3a)

$$\hat{y} = \text{sign}(w^T x + w_0)$$

$w^T \rightarrow [1 \cdot n]$ (labels (training)) $[n \cdot n]$ (kernel) $[n \cdot d]$ (training data)

let $n=3$, $p=2$

$$\hat{y} = \left([1 \cdot 3]\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}_{3\cdot 3}[3 \cdot 2]\right)[2 \cdot 1]_{\text{test data}} + w_0$$

$$w_0 = \hat{y} - [1 \cdot 3][3 \cdot 3][3 \cdot 2][2 \cdot 1]$$

kernel

$$k_{3\cdot 3} = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & k(x_1, x_3) \\ k(x_2, x_1) & - & - \\ k(x_3, x_1) & - & - \end{bmatrix}$$

$$k(x_i, x_j) = (1 + x_i^T x_j)^2$$

$$= 1 + 2x_i^T x_j + (x_i^T x_j)^2$$

plug into k

$$w_0 = \hat{y} - y\,k\,x\,x'$$

- $\hat{y}$: pred labels
- y: test labels
- k: 3.3 kernel
- x: 3.2 x
- x': test x