# Notes on the MSE calculation

Jeffrey J. Early

October 8, 2018

Using our definitions,

$$
\begin{aligned}
\mathrm{MSE}(\lambda) =& \frac{1}{N} || \left( \mathbf{S}_\lambda - I \right) \mathbf{x} ||^2 + \frac{2\sigma^2}{N} \operatorname{Tr} \mathbf{S}_\lambda - \sigma^2 && (1) \\
=& \left( 1 - \frac{1}{d_{\mathrm{var}}} \right) \sigma^2 + \frac{2\sigma^2}{d_{\mathrm{mean}}} - \sigma^2 && (2) \\
=& \sigma^2 \left( \frac{2}{d_{\mathrm{mean}}} - \frac{1}{d_{\mathrm{var}}} \right) && (3)
\end{aligned}
$$

If we can assume that the DOF are the same, then we find that,

$$
\mathrm{MSE} = \frac{\sigma^2}{d} \tag{4}
$$

which is a fairly intuitive result.

Key to note here is that the tension parameter in terms of degrees of freedom, has a minimum of $d = 1$, and starts to asymptote much past $d = 10$. So the range of tension parameters to consider is actually quite small.

Ultimately we want to reduce the MSE of all order of the spline fit. If the higher order information is all garbage, then we should know that.

Roughly speaking, this means we simply compute the MSE for velocity, acceleration, etc.

$$\begin{aligned}
\text{MSE}(\lambda) =& \frac{1}{N}||\mathbf{S}_\lambda \mathbf{x} - \mathbf{x}_{\text{true}}||^2 &(5)\\
=& \frac{1}{N}\left[\mathbf{S}^i{}_j x^j - g^i\right]^2 &(6)\\
=& \frac{1}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j + \mathbf{S}^i{}_j \epsilon^j\right]^2 &(7)\\
=& \frac{1}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j + \mathbf{S}^i{}_j \epsilon^j\right]^2 &(8)\\
=& \frac{1}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]^2 + \frac{1}{N}\left[\mathbf{S}^i{}_j \epsilon^j\right]^2 \\
&+ \frac{2}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]\left[\mathbf{S}^i{}_j \epsilon^j\right] &(9)
\end{aligned}$$

Consider,

$$\begin{aligned}
\left[\mathbf{S}^i{}_j \epsilon^j\right]^2 =& \mathbf{S}^i{}_j \epsilon^j \mathbf{S}^i{}_k \epsilon^k &(10)\\
=& \mathbf{S}^i{}_j \mathbf{S}^i{}_k \epsilon^j \epsilon^k &(11)\\
=& \mathbf{S}^i{}_j \mathbf{S}^i{}_j \epsilon^j \epsilon^j &(12)\\
=& \sigma^2 \mathbf{S}^i{}_j \mathbf{S}^i{}_j &(13)\\
=& E\left[\sum_i \left(\mathbf{S}^i{}_j \epsilon^j\right)^2\right] &(14)
\end{aligned}$$

So this is the sum of the square of all the components in the matrix, which is the the trace of the of the matrix. Now, under expectation we have that,

$$\begin{aligned}
\text{MSE}(\lambda) =& \frac{1}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]^2 + \frac{\sigma^2}{N}\mathbf{S}^i{}_j \mathbf{S}^j{}_i &(15)\\
=& \frac{1}{N}\left[\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right)\left(x^j - \epsilon^j\right)\right]^2 + \frac{\sigma^2}{N}\mathbf{S}^i{}_j \mathbf{S}^j{}_i &(16)
\end{aligned}$$

If we include a derivative in there,

$$\begin{aligned}
\text{MSE}(\lambda) =& \frac{1}{N}\left[D^k{}_i\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]^2 + \frac{1}{N}\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2 &(17)\\
&+ \frac{2}{N}\left[D^k{}_i\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right] &(18)
\end{aligned}$$

The last term vanishes under expectation, as usual. The second to last term looks the same as before, in the sense that the matrix operation and the derivative can be combined, and so nothing really changed.

