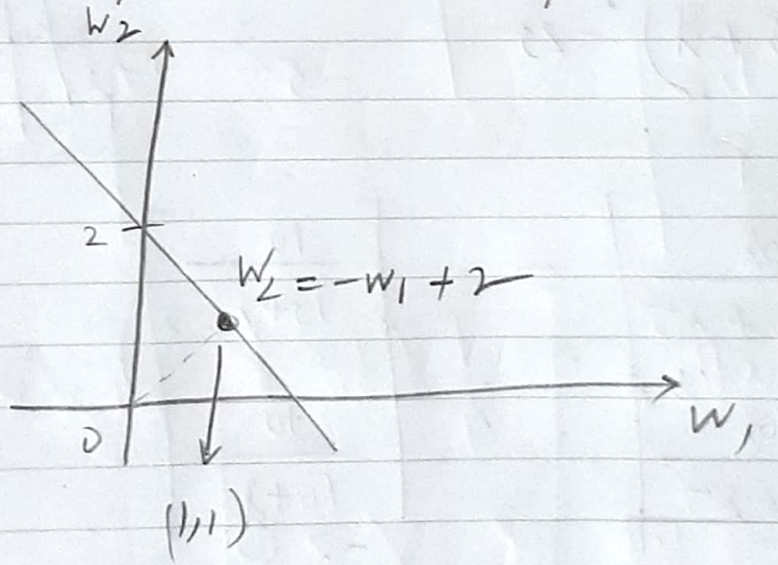
Activity 9

Ayan Deep Hazra

ECE 532

1. a) Since $Xw = y$ also has an exact solution, the least-squares problem has zero-squared error. There are 2 eq^ns and two unknows, but both equations are equivalent to $w_1 + w_2 = 2$.



Solution is not unique

b) The point on the line which is closest to the origin is the solution of the minimum norm.

The point $(1,1)$ is the solution in this case. Every other point has greater $\|w\|_2^2$.

Solution is unique because of the condition of finding min norm solution.

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

$$w = \begin{bmatrix} 5+\lambda & 5 \\ 5 & 5+\lambda \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 10 \end{bmatrix}$$

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

$$= \begin{bmatrix} \dfrac{10\lambda}{10\lambda + \lambda^2} \\ \dfrac{10\lambda}{10\lambda + \lambda^2} \end{bmatrix} = \begin{bmatrix} \dfrac{10}{10+\lambda} \\ \dfrac{10}{10+\lambda} \end{bmatrix}$$

d) The solutions for all λ satisfy $w_1 = w_2$.

when $\lambda \to 0$ at $w_1 = w_2 \to 1$ which has $\|Xw - y\|_2^2 = 0$ & $\|w\|_2^2 = 2$ when $\lambda = 5$, $w_1 = w_2 = 10/15$ which has

Scanned with CamScanner

$\|Xw - y\|_2^2 = 20/9$ & $\|w\|_2^2 = 8/9$.

Then as $\lambda \to \infty$, we have $w_1 = w_2 \to 0$,

so $\|Xw - y\|_2^2 \to \|y\|_2^2 = 20$

& $\|w\|_2^2 = 0$.

As λ increases, the squared error increases, but $\|w_2\|^2$ decreases. The solution to part b) corresponds to $\lambda \to 0$.

2. a) Let the i^{th} column be x_i. Note that $x_1^T x_2 = \gamma - \gamma - \gamma + \gamma = 0$

b) Since the columns of X are orthogonal, we find U by normalizing their 2-norm to unity. This gives $U = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}$

& thus $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 2\gamma \end{bmatrix}$

c) The problem is $\min_w \| U\Sigma w - y\|_2^2$,
so the solution is $w = (\Sigma U^T U \Sigma)^{-1} \Sigma U^T y$
which simplifies to $w = \Sigma^{-1} U^T y$

d) $\Sigma^{-1} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5\gamma^{-1} \end{bmatrix}$ & $U^T y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

so $w = \begin{bmatrix} 1/2 \\ \frac{1}{2\gamma} \end{bmatrix}$

We see that $\|w\|_2^2 = \frac{1}{4}(1+\gamma^{-2})$ so
$\|w\|_2^2$ blows up as $\gamma \to 0$

e) i) for $\gamma = 0.1$, $\|w\|_2^2 = \frac{101}{4}$
condition number is 10

ii) for $\gamma = 10^{-8}$, $\|w\|_2^2 = (10^{16}+1)/4$

Condition number is 10^8

Scanned with CamScanner

f) We previously found,

$$W_0 = \frac{1}{2}\begin{bmatrix} 1 \\ \gamma^{-1} \end{bmatrix}$$

$$W_C = \Sigma^{-1} U^T \begin{bmatrix} \epsilon \\ 0 \\ 0 \\ 0 \end{bmatrix}$$ which simplifies to

$$W_C = \begin{bmatrix} \epsilon/2 \\ \epsilon/2\gamma \end{bmatrix}$$

i. for $\epsilon = 0.01$ & $\gamma = 0.1$,

$$\|W_C\|_2^2 = \frac{10^{-4}}{4} 101 \approx 0.0025$$

ii) for $\epsilon = 0.01$ & $\gamma = 10^{-8}$

$$\|W_C\|_2^2 = \frac{10^{-4}}{4}(10^{6}+1) \approx 25 \times 10^{9}$$

g) $\min_w \|U\Sigma w - y\|_2^2 + \lambda \|w\|_2^2$

implies $w = (\Sigma^2 + \lambda I)^{-1} \Sigma U y$ so

$$w_o = D \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

& $w_c = D \begin{bmatrix} \epsilon \\ \epsilon \end{bmatrix}$ where $D = (\Sigma^2 + \lambda I)^{-1}$

$$\Sigma = \begin{bmatrix} \frac{2}{4+\lambda} & 0 \\ 0 & \frac{2\gamma}{4\gamma^2+\lambda} \end{bmatrix}$$

Thus, $w_o = \begin{bmatrix} \frac{4}{4+\lambda} \\ \frac{4\gamma}{4\gamma^2+\lambda} \end{bmatrix}$ & $w_c = \begin{bmatrix} \frac{2\epsilon}{4+\lambda} \\ \frac{2\gamma\epsilon}{4\gamma^2+\lambda} \end{bmatrix}$

i) $\gamma = 0.1$: $\|w_o\|_2^2 = 9.1$,

$\|w_c\|_2^2 = 2.28 \times 10^{-4}$

ii) $\gamma = 10^{-8}$: $\|w_o\|_2^2 \approx 0.95$,

$\|w_c\|_2^2 \approx 2.38 \times 10^{-5}$