1. $\mathcal{L}^0(\theta_1, \theta_2) = \frac{(\theta_1 - 3)^2}{2} + \frac{(\theta_2 - 1)^2}{3} + \lambda_0 \left(\sum_{i=1}^{n} \theta_i^0 \right)$

$\mathcal{L}^1(\theta_1, \theta_2) = \frac{(\theta_1 - 3)^2}{2} + \frac{(\theta_2 - 1)^2}{3} + \lambda_1 (|\theta_1| + |\theta_2|)$

$\mathcal{L}^2(\theta_1, \theta_2) = \frac{(\theta_1 - 3)^2}{2} + \frac{(\theta_2 - 1)^2}{3} + \lambda_2 (\theta_1^2 + \theta_2^2)$

2. $k_{ij} = (\gamma + x_i^T x_j)^2$

$= (x_i x_j)^2 + 2\gamma x_i^T x_j + \gamma^2$

$k\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}\right) = (x_1 x_1' + \ldots + x_n x_n')^2$

$= (x_1 x_1')^2 + 2(x_1 x_1')(x_2 x_2') + \ldots + (x_n x_n')^2$

$= (x_1^2 \quad \sqrt{2}\, x_1 x_2 \quad \ldots \quad x_n^2) \begin{pmatrix} x_1'^2 \\ \sqrt{2}\, x_1' x_2' \\ \vdots \\ x_n'^2 \end{pmatrix}$

$\Phi(x) = (x_1^2, \sqrt{2}\, x_1 x_2 \ldots x_n^2)$

$k\left(\gamma + \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}\right) = (\gamma + x_1 x_1' + \ldots + x_n x_n'^{\;2})$

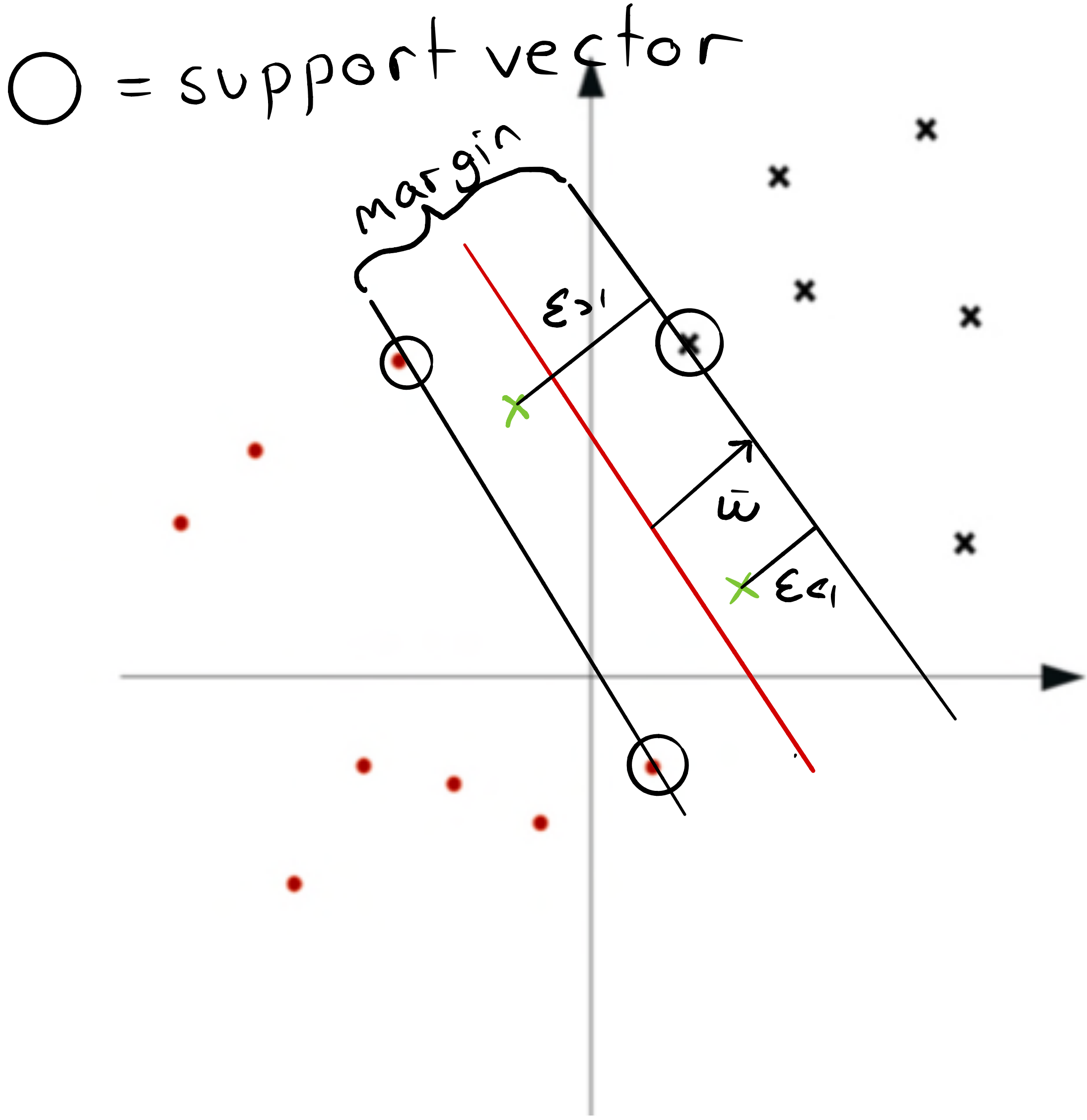
$= \gamma^2 + (x_1 x_1')^2 + 2(x_1 x_1')(x_2 x_2') + \gamma 2(x_1 x_1') + \ldots + (x_n x_n')^2$

