Question 1

Given: $\max_{\alpha} -1/2 \sum \alpha_i\alpha_j y_i y_j K_{ij} + \sum_i \alpha_i$

s.t $\sum \alpha_i y_i = 0$ & $C \geq \alpha_i \geq 0$

Part A:

we are changing α_i and α_j by keeping the sum $\alpha_i y_i + \alpha_j y_j$ as same as before. Therefore,

$$\alpha_i^{new} y_i + \alpha_j^{new} y_j = \alpha_i^{old} y_i + \alpha_j^{old} y_j$$
$$\alpha_j^{new} y_j = \alpha_i^{old} y_i + \alpha_j^{old} y_j - \alpha_i^{new} y_i$$

multiplying both side by y_j,

$$\alpha_j^{new} y_j^2 = \alpha_i^{old} y_i y_j + \alpha_j^{old} y_j^2 - \alpha_i^{new} y_i y_j$$
$$\alpha_j^{new} = \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new}) y_i y_j \qquad \text{since } y_j^2 = 1$$

Therefore new value of α_j in terms of the new value of α_i and the old values of both,

$$\alpha_j^{new} = \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new}) y_i y_j$$

For constraints on α_i^{new} to make sure that $0 \leq \alpha_i^{new} \leq C$ and $0 \leq \alpha_j^{new} \leq C$

There are two possibilities for earlier equation,

$$\alpha_j^{new} = \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new}) y_i y_j$$

1- If $y_i \neq y_j$ i.e $y_i y_j = -1$

$$\alpha_j^{new} = \alpha_j^{old} - (\alpha_i^{old} - \alpha_i^{new})$$

and since

$$0 \leq \alpha_j^{new} \leq C$$

Therefore,

$$0 \leq \alpha_j^{old} - (\alpha_i^{old} - \alpha_i^{new}) \leq C$$
$$0 \leq \alpha_i^{new} - (\alpha_i^{old} - \alpha_j^{old}) \leq C$$
$$(\alpha_i^{old} - \alpha_j^{old}) \leq \alpha_i^{new} \leq C + (\alpha_i^{old} - \alpha_j^{old})$$

correcting bounds considering $0 \leq \alpha_i^{new} \leq C$ as well,

$$\max(0, \alpha_i^{old} - \alpha_j^{old}) \leq \alpha_i^{new} \leq \min(C, C + \alpha_i^{old} - \alpha_j^{old})$$

2 - If $y_i = y_j$ i.e $y_i y_j = 1$

$$\alpha_j^{new} = \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new})$$

and since

$$0 \leq \alpha_j^{new} \leq C$$

Therefore,

$$0 \le \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new}) \le C$$

multiply by -1,

$$-C \le \alpha_i^{new} - (\alpha_i^{old} + \alpha_j^{old}) \le 0$$

$$(\alpha_i^{old} + \alpha_j^{old}) - C \le \alpha_i^{new} \le (\alpha_i^{old} + \alpha_j^{old})$$

correcting bounds considering $0 \le \alpha_i^{new} \le C$ as well,

$$\max(0, \alpha_i^{old} + \alpha_j^{old} - C) \le \alpha_i^{new} \le \min(C, \alpha_i^{old} + \alpha_j^{old})$$

Part B:

For the formula for the extremum of the objective with respect to α_i^{new}

Our objective is to maximize

$$\sum_i \alpha_i - 1/2 \sum_{i,j} \alpha_i \alpha_j y_i y_j K_{ij} \qquad \text{where } K_{ij} = \langle x_i . x_j \rangle$$

From dual problem we also have:

$$w = \sum \alpha_i y_i x_i$$

and since $f(x) = w^T.x+b$

$$f(x) = (\sum \alpha_i y_i x_i).x + b$$

$$f(x) = \sum \alpha_i y_i (x_i.x) + b$$

For ease of calculation we are assuming i=1, j=2.