$$\begin{aligned}
\mathrm{MSE}(\lambda) =& \frac{1}{N}\left[D^k{}_i\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right) g^j\right]^2 + \frac{1}{N}\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2 && (19)\\
=& \frac{1}{N}\left[D^k{}_i\left(\mathbf{S}^i{}_j - \mathbf{I}^i{}_j\right)\left(x^j - \epsilon^j\right)\right]^2 + \frac{1}{N}\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2 \\
=& \frac{1}{N}\left[A^k{}_j\left(x^j - \epsilon^j\right)\right]^2 + \frac{1}{N}\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2 && (20)\\
=& \frac{1}{N}\left[A^k{}_j x^j - A^k{}_j \epsilon^j\right]^2 + \frac{1}{N}\left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2 && (21)\\
=& \left[A^k{}_j x^j\right]^2 + \left[A^k{}_j \epsilon^j\right]^2 - 2\left[A^k{}_j x^j\right]\left[A^k{}_j \epsilon^j\right] && (22)\\
& + \left[D^k{}_i \mathbf{S}^i{}_j \epsilon^j\right]^2
\end{aligned}$$

Ultimately I get that,

$$MSE(\lambda) = ||D(I-S)x||^2 - \sigma^2 \operatorname{Tr}(D - DS)^2 + \sigma^2 \operatorname{Tr}(DS)^2 \qquad (23)$$

# 1 Unknown error

Just eliminate $d$ from the sample variance and sample mean equations, so that,

$$\left(1 - \frac{1}{N}\operatorname{Tr}(\mathbf{S}_\lambda)\right)\sigma^2 = \frac{1}{N}||\left(\mathbf{I} - \mathbf{S}_\lambda\right)\mathbf{x}||^2 \qquad (24)$$

or

$$\sigma^2 = \frac{\frac{1}{N}||\left(\mathbf{I} - \mathbf{S}_\lambda\right)\mathbf{x}||^2}{1 - \frac{1}{N}\operatorname{Tr}(\mathbf{S}_\lambda)} \qquad (25)$$

Now the MSE becomes,

$$\begin{aligned}
\mathrm{MSE}(\lambda) =& \frac{1}{N}||\left(\mathbf{S}_\lambda - I\right)\mathbf{x}||^2 \\
& + \left(\frac{\frac{1}{N}||\left(\mathbf{I} - \mathbf{S}_\lambda\right)\mathbf{x}||^2}{1 - \frac{1}{N}\operatorname{Tr}(\mathbf{S}_\lambda)}\right)\left(\frac{2}{N}\operatorname{Tr}\mathbf{S}_\lambda - 1\right) && (26)\\
=& \frac{1}{N}||\left(\mathbf{I} - \mathbf{S}_\lambda\right)\mathbf{x}||^2\left(\frac{\frac{1}{N}\operatorname{Tr}\mathbf{S}_\lambda}{1 - \frac{1}{N}\operatorname{Tr}(\mathbf{S}_\lambda)}\right) && (27)
\end{aligned}$$

Alternatively, if we normalize the MSE by $\sigma^2$, then

$$\frac{1}{\sigma^2}\mathrm{MSE}(\lambda) = \frac{1}{N}\operatorname{Tr}(\mathbf{S}_\lambda) \qquad (28)$$

which, duh, should have been obvious from my earlier calculation that the MSE is $\sigma^2/d$.

If the error $\sigma$ is not known, we can estimate it. For each value of sigma.

At one extreme the signal could be all noise

$$\sigma_{\max} = \sqrt{\frac{1}{N}\sum(x_i - \bar{x})^2} \tag{29}$$

and thus our total degrees of freedom is $d = N$. At the other extreme, the noise is effectively zero, in that it is much less than the mean distance traversed,

$$\sigma_{\min} = \frac{1}{50}\sqrt{\frac{1}{N-1}\sum(x_i - x_{i-1})^2} \tag{30}$$

which is equivalent to setting $\Gamma = 0.1$