}{2}\right)^2 = \frac{1}{4}$

e) $\cancel{\text{if } r = 0.1,\ \|w\|_2^2 = \left(\frac{r}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{0.01}{4} + \frac{1}{4} = \frac{1}{400}}$

$\cancel{r = 10^{-8},\ \|w\|_2^2 = \left(\frac{10^{-8}}{2}\right)^2 + \left(\frac{1}{2}\right)^2 = \frac{1+10^{-16}}{4}}$

~~f)~~ if $r = 0.1$, $\|w\|_2^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{0.2}\right)^2 = \frac{1}{4} + \frac{1}{0.04}\,\frac{100}{4}$

$r = 10^{-8}$, $\|w\|_2^2 = \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2 \cdot 10^{-8}}\right)^2 = \frac{101}{4}$

$= \frac{1}{4} + \frac{10^{16}}{4} = \frac{1+10^{16}}{4}$

f) $w = w_0 + w_E$

$\frac{1}{2r} \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 0.01 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$\frac{1}{4r} \begin{bmatrix} r & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.01 \\ 0.01 \end{bmatrix} \begin{matrix} 10^{-2} \\ 10^{-2} \end{matrix}$

$= \frac{1}{\cancel{4r}} \begin{bmatrix} \cancel{r} \cdot 10^{-2}/4 \\ 10^{-2}/4r \end{bmatrix} = w_E$

$\|w_\varepsilon\|_2^2$, if $r=0.1$, $\|w_\varepsilon\|_2^2 = \left(\frac{10^{-2}}{4}\right)^2 + \left(\frac{10^{-2}\,10^{-1}}{4\cdot 10^{-1}}\right)^2 = \frac{10^{-4}}{16} + \frac{10^{-2}}{16} = \frac{10^{-4}+10^{-2}}{16}$

$r=10^{-8}$, $\|w_\varepsilon\|_2^2 = \left(\frac{10^{-2}}{4}\right)^2 + \left(\frac{\cancel{10^{-2}}\,10^{-6}}{4\cdot 10^{-8}}\right)^2 = \frac{10^{-4}}{16} + \frac{1}{16\cdot 10^{-12}} = \frac{10^{8}+1}{16\cdot 10^{-12}}$

g)

$\cancel{W = w_0 + w_\varepsilon = \begin{bmatrix}\frac12\\ \frac{1}{2r}\end{bmatrix} + \begin{bmatrix}\frac{10^{-2}}{4}\\ \frac{10^{-2}}{4r}\end{bmatrix}}$, $\lambda = 0.1$, $\varepsilon = 0.01$

$\cancel{\text{if } r=0.1,\ \|W\|_2^2 = \frac{101}{4} + \frac{10^{-4}+10^{-2}}{16}}$

$\cancel{r=10^{-8},\ \|W\|_2^2 = \frac{1+10^{16}}{4} + \frac{10^{8}+1}{16\cdot 10^{-12}}}$

$\hat{w} = (X^TX+\lambda I)^{-1}X^T(y_0+y_\varepsilon)$, $X = U\Sigma$

$= (\Sigma^T \cancel{U^T}\cancel{U}\Sigma + \lambda I)^{-1}\Sigma^T U^T(y_0+y_\varepsilon)$

$= (\Sigma^{-1}U^T + \lambda\Sigma^{T}U^{-1})(y_0+y_\varepsilon)$

$= \left(\frac{1}{\cancel{4}r\,2}\underset{2\times2}{\begin{bmatrix}\cancel{2}r & 0\\ 0 & \cancel{2}1\end{bmatrix}}\frac12\underset{2\times4}{\begin{bmatrix}1&1&1&1\\1&-1&-1&1\end{bmatrix}} + \lambda\begin{bmatrix}\cancel{2}1 & 0\\ 0 & \cancel{2}r\end{bmatrix}\frac{1}{\cancel{2}1}\begin{bmatrix}1&1&1&1\\1&-1&-1&1\end{bmatrix}\right)(y_0+y_\varepsilon)$

$= \left(\frac14\begin{bmatrix}r&r&r&r\\1&-1&-1&1\end{bmatrix} + \lambda\begin{bmatrix}1&1&1&1\\r&-r&-r&r\end{bmatrix}\right)(y_0+y_\varepsilon)$

$w_0$ $\qquad w_\varepsilon$

$\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\ \begin{bmatrix}\varepsilon\\0\\0\\0\end{bmatrix}$

$= \frac14\begin{bmatrix}r&r&r&r\\1&-1&-1&1\end{bmatrix}\begin{bmatrix}1\\0\\0\\1\end{bmatrix} + \lambda\begin{bmatrix}1&1&1&1\\r&-r&-r&r\end{bmatrix}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$

