[titlesec]

0.01em

Problem 1.1

We use a slightly better notation to write this problem. Let X be the matrix of the form

$$X = \begin{bmatrix} x_1^0 & x_1^1 & \cdots & x_1^M \\ x_2^0 & x_2^1 & \cdots & x_2^M \\ \vdots & \vdots & \ddots & \vdots \\ x_N^0 & x_N^1 & \cdots & x_N^M \end{bmatrix}, \quad t = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$$

The the problem can be rewritten in the following form:

$$E(w) = \frac{1}{2}\left((Xw-t)^T (Xw-t)\right).$$

Now we differentiate w.r.t w, note that

$$\begin{aligned} E(w+h) &= \frac{1}{2}\left(X(w+h)-t\right)^T \left(X(w+h)-t\right) \\ &= \frac{1}{2}\left((Xw-t)^T + (Xh)^T\right)(Xw-t+Xh) \\ &= \frac{1}{2}\left[(Xw-t)^T (Xw-t) + (Xw-t)^T Xh + (Xh)^T (Xw-t) + (Xh)^T (Xh)\right] \\ &= E(w) + \left\langle (Xw-t)^T, Xh\right\rangle + \frac{1}{2}\langle Xh, Xh\rangle \\ &= E(w) + \left\langle X^T(Xw-t), h\right\rangle + \frac{1}{2}\langle Xh, Xh\rangle. \end{aligned}$$

Note that $\left\langle X^T(Xw-t), h\right\rangle \in \mathrm{Hom}(\mathbb{R}^{M+1}, \mathbb{R})$ and

$$\frac{1}{2}\langle Xh, Xh\rangle \leq \frac{1}{2}\|Xh\|\|Xh\| \leq \frac{C}{2}\|X\|_\infty^2 \|h\| \xrightarrow{\|h\|\to 0} 0,$$

it follows that $\nabla E(w) = X^T(Xw-t)$. Set it to zero and we get

$$X^T(Xw-t) = 0 \iff X^T Xw = X^\top t.$$

So $X^T X$ is the A proposed in the problem.

$$\left[X^T X\right]_{ij} = \sum_{n=1}^N \left(x_n^i x_n^j\right) = \sum_{n=1}^N x_n^{i+j}, \text{ and } \left[X^T t\right]_i = \sum_{n=1}^N x_n^i t_n,$$

as desired.

Problem 1.2

We use the same notation as in the previous problem and still rewrite the loss function in matrix form as follows:

$$\widetilde{E}(w) = \frac{1}{2}\langle Xw-t, Xw-t\rangle + \frac{\lambda}{2}\langle w, w\rangle.$$

Still we differentiate the expression. Note that if we let $\varphi(w) = \frac{\lambda}{2}\langle w, w\rangle$,we have that

$$\begin{aligned} \varphi(w+h) &= \frac{\lambda}{2}(w+h)^T(w+h) \\ &= \frac{\lambda}{2}\left(w^T w + w^T h + h^T x + \|h\|\right) \\ &= \varphi(w) + \langle \lambda w, h\rangle + \frac{\lambda}{2}\|h\|. \end{aligned}$$