1 Effect of regularisation

For the unregularised case, we have that

$$\begin{aligned}
\frac{\partial E}{\partial w_i} = \sum_{n=1}^{N}\left[\sum_{j=0}^{M} w_j x_n^j - t_n\right] x_n^i = 0 \\
\sum_{n=1}^{N}\left[\sum_{j=0}^{M} w_j x_n^{i+j}\right] = \sum_{n=1}^{N} t_n x_n^i \\
\sum_{j=0}^{M}\left[w_j \sum_{n=1}^{N} x_n^{i+j}\right] = \sum_{n=1}^{N} t_n x_n^i \\
\sum_{j=0}^{M} w_j A_{ij} = T_i
\end{aligned} \tag{1}$$

With regularisation, we get

$$\begin{aligned}
\frac{\partial E}{\partial w_i} = \sum_{n=1}^{N}\left[\sum_{j=0}^{M} w_j x_n^j - t_n\right] x_n^i + \lambda w_i = 0 \\
\sum_{n=1}^{N}\left[\sum_{j=0}^{M} w_j (x_n^{i+j} + \delta_{ij} w_j)\right] = \sum_{n=1}^{N} t_n x_n^i \\
\sum_{j=0}^{M}\left[w_j \sum_{n=1}^{N} x_n^{i+j} + \delta_{ij}\right] = \sum_{n=1}^{N} t_n x_n^i \\
\sum_{j=0}^{M} w_j (A_{ij} + \lambda I) = T_i
\end{aligned} \tag{2}$$

This also shows that A being singular is solved by this form of regularisation, since it adds a constant along the diagonal thus making A invertible. It is not impossible either that this operation would *result* in a singular matrix, in that case, choosing a slightly different value of λ solves this.

A being singular implies that $\exists c_0, ...c_M \in \mathbb{R}^M$ such that:

$$\begin{aligned}
&c_0 A_{1,j} + c_1 A_{2,j} + ... + c_M A_{M+1,j} = 0 \\
\Leftrightarrow &c_0 \left[\sum_{n=1}^{N} x_n^0, \sum_{n=1}^{N} x_n^1, ..., \sum_{n=1}^{N} x_n^M \right] \\
+&c_1 \left[\sum_{n=1}^{N} x_n^1, \sum_{n=1}^{N} x_n^2, ..., \sum_{n=1}^{N} x_n^{M+1} \right] \\
+&... \\
+&c_M \left[\sum_{n=1}^{N} x_n^M, \sum_{n=1}^{N} x_n^{M+1}, ..., \sum_{n=1}^{N} x_n^{2M} \right] = 0 \\
\Leftrightarrow &\left[\sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^j, \sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^{j+1}, ..., \sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^{j+M} \right] = 0
\end{aligned} \tag{3}$$

If the (squared) magnitude of this vector is zero, A is singular:

$$\sum_{l=0}^{M} \left(\sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^{j+l} \right)^2 = 0 \tag{4}$$

Meaning each element must be zero:

$$\sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^j = ... = \sum_{n=1}^{N} \sum_{j=0}^{M} c_j x_n^{j+M} = 0 \tag{5}$$

For any n, this represents a linear equation with $M + 1$ unknowns c_j. The system is underdetermined if $N \leq M$. Working back to our original assumption, this shows that A is singular if there are more degrees of freedom in the fit than there are data points. In addition, if any of the x_n are duplicate, this results in a duplicate row in the system of equations (thus rendering A singular again). In other words, *A is singular if there are more degrees of freedom in the fit than there are independent data points.*

What if $x_n = 0$?