Q3 (a)

$$\max f(x,y) = xy \quad \text{s.t} \quad x+y^2 \leq 2$$
$$x, y > 0$$

using KKT

Converting to lagrange

$$L(x,y,\lambda) = xy - \lambda(x+y^2-3)$$

Applying KKT

No feasible descent

(1)

(1a) $\nabla_x L = 0 \Rightarrow y - \lambda = 0 \Rightarrow \boxed{y = \lambda}$

(1b) $\nabla_y L = 0 \Rightarrow x - 2\lambda y = 0 \Rightarrow$

$x = 2\lambda y$ or

$$\boxed{x = 2y^2}$$

(2)

$\lambda(x+y^2-2) = 0$ (complementary slackness)

(3) $x+y^2-2 \le 0$ (feasible constraints)

(4) $\lambda \ge 0$ (Positive lagrange multiplier)

<u>case 1</u> $\lambda = 0$

Subs in (1a), (1b),

$y = \lambda = 0$

$x = 2y^2 = 0$

but $x, y > 0$ given

→ Invalid

<u>Case 2</u> $\lambda > 0$

Subs in (2)

$x+y^2-2=0$

Subs (1b) here

$2y^2+y^2-2=0$

$y = \sqrt{\frac{2}{3}}$ $x = 4/3$

$\lambda = \sqrt{2/3} > 0$

So $x, y > 0$ and $\lambda \ge 0$ satisfied optimal values of (x, y)

$= \left(4/3, \sqrt{2/3}\right)$

all 4 KKT conditions satisfied

Q3 (b)

True

SVM draws a hyperplane which maximises the minimum possible geometric margin for all points w.r.t the decision boundary, ie to draw a decision boundary that maximizes the distance b/w the closest set of points of different classes (support vectors)

Q 4

(a) $k(x,x') = c k^{1}(x,x') \qquad c \geq 0$

$k^{(1)}$ is a valid kernal so $\forall z \quad z \in R^{n}$

$z^T k^{(1)} z \geq 0$

$$k(x,x') = c k^{1}(x,x')$$

$$= c \cdot z^T k^{(1)} z \geq 0 \quad (\because c \geq 0 \ \& \ z^T k^{(1)} z \geq 0)$$

$$= z^T c k^{(1)} z \geq 0$$

$\Rightarrow$ $ck^{(1)}$ or $k(x,x')$ is a valid kernal it satisfies the +ve semi-definite property of mercer's theoram.

(b) $k(x,x') = k^{(1)}(x,x') + k^{(2)}(x,x')$

$k^{(1)}$ & $k^{(2)}$ are valid kernels

$z^T k^{(1)} z \geq 0$ & $z^T k^{(2)} z \geq 0$

$\forall z \in R^{n}$

$$\Rightarrow z^T K^{(1)} z + z^T K^{(2)} z \geq 0$$

$$\Rightarrow z^T (K^{(1)} + K^{(2)}) z \geq 0$$

$(K^{(1)} + K^{(2)})$ follows property of +ve semidefinite $K = K^{(1)} + K^{(2)}$ is a valid kernal

(e) $K(x, x') = f(x) K^{(1)}(x, x) f(x')$

$f : R^m \to R$

f maps an n dimensional vector $x \langle x_1 \cdots x_m \rangle$ into a 1D vector $f(x)$ ie a scalar value

Let $f(x) = c_1$ & $f(x') = c_2$

$K^{(1)}$ is a valid kernel.

$$\Rightarrow z^T K^{(1)} z \geq 0 \quad \forall z \in R^n$$

$$\Rightarrow c_1 c_2 z^T K^{(1)} z \geq 0 \quad c_1, c_2 \geq 0$$

$$\Rightarrow z^T (c_1 K^{(1)} c_2) z \geq 0$$

$\Rightarrow \quad z^T \left(f(x) \cdot K^{(1)} \cdot f'(x) \right) z \geq 0$

$\Rightarrow \quad f(x) \cdot K^{(1)} f(x')$ follows +ve semi def matrix property

$\Rightarrow \quad K = f(x) K^{(1)} f(x)$ is a valid kernel

(d) $K(x, x') = K^{(1)}(x, x') \cdot K^{(2)}(x, x')$

$K^{(1)}$ & $K^{(2)}$ are valid kernels

$\therefore$ they corresponds to remapping of i/p to new feature space.

ie $K(x, y) = \Sigma_i \phi_i(x) \phi_i(y)$ for some large (may be ∞)

set of basis functions

so $K^{(1)}(x, x') = \Sigma_i \phi_i^{(1)}(x) \phi_i^{(1)}(x')$

$K^{(1)}(x, x') = \Sigma_j \phi_j^{(2)}(x) \phi_j^{(2)}(x')$

$K^{(1)}(x, x') \cdot K^{(2)}(x, x')$

$= \Sigma_i \phi_i^{(1)}(x) \cdot \phi_i^{(1)}(x') \cdot \Sigma_j \phi_j^{(2)}(x) \cdot \phi_y^{(2)}(x')$

$= \Sigma_i \Sigma_j \phi_i^{(1)}(x) \phi_j^{(2)} x \; \phi_i^{(1)}(x') \cdot \phi_j^{(2)}(x')$

$\therefore$ each ϕ function outputs a scalar
we can define $\phi_K(x) = \phi_i^{(1)}(x) \cdot \phi_j^{(2)}(x)$

similarly for $\phi_K(x')$

finally $K^1(x \cdot x') \cdot K^{(2)}(x, x')$

$$= \Sigma_K \phi_K(x) \cdot \phi_K(x')$$

So the product of 2 kernels creates a function with same invariants which again corresponds to remapping of input to new feature space

$$K(x, x') = K^{(1)}(x, x') \cdot K^{(2)}(x, x')$$

Q5 (a) Classes are not linearly seperable

(b) After mapping to $(1,0,0)$, $(1,-\sqrt{2},1)$, $(1,\sqrt{2},1)$ the points are seperable in 3D

seperating hyperplane is given by weight vector $(0,0,1)$ in new space

(c) Since the pts are support vectors itself we can use langrange multiplier instead of KKT

$$\min_{w,b} \frac{1}{2}|w|_2^2 \quad s.t$$

$$y_i(w^T \phi(x_i)+b) = 1$$

$$i = 1, 2, 3 \dots$$