Activity 17 ECE 532.

Ayan Deep Hazra.

1.

a) Given, $A^T A A^T + \lambda A^T$

$$A^T(AA^T + \lambda I) = (A^TA + \lambda I)A^T$$

$$(A^TA + \lambda I)^{-1} A^T \cancel{(AA^T + \lambda I)(AA^T + \lambda I)^{-1}} =$$

$$\cancel{(A^TA + \lambda I)^{-1}(A^TA + \lambda I)} A^T (AA^T + \lambda I)^{-1}$$

$\therefore$, $(A^TA + \lambda I)^{-1} A^T = A^T (AA^T + \lambda I)^{-1}$

b) Since $A \in \mathbb{R}^{8000 \times 100}$

we have $AA^T \in \mathbb{R}^{8000 \times 8000}$

and $A^TA \in \mathbb{R}^{100 \times 100}$

Thus the $(A^TA + \lambda I)^{-1}A^T$ ~~f~~ formula will calculate inverse faster. This is because operating on a 100×100 matrix will be faster than doing the same operation of 8000×8000 matrix

c) i) $y_i = \text{sign}\{g_i^T w\}$

$y = 100 \times 1 \quad g = 8000 \times 100 \quad w = 8000 \times 1$

$\min_w \|g^T w - y\|_2^2 \Rightarrow w = (A^T A)^{-1} A^T y$

$= (y^T g) g^T y$

As the number of columns outweighs the number of rows, we conclude that the system has no unique solutions due to no linearly independent columns.

ii) $g^T g = 100 \times 100$ as stated.

Thus for Tikhonov we have

$\min_w \|g^T w + \lambda I - y\|_2^2$

$\Rightarrow w = (g^T g + \lambda I)^{-1} g^T y$

Solution.

$(A^T A + \lambda I)^{-1} A^T y$ form is more computationally efficient as $g^T g$ is 100×100

2\. a) Given, $\min_{w} \|z - w\|_2^2 + \lambda \|w\|_2^2$

$$= \min_{w} \left[\sum_{i=1}^{n} (z_i - w_i)^2 + \lambda w_i^2 \right]$$

Clearly the problem is seperable as the ith term does not depend on any other i-index terms.

$$= \sum \min_{w_i} \left[(z_i - w_i)^2 + \lambda w_i^2 \right]$$

$$= \min_{w_1} \left[(z_1 - w_1)^2 + \lambda w_1^2 \right] + \min_{w_2} \left[(z_2 - w_2)^2 + \lambda w_2^2 \right]$$

$$+ \cdots\cdots + \min_{w_n} \left[(z_n - w_n)^2 + \lambda w_n^2 \right]$$

b) Given, $\min_{w} \|z - w\|_2^2 + \lambda \|w\|_1$

$$= \min_{w} \sum_{i=1}^{n} (z_i - w_i)^2 + \lambda |w_i|$$

Clearly the problem is seperable as the the i^{th} term does not depend on any other i-index terms.

$$= \sum \min_{w_i} \left[(z_i - w_i)^2 + \lambda |w_i| \right]$$

$$= \min_{w_1} \left[(z_1 - w_1)^2 + \lambda |w_1| \right] + \min_{w_2} \left[(z_2 - w_2)^2 + \lambda |w_2| \right]$$

$$+ \cdots\cdots + \min_{w_3} \left[(z_3 - w_3)^2 + \lambda |w_3| \right]$$