So

$$\alpha_1^{new} y_1 + \alpha_2^{new} y_2 = \alpha_1^{old} y_i + \alpha_2^{old} y_j = \zeta \text{ (Constant)}$$

since we are trying to derive the formula for the extremum of objective with respect to α_i therefore we will try to maximize with respect to α_1 .

$$\alpha_1^{new} y_1 + \alpha_2^{new} y_2 = \zeta$$

$$\alpha_2^{new} y_2 = \zeta - \alpha_1^{new} y_1$$

$$\alpha_2^{new} = \zeta - \alpha_1^{new} y_1 y_2$$

$$\alpha_2^{new} = \zeta - \gamma\alpha_1^{new} \qquad \gamma = y_1 y_2$$

our objective in terms of α_1^{new} & α_2^{new}

$$L_D = \sum \alpha_i - 1/2 \sum \alpha_i \alpha_j y_i y_j K_{ij}$$

After substitution we have to create equation in form of

$$L_D = k_1 (\alpha_1^{new})^2 + k_2 (\alpha_1^{new}) + \text{Constant}$$

$$L_D = \alpha_1^{new} + \alpha_2^{new} + const -1/2(\alpha_1^{new}\alpha_1^{new} y_1 y_1 K_{11} + \alpha_2^{new}\alpha_2^{new} y_2 y_2 K_{22}$$
$$+ 2.\alpha_1^{new}\alpha_2^{new} y_1 y_2 K_{12} + 2\,\alpha_1^{new} y_1 x_1 (\textstyle\sum_{k=3} \alpha_k y_k x_k)$$
$$+ 2\,\alpha_2^{new} y_2 x_2 (\textstyle\sum_{k=3} \alpha_k y_k x_k) + const\ )$$

$$L_D = \alpha_1^{new} + \alpha_2^{new} -1/2((\alpha_1^{new})^2 y_1^2 K_{11}) -1/2((\alpha_2^{new})^2 y_2^2 K_{22})$$
$$- \alpha_1^{new}\alpha_2^{new} y_1 y_2 K_{12} - \alpha_1^{new} y_1 x_1 (\textstyle\sum_{k=3} \alpha_k y_k x_k)$$
$$- \alpha_2^{new} y_2 x_2 (\textstyle\sum_{k=3} \alpha_k y_k x_k) + \lambda$$

Since y_1^2 & $y_2^2 = 1$,

$$L_D = \alpha_1^{new} + \alpha_2^{new} -1/2((\alpha_1^{new})^2 K_{11}) -1/2((\alpha_2^{new})^2 K_{22}) - \alpha_1^{new}\alpha_2^{new} y_1 y_2 K_{12}$$
$$- \alpha_1^{new} y_1 x_1 (\textstyle\sum_{k=3} \alpha_k y_k x_k) - \alpha_2^{new} y_2 x_2 (\textstyle\sum_{k=3} \alpha_k y_k x_k) + \lambda$$

$\alpha_i^{new} y_i x_i (\sum_{k=3} \alpha_k y_k x_k)$ can be written as $\alpha_1^{new} y_1 (\sum_{k=3} \alpha_k y_k x_k .x_1)$
and $f(x) = \sum \alpha_i y_i (x_i.x) + b$, ignoring b, $(\sum_{k=3} \alpha_k y_k x_k .x_1) = f(x_1)$

Therefore, after substitution

$$L_D = \alpha_1^{new} + \alpha_2^{new} -1/2((\alpha_1^{new})^2 K_{11}) -1/2((\alpha_2^{new})^2 K_{22}) - \alpha_1^{new}\alpha_2^{new} y_1 y_2 K_{12}$$
$$- \alpha_1^{new} y_1 f(x_1) - \alpha_2^{new} y_2 f(x_2) + \lambda$$

Substituting $\alpha_2^{new} = \zeta - \gamma\alpha_1^{new}$; $\gamma = y_1 y_2$

$$L_D = \alpha_1^{new} + \zeta - \gamma\alpha_1^{new} -1/2((\alpha_1^{new})^2 K_{11}) -1/2((\zeta - \gamma\alpha_1^{new})^2 K_{22})$$
$$- \alpha_1^{new}(\zeta - \gamma\alpha_1^{new}) y_1 y_2 K_{12} - \alpha_1^{new} y_1 f(x_1) - (\zeta - \gamma\alpha_1^{new}) y_2 f(x_2) + \lambda$$

$$L_D = \alpha_1^{new} + \zeta - \gamma\alpha_1^{new} -1/2((\alpha_1^{new})^2 K_{11}) -1/2((\zeta - \gamma\alpha_1^{new})^2 K_{22})$$
$$- \alpha_1^{new}(\zeta - \gamma\alpha_1^{new})\gamma K_{12} - \alpha_1^{new} y_1 f(x_1) - (\zeta - \gamma\alpha_1^{new}) y_2 f(x_2) + \lambda$$

