Matrices

$$\left(E - FH^{-1}G\right)^{-1}FH^{-1} \equiv E^{-1}F\left(H - GE^{-1}F\right)^{-1}$$

For ridge regression

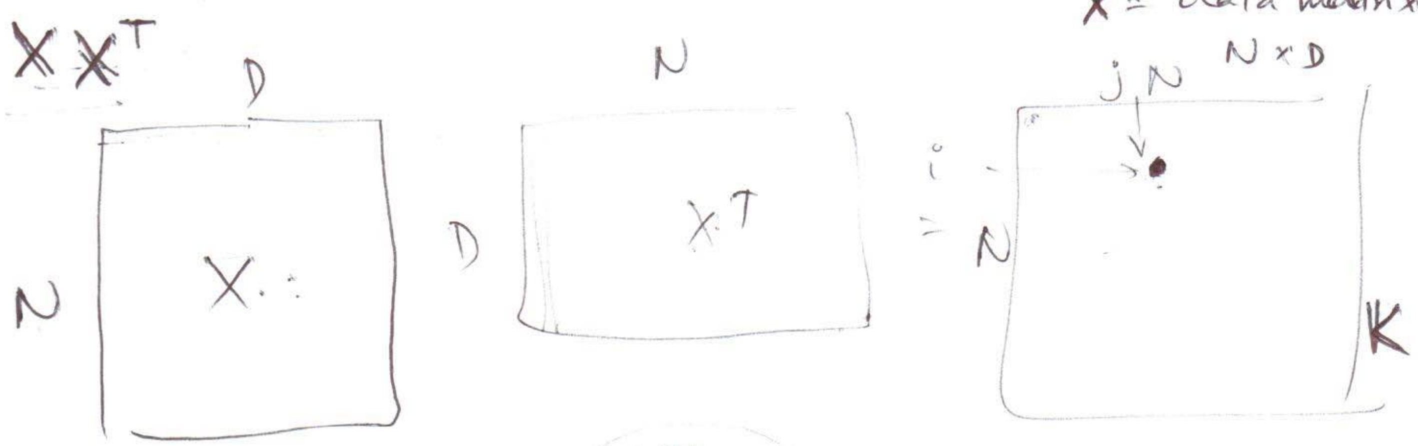
$$y^* = w^T x^* = \left(\left(\lambda I_D + X^T X\right)^{-1} X^T \mathbf{y}\right)^T \mathbf{x}^*$$

test instance

$$= \mathbf{y}^T \left(\lambda I_N + X X^T\right)^{-1} \mathbf{X} x^* \rightarrow \text{test instance } D \times 1$$

$\mathbf{y}$ = vector of targets $(N \times 1)$

$\mathbf{X}$ = data matrix $N \times D$



$$K[i][j] = \mathbf{x}_i^T \mathbf{x}_j$$

dot product between the i^{th} row and j^{th} column.

= dot product between i^{th} and j^{th} data points.

If we replace $x_i^T x_j = k(x_i, x_j)$

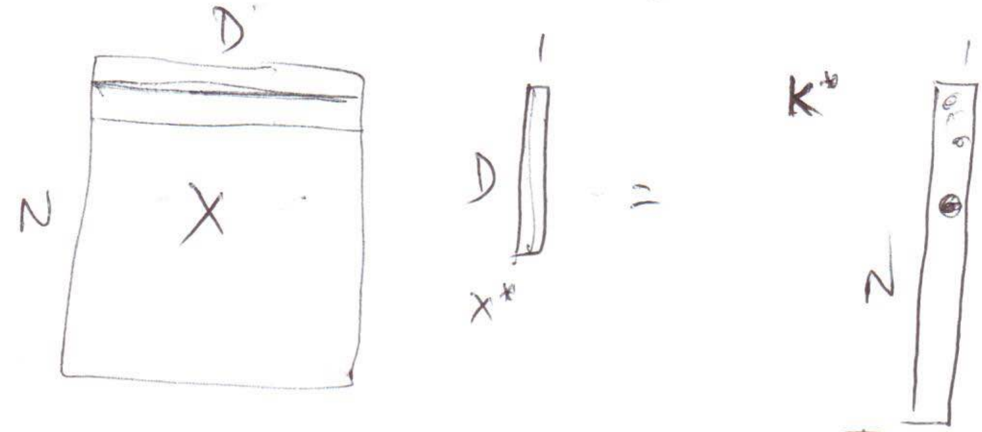
Even if x_i and x_j are not D-length vector

We replace XX^T with K.

$$y^* = y^T(\lambda I_N + \underline{XX^T})^{-1} X x^*$$

$$= y^T(\lambda I_N + K)^{-1} X x^*$$

$\underline{X x^*}$



$$\underline{k^*[i] = x_i^T x^*}$$

$$\equiv k(x_i, x^*)$$

Replace $X x^*$ with k^*

$$\boxed{y^* = y^T(\lambda I_N + K)^{-1} k^*} \rightarrow \underline{\text{Kernel Regression}}$$

where

$$\boxed{\begin{aligned} K[i,j] &= k(x_i, x_j) \\ k^*[i] &= k(x_i, x^*) \end{aligned}}$$

$\underline{k \text{ — kernel function.}}$

---

for any ML method $\rightarrow \mathcal{M}$

If $\underline{x_i, x_j}$ occurs as $\underline{x_i^T x_j}$

Then replace $x_i^T x_j$ with $k(x_i, x_j)$

and you get : $\boxed{\text{Kernel } \mathcal{M}}$

$\underbrace{k(x_i, x_j)}_{\text{Kernel fn.}} \rightarrow \mathbb{R}$ (a scalar value)

data objects.

$(x_1, x_2, \ldots \, x_N)$

$$K[i][j] = k(x_i, x_j)$$

Gram/Kernel Matrix

$$\boxed{K}$$

p.s.d

symmetric.