$= \frac{1}{\cancel{4}\,2}\begin{bmatrix}\cancel{2}r\\ \cancel{2}1\end{bmatrix} + \lambda\begin{bmatrix}2\\2r\end{bmatrix} = \underbrace{\begin{bmatrix}\frac r2 + \lambda 2\\ \frac12 + 2r\lambda\end{bmatrix}}_{w_0}$

$= \frac14\begin{bmatrix}r&r&r&r\\1&-1&-1&1\end{bmatrix}\begin{bmatrix}\varepsilon\\0\\0\\0\end{bmatrix} + \lambda\begin{bmatrix}1&1&1&1\\r&-r&-r&r\end{bmatrix}\begin{bmatrix}\varepsilon\\0\\0\\0\end{bmatrix}$

$= \underbrace{\frac14\begin{bmatrix}r\varepsilon\\ \varepsilon\end{bmatrix} + \lambda\begin{bmatrix}\varepsilon\\ r\varepsilon\end{bmatrix}}_{w_\varepsilon}$

$\lambda = 10^{-1}$, $\varepsilon = 10^{-2}$,

$w_\varepsilon = \varepsilon\begin{bmatrix}\frac{r\varepsilon}{4} + \lambda\varepsilon\\ \frac{\varepsilon}{4} + \lambda r\varepsilon\end{bmatrix}$

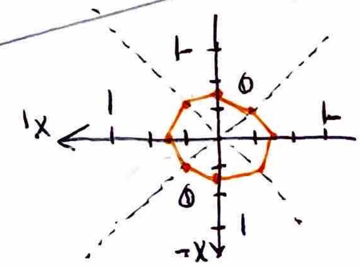
② $r = 10^{-8}$

$\cancel{w_0 = \begin{bmatrix}\frac{10^{-8}}{2} + 10^{-1}\cdot 2\\ \frac12 + 2\cdot10^{-1}\cdot10^{-1}\end{bmatrix}}$

① $r = 10^{-1}$

$w_0 = \begin{bmatrix}\frac{10^{-1}}{2} + 10^{-1}\cdot 2\\ \frac12 + 2\cdot10^{-1}\cdot10^{-1}\end{bmatrix} = \begin{bmatrix}\frac{10^{-1}+4\cdot10^{-1}}{2} = \frac{5\cdot10^{-1}}{2}\\ \frac{1+4\cdot10^{-2}}{4}\end{bmatrix}$

$w_\varepsilon = \begin{bmatrix}\frac{10^{-1}\cdot10^{-2}}{4} + 10^{-1}\cdot10^{-2}\\ \frac{10^{-2}}{4} + 10^{-1}\cdot10^{-1}\cdot10^{-2}\end{bmatrix} = \begin{bmatrix}\frac{5\cdot10^{-3}}{4}\\ \frac{10^{-2}+4\cdot10^{-4}}{4}\end{bmatrix}$



$(\frac{\varepsilon}{1}, \frac{\varepsilon}{1})$

$(\frac{\varepsilon}{1}, 0)$ ⓪

$1 < |{}^2x| + |{}^1x| < \Leftarrow |{}^2x| < |{}^1x|$

① $-1 = |{}^1x| + |{}^2x| < \Leftarrow |{}^1x| < |{}^2x|$ $\Big\}\ \Leftarrow$

$|{}^1x| \Leftarrow |{}^2x| < |{}^1x|$

$|{}^2x| \Leftarrow |{}^1x| < |{}^2x|$ $\Big\}$ (xvw) $= \infty\|x\|$

$|{}^2x| + |{}^1x| = {}^1\|x\|$

② $r = 10^{-8}$

$$w_0 = \begin{bmatrix} \frac{10^{-8}}{2} + 10^{-1} \cdot 2 \\ \frac{1}{2} + 2 \cdot 10^{-8} \cdot 10^{-1} \end{bmatrix} = \begin{bmatrix} \frac{10^{-8} + 4 \cdot 10^{-1}}{2} \\ \frac{1 + 4 \cdot 10^{-9}}{2} \end{bmatrix}$$

$$w_E = 10^{-2} \begin{bmatrix} \frac{10^{-8}}{4} + 10^{-1} \\ \frac{1}{4} + 10^{-1} \cdot 10^{-8} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{10^{-10} + 4 \cdot 10^{-3}}{4} \\ \frac{10^{-2} + 4 \cdot 10^{-11}}{4} \end{bmatrix}$$

① $(w_0, w_E)$

vs

② $(w_0, w_E)$

when $r \downarrow$, $w_0$ & $w_E \downarrow$