$$L_D = \alpha_1^{new} + \zeta - \gamma\alpha_1^{new} -1/2((\alpha_1^{new})^2 K_{11}) -1/2(\zeta^2 + (\gamma\alpha_1^{new})^2 - 2\zeta\gamma\alpha_1^{new}) K_{22}$$
$$-(\zeta\,\alpha_1^{new} - \gamma(\alpha_1^{new})^2)\,\gamma\, K_{12} - \alpha_1^{new} y_1 f(x_1) - (\zeta - \gamma\alpha_1^{new}) y_2 f(x_2) + \lambda$$

$$L_D = -1/2(\alpha_1^{new})^2 K_{11} -1/2\,\zeta^2 K_{22} -1/2(\gamma\alpha_1^{new})^2 K_{22} + \zeta\gamma\alpha_1^{new} K_{22} + \alpha_1^{new} + \zeta$$
$$- \gamma\alpha_1^{new} - \zeta\,\alpha_1^{new}\gamma K_{12} + \gamma^2(\alpha_1^{new})^2 K_{12} - \alpha_1^{new} y_1 f(x_1) - \zeta y_2 f(x_2)$$
$$+ \gamma\alpha_1^{new} y_2 f(x_2) + \lambda$$

$$L_D = -1/2(\alpha_1^{new})^2 K_{11} -1/2(\gamma\alpha_1^{new})^2 K_{22} + \gamma^2(\alpha_1^{new})^2 K_{12}$$
$$+ \zeta\gamma\alpha_1^{new} K_{22} + \alpha_1^{new} - \gamma\alpha_1^{new} - \zeta\gamma\alpha_1^{new} K_{12} - \alpha_1^{new} y_1 f(x_1) + \gamma\alpha_1^{new} y_2 f(x_2)$$
$$- \zeta y_2 f(x_2) -1/2\,\zeta^2 K_{22} + \zeta + \lambda$$

$$L_D = \left(-1/2K_{11} - 1/2\gamma^2 K_{22} + \gamma^2 K_{12}\right) (\alpha_1^{new})^2$$
$$+ \left(1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2)\right) \alpha_1^{new}$$
$$- \zeta y_2 f(x_2) - 1/2\, \zeta^2 K_{22} + \zeta + \lambda$$

$$L_D = \left(-1/2K_{11} - 1/2K_{22} + K_{12}\right) (\alpha_1^{new})^2$$
$$+ \left(1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2)\right) \alpha_1^{new} + \lambda'$$

This is now in the form of

$$k_1 (\alpha_1^{new})^2 + k_2 (\alpha_1^{new}) + C$$

The First derivative of L_D is

$$\partial L_D / \partial \alpha_1^{new} = 2(-1/2K_{11} - 1/2K_{22} + K_{12})\, \alpha_1^{new} + (1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2))$$

Let $\partial L_D / \partial \alpha_1^{new} = 0$ and we have,

$$(2K_{12} - K_{11} - K_{22})\, \alpha_1^{new} = -(1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2))$$

$$\alpha_1^{new} = -(1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2)) / (2K_{12} - K_{11} - K_{22})$$

Simplifying $1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2)$

$$= 1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + y_1 f(x_2) \qquad \text{since } \gamma y_2 = y_1$$

Since $\alpha_2^{new} + \gamma\alpha_1^{new} = \alpha_2^{old} + \gamma\alpha_1^{old} = \zeta$

$$= 1 - \gamma + (\alpha_2^{old} + \gamma\alpha_1^{old})\gamma K_{22} - (\alpha_2^{old} + \gamma\alpha_1^{old})\gamma K_{12} - y_1 f(x_1) + y_1 f(x_2)$$
$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} + \alpha_1^{old} K_{22} - \alpha_2^{old}\gamma K_{12} - \alpha_1^{old} K_{12} - y_1 f(x_1) + y_1 f(x_2)$$

$$f(x_1) = \sum \alpha_i^{new} y_i (x_i . x_1)$$
$$= w^{old}.x_1 - \alpha_1^{old} y_1\, x_1.x_1 - \alpha_2^{old} y_2\, x_2.x_1$$
$$= w^{old}.x_1 - \alpha_1^{old} y_1 K_{11} - \alpha_2^{old} y_2 K_{12}$$
$$= f(x_1)^{old} - \alpha_1^{old} y_1 K_{11} - \alpha_2^{old} y_2 K_{12}$$

