Ex 1.

1. K is psd. this mean eigen value of $k \geq 0$

$\Rightarrow \Lambda$ is psd. has ≥ 0 diagonal elements.

2. $\min \frac{1}{2} \| y - \beta \|_2^2 + \beta^T K \beta$

$\Rightarrow \min_{\beta} \frac{1}{2} (y-\beta)^T (y-\beta) + \beta^T k \beta$ consider the gradient

minimizer at $\hat{\beta}$ = $\Rightarrow$ $-y + (I+K)\hat{\beta} = 0$

$\Rightarrow \hat{\beta} = (I + k)^{-1} y$ $I+k$ is invertible eigenvalues ≥ 1

$\Rightarrow \hat{\beta} = U(I+\Lambda)^{-1} U^T y$ //

consider λ_i the eigenvalue of K to different K matrices:

If we consider the difference of $\hat{\beta}_1$ and $\hat{\beta}_2$ where

$\hat{\beta}_\ell = U(I+\Lambda_\ell)^{-1} U^T y$,

$$\hat{\beta}_1 - \hat{\beta}_2 = \left(\frac{1}{1+\lambda_{i,1}} - \frac{1}{1+\lambda_{i,2}} \right) u_i u_i^T y$$

$\Rightarrow$ increase of an eigen value of $\lambda_{i,2}$ to $\lambda_{i,1}$ has the above effect for $\hat{\beta}$.

3. i) $k(x, x') = x^T x' \Rightarrow K = XX^T$

let $a \in \mathbb{R}^n$
consider $a^T K a = a^T X X^T a = (X^T a)^T (X^T a) \geq 0$

$\Rightarrow K \succeq 0 \Rightarrow$ mercer kernel ✓

ii) $k(x, x') = x_1 x_1' - x_2 x_2'$

$\Rightarrow K = X_{n \times 2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} X^T_{2 \times n}$

not psd $\Rightarrow$ $\underline{\underline{\text{not}}}$ a mercer kernel
(Ex let $X = I_{2 \times 2}$ $\Rightarrow K = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
n=2 and
not psd //

iii) $k(s, s') = 1$

$\Rightarrow K$ has diagonal 1 and symmetric

$\Rightarrow K$ has the form of a covariance matrix.

let $K = E((X - \bar{X})(X - \bar{X})^T)$

$a \in \mathbb{R}^n$ $a^T K a = a^T E((X - \bar{X})(X - \bar{X})^T) a$

$= E(((X - \bar{X})^T a)^T ((X - \bar{X})^T a))$
≥ 0

≥ 0

$\Rightarrow$ K is a mercer kernel ✓