13.

The soft margin SVM dual is defined as:

$$\min_{\alpha} \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^{N} \alpha_n$$
$$\text{subject to} \sum_{n=1}^{N} y_n \alpha_n = 0$$
$$0 \leq \alpha_n \leq C, \quad n = 1, \cdots, N$$

We're asked to derive a dual of the above problem.

So first we should rewrite the objective function and the constraints into the canonical form:

Let the objective function be $f_0(\alpha)$:

$$f_0(\alpha) = \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^{N} \alpha_n$$

and the constraints be $f_1(\alpha)$ for the equality constraint and $h_j(\alpha)$ for the inequality constraint:

$$f_1(\alpha) = \sum_{n=1}^{N} y_n \alpha_n$$

For $0 \leq \alpha_n \leq C$, we can split it into two constraints:

$$0 \leq \alpha_n \quad n = 1, \cdots, N$$
$$\Rightarrow \ -\alpha_n \leq 0 \quad n = 1, \cdots, N$$

and

$$\alpha_n - C \leq 0 \quad n = 1, \cdots, N$$

So let:

1. $g_n(\alpha_n) = -\alpha_n, \quad n = 1, \cdots, N$
2. $h_n(\alpha_n) = \alpha_n - C, \quad n = 1, \cdots, N$

Therefore, we can construct the Lagrangian:

$$
\begin{aligned}
L(\alpha, \lambda, \mu, \gamma) &= f_0(\alpha) + \lambda f_1(\alpha) + \sum_{n=1}^{N} \mu_n g_n(\alpha) + \sum_{n=1}^{N} \gamma_n h_n(\alpha) \\
&= \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^{N} \alpha_n + \lambda \sum_{n=1}^{N} y_n \alpha_n + \sum_{n=1}^{N} \mu_n(-\alpha_n) + \sum_{n=1}^{N} \gamma_n(\alpha_n - C)
\end{aligned}
$$

We need to compare this $L(\alpha, \lambda, \mu, \gamma)$ with the soft-margin SVM primal, which is:

$$
\begin{aligned}
&\min_{\alpha} \frac{1}{2} \mathbf{w}^T \mathbf{w} - C \sum_{n=1}^{N} \xi_n \\
\text{subject to} \quad & y_n(\mathbf{w}^T \mathbf{z}_n + b) \geq 1 - \xi_n, \quad n = 1, \cdots, N \\
& \xi_n \geq 0, \quad n = 1, \cdots, N
\end{aligned}
$$

To check if the lagrange dual we derived is the same or similar to the primal, we can check if they would give the same optimal solution.

So first we consider the stationary condition of $L(\alpha, \lambda, \mu, \gamma)$ by taking partial derivatives:

$$
\begin{aligned}
\frac{\partial L}{\partial \alpha_n} &= \frac{\partial}{\partial \alpha_n} \left(\frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \alpha_n \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - \sum_{n=1}^{N} \alpha_n + \lambda \sum_{n=1}^{N} y_n \alpha_n + \sum_{n=1}^{N} \mu_n(-\alpha_n) + \sum_{n=1}^{N} \gamma_n(\alpha_n - C) \right) \\
&= \sum_{m=1}^{N} \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - 1 + \lambda y_n + \mu_n - \gamma_n
\end{aligned}
$$

Setting the result to 0 gives:

$$
\sum_{m=1}^{N} \alpha_m y_n y_m \mathbf{z}_n^T \mathbf{z}_m - 1 + \lambda y_n + \mu_n - \gamma_n = 0 \qquad (*)
$$

And the primal constraints must hold, which are:

$$
\sum_{n=1}^{N} y_n \alpha_n = 0 \qquad \text{(1-1)}
$$

$$
0 \leq \alpha_n \leq C, \quad n = 1, \cdots, N \qquad \text{(1-2)}
$$

The dual feasibility requires the lagrange multipliers to satisfy:

$$\lambda \geq 0, \quad \mu_n \geq 0, \quad \gamma_n \geq 0, \quad n = 1, \cdots, N \tag{2}$$

Also, the complementary slackness conditions are:

$$\mu_n(-\alpha_n) = 0, \quad \gamma_n(\alpha_n - C) = 0, \quad n = 1, \cdots, N \tag{3}$$

I think that the problems are similar in some sense, the primal problem generates the solution by directly finding the optimal hyperplane, while the dual of the dual deals with the values of the lagrange multipliers α_n, focuses on the weight (importance) of the support vectors.