Machine Learning, Spring 2019: Exercise Sheet 4

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 $a_1 = ... = a_m = 0$.)

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?