Ex 7.1

Since Ex 7.4 Proves expression (7.21), for this question we will take that as given and Prove that for a linear fit with d basis functions:

$$\frac{2}{N}\sum_{i=1}^{N} \mathrm{cov}(\hat{y}_i, y_i) = \frac{2d}{N}\sigma_\varepsilon^2 \qquad \text{thus Proving (7.24)}$$

Where $Y = f(X) + \varepsilon$ with $\varepsilon \overset{iid}{\sim} N(0, \sigma_\varepsilon^2)$ as usual.

Now, we notice that $\sum_{i=1}^{N} \mathrm{cov}(\hat{y}_i, y_i)$ is just the sum of the diagonal elements of the covariance matrix of the vectors $\hat{y}, y$, explicitly trace(Σ) where $\Sigma = \mathrm{cov}(\hat{y}, y)$

Now we just need to find an expression for Σ.

$$\Sigma = \mathrm{Cov}(\hat{y}, y) = \mathrm{cov}(X(X^TX)^{-1}X^Ty, y) \quad \text{(Since linear model)}$$

$$= X(X^TX)^{-1}X^T\,\mathrm{cov}(y,y)$$

$$= X(X^TX)^{-1}X^T\,\mathrm{Var}(y)$$

$$= X(X^TX)^{-1}X^T\sigma_\varepsilon^2 \quad \text{(see aside)}$$

aside//

$$\mathrm{Var}(y) = \mathbb{E}[(y - \mathbb{E}(y))^2]$$

$$= \mathbb{E}[(f(X) + \varepsilon - \mathbb{E}[f(X) + \varepsilon])^2]$$

$$= \mathbb{E}[(f(X) + \varepsilon - f(X))^2]$$

$$= \mathbb{E}[\varepsilon^2] = \sigma_\varepsilon^2$$

Now we may use this expression to solve the original Problem e.g.

$$\frac{2}{N}\,\mathrm{trace}(\Sigma)$$

$$= \frac{2}{N}\,\mathrm{trace}(X(X^TX)^{-1}X^T\sigma_\varepsilon^2)$$

$$= \frac{2\sigma_\varepsilon^2}{N} \operatorname{trace}\left(X(X^T X)^{-1} X^T\right) \qquad (\text{since } \sigma_\varepsilon^2 \text{ constant})$$

$$= \frac{2\sigma_\varepsilon^2}{N} \operatorname{trace}\left((X^T X)^{-1}(X^T X)\right) \qquad (\operatorname{trace}(AB) = \operatorname{trace}(BA))$$

$$= \frac{2\sigma_\varepsilon^2}{N} \operatorname{trace}(\mathbb{I}_d)$$

$$= 2 \cdot \frac{d}{N} \sigma_\varepsilon^2 \qquad (\text{as required})$$