CS-E4830
Assignment 3
Question 1

Let $f(\cdot) = \|\cdot\| : \mathbb{R}^n \to \mathbb{R}^+$.

So, f is a norm function on $\mathbb{R}^n$, which by definition is non-negative and homogeneous.

f is a convex function iff it is sub-additive.

Proof.

$$f(\theta x + (1-\theta)y) \leq f(\theta x) + f((1-\theta)y) \quad \text{// triangle ineq.}$$
$$= \theta f(x) + (1-\theta) f(y) \quad \text{// abs. homogeneity}$$

$\blacksquare$

Question 2

Claim:
A set C is convex and $x_i \in C$, for $i = \{1,2,3\}$.
$\theta_i \geq 0$, for $i = \{1,2,3\}$ and since $\theta_1 + \theta_2 + \theta_3 = 1$,
$\theta_i \leq 1$, for $i = \{1,2,3\}$.

Then also $\sum_{i=1}^{3} \theta_i x_i \in C$.

Proof.

We can prove the above claim by induction.

$$\sum_{i=1}^{3} \theta_i x_i = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3$$
$$= \theta_1 x_1 + \sum_{i=2}^{3} \theta_i \left(\frac{\theta_2}{\theta_2+\theta_3} x_2 + \frac{\theta_3}{\theta_2+\theta_3} x_3 \right)$$

$$= \theta_1 x_1 + \sum_{i=2}^{3} \theta_i \bar{x} \qquad \text{// By induction } \bar{x} \in C$$

$$= \theta_1 x_1 + (1 - \theta_1) \bar{x} \qquad \blacksquare$$

Question 3

a) $L_P = \frac{1}{2} w^T w + C \sum_i \xi_i - \sum_i \alpha_i (y_i (w^T \phi(x_i) + b) - 1 + \xi_i) - \sum_i \mu_i \xi_i$

$$= \frac{1}{2} w^T w + C \sum_i \xi_i - \sum_i \alpha_i y_i w^T \phi(x_i) - \sum_i \alpha_i y_i b + \sum_i \alpha_i$$

$$- \sum_i \alpha_i \xi_i - \sum_i \mu_i \xi_i$$

$$= \frac{1}{2} w^T w - \sum_i \alpha_i y_i w^T \phi(x_i) - \sum_i \alpha_i y_i b + \sum_i (C - (\alpha_i + \mu_i)) \xi_i + \sum_i \alpha_i$$

b) $\frac{\partial L_P}{\partial w} = w - \sum_i \alpha_i y_i \phi(x_i) = 0$

$$w = \sum_i \alpha_i y_i \phi(x_i)$$

$$\frac{\partial L_P}{\partial b} = - \sum_i \alpha_i y_i = 0$$

$$\sum_i \alpha_i y_i = 0$$

$$\frac{\partial L_P}{\partial \xi_i} = \sum_i C - (\alpha_i + \mu_i) = 0$$

$$C = (\alpha_i + \mu_i)$$

c) We substitute the K.-T.-K. conditions into primal Lagrangian, to get the dual form of C-SVM as follows:

$$L_D = \frac{1}{2} w^T w - w^T \cdot \underbrace{\sum_i \alpha_i y_i \phi(x_i)}_{=w} - \underbrace{\sum_i \alpha_i y_i}_{=0} b + \sum_i ((C - (\alpha_i + \mu_i)) \xi_i$$

$$+ \sum_i \alpha_i$$

$$= \frac{1}{2} w^T w - w^T w + \sum_i ((\alpha_i + \mu_i) - (\alpha_i + \mu_i)) \xi_i) + \sum_i \alpha_i$$

$$= \sum_i \alpha_i - \frac{1}{2} w^T w$$

$$= \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j y_i y_j \alpha_i \alpha_j \phi(x_i) \phi(x_j)$$

$$= \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$

→ The dual form of the C-SVM problem is posed as

$$\max_{0 \le \alpha_i \le C} \sum_i \alpha_i - \frac{1}{2} \sum_i \sum_j y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$