Task 4.2

Lemma 1. *For a matrix $A \in \mathbb{R}^{n\times n}$, $w = \frac{1}{n}1_n \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$ the following equality holds:*

$$tr(Awz^\intercal) = z^\intercal Aw$$

Proof. We can ignore the $\frac{1}{n}$, because it commutes with all of our operations. Therefore we show:

$$\operatorname{tr}(A1_n z^\intercal) = z^\intercal A1_n$$

using the well know fact, that the trace operator is invariant under cyclic permutations, which, for two matrices $A, B \in \mathbb{R}^{n\times n}$, implies

$$\operatorname{tr}(AB) = \operatorname{tr}(BA).$$

Therefore

$$\begin{aligned}\operatorname{tr}(A1z^\intercal) &= \operatorname{tr}(1z^\intercal A)\\ &\overset{\star}{=} \sum_{i=1}^{n} (z^\intercal A)_i \overset{\star\star}{=} \underbrace{(z^\intercal A)1_n}_{\text{scalar product}}\end{aligned}$$

$\star$ holds, because (by associativity) $1_n z^\intercal A = 1_n \underbrace{(z^\intercal A)}_{=:a,\text{ row vector}}$, i.e. $1_n z^\intercal A$ can be expressed as a product of the column vector 1 and a, which is just a $n \times n$ matrix M, where each of the n rows is given by a. Therefore the trace (sum of the diagonal elements) is just the sum of the elements of a:

$$\operatorname{tr}(M) = \sum_{i=1}^{n} M_{ii} = \sum_{i=1}^{n} a_i$$

$\star\star$ holds by the definition of the scalar product:

$$\sum_{i=1}^{n} (z^\intercal A)_i = \sum_{i=1}^{n} (z^\intercal A)_i \cdot 1 = \sum_{i=1}^{n} (z^\intercal A)_i (1_n)_i = (z^\intercal A)1_n$$

□

Corollary 2. *For a data matrix X, $w = \frac{1}{n}1_n \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$ the following equality holds:*

$$tr(X^\intercal Xwz^\intercal) = z^\intercal X^\intercal Xw$$

Proof. The claim is the result of lemma 1 for $A = X^\intercal X$. □

Corollary 3. *For a data matrix X, $w = \frac{1}{n}1 \in \mathbb{R}^n$ and $z \in \mathbb{R}^n$ the following equality holds:*

$$tr(zw^\intercal X^\intercal Xwz^\intercal) = z^\intercal zw^\intercal X^\intercal Xwz^\intercal$$

Proof. The claim is the result of lemma 1 for $A = zw^\intercal X^\intercal X$, since $zw^\intercal \in \mathbb{R}^{n\times n}$ and therefore $zw^\intercal \underbrace{X^\intercal X}_{\in\mathbb{R}^{n\times n}} \in \mathbb{R}^{n\times n}$ □