Machine Learning, Spring 2019: Exercise Sheet 4 – Solutions

Problem 1 (linear algebra training). Let $x_1, ..., x_m \in \mathbb{R}^n$ be m linearly independent n-dimensional vectors, and let μ be their mean. Prove that the centered points $\bar{x}_1 = x_1 - \mu, \dots, \bar{x}_m = x_m - \mu$ span an m–1 dimensional subspace of $\mathbb{R}^n$. (Recall that a set $x_1, ..., x_m$ of vectors is called linearly independent if $a_1 x_1 + ... + a_m x_m = \mathbf{0}$ implies $a_1 = ... = a_m = 0$.)

Solution. First we show that $\bar{x}_1, \dots, \bar{x}_m$ are linearly dependent. Using unit combination coefficients $a_i = 1$ for all $1 \le i \le m$, we find that

$$\bar{x}_1 + \cdots + \bar{x}_m = (x_1 - \mu) + \ ... + (x_m - \mu) = \sum_{i=1}^m x_i - m\frac{1}{m}\sum_{i=1}^m x_i = \mathbf{0},$$

hence $\bar{x}_1, \dots, \bar{x}_m$ are linearly dependent and thus span a subspace of dimension less than m.

Next we show that $\bar{x}_1, \dots, \bar{x}_{m-1}$ are linearly independent (then we are done, because then $\bar{x}_1, \dots, \bar{x}_m$ span an m–1 dimensional subspace). Consider a linear combination satisfying $a_1\bar{x}_1 + \cdots + a_{m-1}\bar{x}_{m-1} = 0$. We show that this implies that all a_j are zero:

$$0 = a_1\bar{x}_1 + \cdots + a_{m-1}\bar{x}_{m-1} =$$
$$= \sum_{i=1}^{m-1} a_i x_i - (\sum_{i=1}^{m-1} a_i)\, \mu$$
$$= \sum_{i=1}^{m-1} a_i x_i - (\sum_{i=1}^{m-1} a_i)\, \frac{1}{m}\sum_{i=1}^{m} x_i$$
$$= -\frac{1}{m}(\sum_{i=1}^{m-1} a_i) x_m + \sum_{i=1}^{m-1} \left(a_i - \frac{1}{m}(\sum_{i=1}^{m-1} a_i)\right) x_i$$

$\Rightarrow -\frac{1}{m}(\sum_{i=1}^{m-1} a_i) = 0$ because $x_1, ..., x_m$ are linearly independent and hence $a_i - \frac{1}{m}(\sum_{i=1}^{m-1} a_i) = 0$ for all $1 \le i \le m - 1$, from which $a_i = 0$ for all $1 \le i \le m - 1$ follows.

Problem 2 (A very toy-ish demo of PCA) Assume you have a sample S of four 2-dimensional datapoints from $\mathbb{R}^2$, $S = \{(1,1)', (0,0)', (0,0)', (-1, -1)'\}$. What are the two principal component vectors $\mathbf{u}_1$, $\mathbf{u}_2$ of this dataset?

Solution. The first PC vector is $\mathbf{u}_1 = (1,1)' / \| (1,1)' \| = (1,1)' / \sqrt{2}$. The second PC vector is a unit-length vector orthogonal to $\mathbf{u}_1$, that is $\mathbf{u}_2 = (1, -1)' / \sqrt{2}$. Note: the negatives of $\mathbf{u}_1$, $\mathbf{u}_2$ would likewise qualify as PC vectors, because PC vectors are unique only up to their sign.