Ex 7.3 (a)

This exercise has already been solved in Ex 5.13. However, Part (c) of this exercise looks for this to be proved for a more general smoother S rather than the specific case of a cubic smoothing spline. Thus, here we will prove (7.64) for a general smoother $S = X(X^TX + K)^{-1}X^T$ which includes cubic smoothing splines.

We will also assume that $(X^TX+K)^{-1}$ is positive semidefinite.

Notation: $X^{(-i)}$ denotes the matrix X with the i^{th} row removed.

and a few facts we will use:

① $X^TX = X^{(-i)T}X^{(-i)} + x_ix_i^T$ and $X^{(-i)T}y^{(-i)} = X^Ty - x_iy_i$

② The i^{th} diagonal element of S (e.g. S_{ii}) is given by: $S_{ii} = x_i^T(X^TX+K)^{-1}x_i$ and $S_{ii} \geq 0$ by the definition of a positive semi definite matrix on $(X^TX+K)^{-1}$.

③ $\hat{f}(x_i) = x_i^T(X^TX+K)^{-1}X^Ty$

④ The Sherman-Morrison formula (Proved in Ch 6):

$$(A + \lambda uv^T)^{-1} = \left(A^{-1} - \frac{\lambda A^{-1}uv^TA^{-1}}{1+\lambda v^TA^{-1}u}\right)$$

Finally, to begin the proof we will find an expression for $\hat{f}^{(-i)}$.

$$\hat{f}^{(-i)}(x_i) = x_i^T(X^{(-i)T}X^{(-i)} + K)^{-1}X^{(-i)T}y^{(-i)} \quad \text{(by ③)}$$

$$= x_i^T(X^TX - x_ix_i^T + K)^{-1}(X^Ty - x_iy_i) \quad \text{(by ①)}$$

and we can expand this using ④ with: $u = v = x_i$, $\lambda = -1$
and $A = X^TX + K$

$$= x_i^T\left((X^TX+K)^{-1} + \frac{(X^TX+K)^{-1}x_i x_i^T(X^TX+K)^{-1}}{1 - x_i^T(X^TX+K)^{-1}x_i}\right)(X^Ty - x_iy_i)$$

(*) $$= \left(x_i^T(X^TX+K)^{-1} + \frac{S_{ii}\,x_i^T(X^TX+K)^{-1}}{1-S_{ii}}\right)\left(X^Ty - x_iy_i\right) \quad \text{(by ②)}$$

$$= x_i^T(X^TX+K)^{-1}X^Ty + \frac{S_{ii}\,x_i^T(X^TX+K)^{-1}X^Ty}{1-S_{ii}} - x_i^T(X^TX+K)^{-1}x_iy_i - \frac{S_{ii}\,x_i^T(X^TX+K)^{-1}x_iy_i}{1-S_{ii}}$$

$$= \hat{f}(x_i) + \frac{S_{ii}\hat{f}(x_i)}{1-S_{ii}} - S_{ii}y_i - \frac{S_{ii}^2 y_i}{1-S_{ii}}$$

$$= \hat{f}(x_i) - S_{ii}y_i + \frac{S_{ii}(\hat{f}(x_i) - S_{ii}y_i)}{1-S_{ii}}$$

$$= (\hat{f}(x_i) - S_{ii}y_i)\left(1 + \frac{S_{ii}}{1-S_{ii}}\right)$$

$$= \frac{\hat{f}(x_i) - S_{ii}y_i}{1-S_{ii}} \quad \text{(as required)}$$

Since $$y_i - \hat{f}^{(-i)}(x_i) = y_i - \frac{\hat{f}(x_i) - S_{ii}y_i}{1-S_{ii}}$$

$$= \frac{y_i - S_{ii}y_i - \hat{f}(x_i) + S_{ii}y_i}{1-S_{ii}}$$

$$= \frac{y_i - \hat{f}(x_i)}{1-S_{ii}}$$

Ex 7.3 (b)

By ② we know that $x_i^T(X^{(-i)T}X^{(-i)}+K)^{-1}x_i \geq 0$ by definition of a positive semidefinite matrix.

Then, as we did in (*) we may use ④ to expand this expression:

$$x_i^T(X^{(-i)T}X^{(-i)}+K)^{-1}x_i = \left(x_i^T(X^TX+K)^{-1} + \frac{S_{ii}\,x_i^T(X^TX+K)^{-1}}{1-S_{ii}}\right)x_i$$

$$= x_i^T(X^TX+K)^{-1}x_i + \frac{S_{ii}\,x_i^T(X^TX+K)^{-1}x_i}{1-S_{ii}}$$

$$= S_{ii} + \frac{S_{ii}^2}{1-S_{ii}} = \frac{S_{ii}}{1-S_{ii}} \geq 0$$

Bounds to investigate at $S_{ii} \in \overset{(a)}{(-\infty,0)}, \overset{(b)}{[0,1)}, \overset{(c)}{(1,\infty)}$

In (a) and (c) $\frac{S_{ii}}{1-S_{ii}} < 0$, therefore $S_{ii} \in [0,1)$.

This implies that $1 - S_{ii} \in (0,1]$.

And since we proved in part (a) that:

$$y_i - \hat{f}^{(-i)}(x_i) = \frac{y_i - \hat{f}(x_i)}{1-S_{ii}}$$

We may conclude that:

$$|y_i - \hat{f}^{(-i)}(x_i)| \geq |y_i - \hat{f}(x_i)|$$

Ex 7.3 (c)

To Prove (7.64) we have assumed a smoother of the form $\hat{y} = Sy$ where the smoother matrix S is of the form:

$$S = X(X^T X + K)^{-1} X^T$$

Such that K is not a function of y and therefore S is not a fuction of y. in other words S is only a function of the data X and smoothing Parameter(s).

Additionally, we required that $(X^T X + K)^{-1}$ is positive semidefinite (only for part (b)).

These are the general conditions we have assumed to prove that (7.64) holds.