Hint to Exercise 5.13

First part:

You can argue directly from (5.9) without considering the involved matrices.

Second part:

The goal is to show that (referring to the notation in (5.26) and (5.27)):

$$y_i - \hat{f}_\lambda^{(-i)}(x_i) = \frac{y_i - \hat{f}_\lambda(x_i)}{1 - S_\lambda(i,i)} \tag{1}$$

To prove this you shall use the first part of the exercise by assuming that the data including x_0 are the full data, while the data without the x_0 corresponds to the "$(-i)$"-case (and you may hence write it as "(-0)").

The key is to show that - assuming $S(\lambda)$ is the $N+1 \times N+1$ matrix which includes the x_0, and where $\hat{f}_\lambda(x_0)$ has the meaning as in the first part of the exercise, -

$$\hat{f}_\lambda(x_0) = \sum_{j=1}^{N} S_{0j}(\lambda) y_j + S_{00}(\lambda)\hat{f}_\lambda(x_0) \tag{2}$$

The result (1) (with $i = 0$) can now be proved from (2). You must then assume that the value of y at x_0 is a general observation y_0, and $\hat{f}_\lambda(x_0)$ should now be interpreted as $\hat{f}_\lambda^{(-0)}(x_0)$.

Extra question

The relation (1) will hold also in many other cases than for smoothing splines, e.g. for usual least squares regression. Why? Other cases?