$$f(x_2) = \sum \alpha_i^{new} y_i (x_i . x_2)$$
$$= w^{old}.x_2 - \alpha_1^{old} y_1\, x_1.x_2 - \alpha_2^{old} y_2\, x_2.x_2$$
$$= w^{old}.x_2 - \alpha_1^{old} y_1 K_{12} - \alpha_2^{old} y_2 K_{22}$$
$$= f(x_2)^{old} - \alpha_1^{old} y_1 K_{11} - \alpha_2^{old} y_2 K_{12}$$

substituting $f(x_1)$ & $f(x_2)$

$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} + \alpha_1^{old}K_{22} - \alpha_2^{old}\gamma K_{12} - \alpha_1^{old}K_{12} - y_1(f(x_1)^{old} - \alpha_1^{old}y_1K_{11} - \alpha_2^{old}y_2K_{12}) + y_1(f(x_2)^{old} - \alpha_1^{old}y_1K_{12} - \alpha_2^{old}y_2K_{22})$$

$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} + \alpha_1^{old}K_{22} - \alpha_2^{old}\gamma K_{12} - \alpha_1^{old}K_{12} - y_1 f(x_1)^{old} + y_1\alpha_1^{old}y_1K_{11} + y_1\alpha_2^{old}y_2K_{12} + y_1 f(x_2)^{old} - y_1\alpha_1^{old}y_1K_{12} - y_1\alpha_2^{old}y_2K_{22}$$

$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} + \alpha_1^{old}K_{22} - \alpha_2^{old}\gamma K_{12} - \alpha_1^{old}K_{12} + \alpha_1^{old}K_{11} + y_1y_2\alpha_2^{old}K_{12} - \alpha_1^{old}K_{12} - y_1y_2\alpha_2^{old}K_{22} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} + \alpha_1^{old}K_{22} - \alpha_2^{old}\gamma K_{12} - \alpha_1^{old}K_{12} + \alpha_1^{old}K_{11} + \gamma\alpha_2^{old}K_{12} - \alpha_1^{old}K_{12} - \gamma\alpha_2^{old}K_{22} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= 1 - \gamma + \alpha_2^{old}\gamma K_{22} - \alpha_2^{old}\gamma K_{12} + \gamma\alpha_2^{old}K_{12} - \gamma\alpha_2^{old}K_{22} + \alpha_1^{old}K_{22} - \alpha_1^{old}K_{12} + \alpha_1^{old}K_{11} - \alpha_1^{old}K_{12} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= 1 - \gamma + (\gamma K_{22} - \gamma K_{12} + \gamma K_{12} - \gamma K_{22})\alpha_2^{old} + (K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= 1 - \gamma + (K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= y_1^2 - y_1y_2 + (K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1(f(x_2)^{old} - f(x_1)^{old})$$

$$= (K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1(y_1 - y_2 + f(x_2)^{old} - f(x_1)^{old})$$

$$= (K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1((f(x_2)^{old} - y_2) - (f(x_1)^{old} - y_1))$$

since,

$$\alpha_1^{new} = -(1 - \gamma + \zeta\gamma K_{22} - \zeta\gamma K_{12} - y_1 f(x_1) + \gamma y_2 f(x_2)) / (2K_{12} - K_{11} - K_{22})$$

$$\alpha_1^{new} = -(K_{22} - 2K_{12} + K_{11})\alpha_1^{old} + y_1((f(x_2)^{old} - y_2) - (f(x_1)^{old} - y_1)) / (2K_{12} - K_{11} - K_{22})$$

$$\alpha_1^{new} = \alpha_1^{old} - y_1((f(x_2)^{old} - y_2) - (f(x_1)^{old} - y_1)) / (2K_{12} - K_{11} - K_{22})$$

Generalizing for i & j,

$$\alpha_i^{new} = \alpha_i^{old} - y_i((f(x_j)^{old} - y_j) - (f(x_i)^{old} - y_i)) / (2K_{ij} - K_{ii} - K_{jj})$$

$$\alpha_j^{new} = \alpha_j^{old} + (\alpha_i^{old} - \alpha_i^{new})y_iy_j$$

Part C:

```
function extremum(i,j):
    αi = Lagrange multiplier for i
    αj = Lagrange multiplier for j
    f(xi) = SVM output for i
    yi = value for i
    f(xi) = SVM output for j
    yi = value for j
    Kij = Kernel(i,j)
    Kii = Kernel(i,i)
    Kjj = Kernel(j,j)
    αi^new = αi - yi ((f(xj)-yj)-(f(xi)-yi) )/(2Kij-Kii -Kjj )
    αj^new = αj + (αi - αi^new )yiyj
    update Lagrange multipliers
    update Weight Vector with new αi & αj
end function
```