$= (\gamma, x_1^2, \sqrt{2}\, x_1 x_2, \sqrt{\gamma}\sqrt{2}\, x_1 \ldots x_n^2) \begin{pmatrix} \gamma \\ x_1'^2 \\ \sqrt{2}\, x_1' x_2' \\ \sqrt{2}\sqrt{\gamma}\, x_1' \\ \vdots \\ x_n'^2 \end{pmatrix}$

$\Phi(x) = (\gamma, x_1^2, \sqrt{2}\, x_1 x_2, \sqrt{\gamma}\sqrt{2}\, x_1 \ldots x_n^2)$

We can see that γ has the higher exponential multiplier on lower order terms. Thus it can be used to increase or decrease the weight of lower order terms

3.



b) $\mathcal{L}(w, \alpha, b) = [\frac{1}{2} w^T w - \sum_i \alpha_i ((w^T x_i + b) y_i - 1)]$

Calculate gradients w.r.t w and b

$\nabla_b = 0 - \sum_i \alpha_i ((0+1) y_i - 0)$

$= \sum_i \alpha_i y_i$

$\nabla_w = w - \sum_i \alpha_i (x_i + 0) y_i - 0)$

$= w - \sum_i \alpha_i x_i y_i$

$\nabla_b = 0$

$\sum_i \alpha_i y_i = 0$

$$\nabla_w = 0$$

$$w = \sum_i \alpha_i x_i y_i$$

$$f(\alpha) = \frac{1}{2}\left(\sum_i \alpha_i y_i x_i\right)^T\left(\sum_i \alpha_i y_i x_i\right) - \sum_i \alpha_i \left(\left(\left(\sum_i \alpha_i x_i y_i + \sum_i \alpha_i y_i\right) y_i x_i - 1\right)\right)$$

$$= \frac{1}{2}\left(\sum_i \alpha_i y_i x_i\right)^T\left(\sum_i \alpha_i y_i x_i\right) - \Sigma_i\left(\alpha_i \left(\sum_i \alpha_i x_i y_i\right) y_i x_i\right) + \sum_i \alpha_i$$

$$= \frac{1}{2}\left(\sum_i \alpha_i y_i x_i\right)^T\left(\sum_i \alpha_i y_i x_i\right) - \Sigma_i\left(\alpha_i x_i y_i\right) \sum_i \alpha_i x_i y_i\big) + \sum_i \alpha_i$$

$$= -\frac{1}{2}\left(\sum_i \alpha_i x_i y_i\right)^T \Sigma_i\left(\alpha_i y_i x_i\right) + \sum_i \alpha_i$$

$$= \sum_i \alpha_i - \frac{1}{2}\left(\sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j\right)$$