Task 1 - A PCA example

Consider a data set with two data points: $\begin{bmatrix}-1\\0\\0\end{bmatrix}$ and $\begin{bmatrix}1\\0\\0\end{bmatrix}$, $X = \begin{bmatrix}-1 & 1\\0 & 0\\0 & 0\end{bmatrix}$.

1. Recall an eigenvector of a square matrix $A \in \mathbb{R}^{n\times n}$ is defined as a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ such that $A\mathbf{v} = \lambda\mathbf{v}$ where $\lambda \in \mathbb{C}$ is called the eigenvalue of A corresponding to $\mathbf{v}$.

   If A is a diagonal matrix, what are the eigenvectors and corresponding eigenvalues of A?

   Solution: Remember that if A is a real, symmetric matrix, you can decompose it into its eigenvalue decomposition $A = VDV^\top$ where D is a diagonal matrix with the eigenvalues of A and V is an orthogonal matrix containing the corresponding eigenvectors. Then it is easy to see that the eigenvalues of A are just its elements and the unit vectors are the corresponding eigenvectors.

2. Use standard PCA to project the data onto a 1-dimensional subspace. Recall you have to do the following steps: (1) Compute the covariance matrix $S \in \mathbb{R}^{3\times 3}$ of the data. (2) The first principal direction $\mathbf{w} \in \mathbb{R}^3, \|\mathbf{w}\| = 1$ is given by the eigenvector of S corresponding to the largest eigenvalue. (3) Compute the projected data $H = \mathbf{w}^T X$.

   Solution:

   (a) $S = \frac{1}{N}XX^\top = \frac{1}{2}\begin{pmatrix}2 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}$

   (b) $det(S - \lambda I) = (1-\lambda_1)\cdot\lambda_2\cdot\lambda_3$, eigenvalues are $\lambda_1 = 1, \lambda_{2,3} = 0$

   $\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & 0\end{pmatrix}\begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix} = \begin{pmatrix}w_1\\w_2\\w_3\end{pmatrix}$, eigenvector to λ_1 is $w = \begin{pmatrix}1\\0\\0\end{pmatrix}$

   (c) $H = w^\top X = (-1\ 1)$

3. Use linear Kernel PCA to obtain exactly the same result. Recall you have to do the following steps: (1) Compute the kernel matrix $\mathbb{R}^{2\times 2} \ni K = X^T X$ of the data. (2) Compute a linear combination of the data points $\alpha \in \mathbb{R}^2$ as the eigenvector of K corresponding to the largest eigenvalue. (Try to find the eigenvectors of K by educated guessing, use a compute programm, or ask the lecturer) (3) Compute $\mathbf{w} = X\alpha$. You should obtain a scaled version of the standard PCA result. **Solution:**

   (a) $K = X^\top X = \begin{pmatrix}1 & -1\\-1 & 1\end{pmatrix}$

   (b) $det(K - \lambda I) = (1-\lambda_1)\cdot(1-\lambda_2) + 1$, eigenvalues are $\lambda_1 = 2, \lambda_2 = 0$

   (c) $w = X \cdot \alpha = \begin{pmatrix}-1 & 1\\0 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}2\\0\\0\end{pmatrix}$

Task 2 - Covariance matrix and eigenvalues

For each of the three scenarios below, sketch a 2-dimensional Gaussian data set whose covariance matrix has the following two eigenvalues:

1. $\lambda_1 = 1, \lambda_2 = 1$
2. $\lambda_1 = 1, \lambda_2 = 5$
3. $\lambda_1 = 1, \lambda_2 = 0$

Hint: Remember the following fact from the lecture. The variance of a data set projected onto a direction $\mathbf{w} \in \mathbb{R}^D, ||\mathbf{w}|| = 1$ can be computed as $\mathbf{w}^T S\mathbf{w}$, where S denotes the data covariance matrix. If $\mathbf{w}$ is an eigenvector of S with corresponding eigenvalue λ, then $\mathbf{w}^T S\mathbf{w} = \mathbf{w}^T\lambda\mathbf{w} = \lambda\mathbf{w}^T\mathbf{w} = \lambda$.

Task 3 - Covariance matrix and eigenvalues II

1. Consider a square matrix $S \in \mathbb{R}^{D\times D}$, and an eigenvector $\mathbf{v}$ of S with corresponding eigenvalue λ, i.e. $S\mathbf{v} = \lambda\mathbf{v}$.

   Show: $\mathbf{v}$ is also an eigenvector of the scaled matrix $S_c := cS$ for any $c \in \mathbb{R}$.

   Solution:
   $$S\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow \underbrace{cS}_{=S_c}\mathbf{v} = c\lambda\mathbf{v}$$

   Thus, $\mathbf{v}$ is eigenvector of S_c with corresponding eigenvalue $c\lambda$.

2. Consider a centered data set $X \in \mathbb{R}^{D\times N}$ with corresponding covariance matrix $S = \frac{1}{N}XX^T$, and an eigenvector $\mathbf{v}$ of S with corresponding eigenvalue λ, i.e. $S\mathbf{v} = \lambda\mathbf{v}$. Suppose we rotate the data, $X \mapsto UX$ where $U \in \mathbb{R}^{D\times D}$ is a rotation matrix ($UU^T = U^TU = I$).

   Show: λ is an eigenvalue of the covariance matrix of the rotated data set, $S_U = \frac{1}{N}UX(UX)^T$.

   Solution: We have
   $$S_U = \frac{1}{N}UX(UX)^\top = U\frac{1}{N}XX^\top U^\top = USU^\top$$

   Thus (remember $U^\top U = I$)
   $$S\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow US\mathbf{v} = \lambda U\mathbf{v} \Leftrightarrow \underbrace{USU^\top}_{S_U}\underbrace{U\mathbf{v}}_{:=\mathbf{z}} = \lambda\underbrace{U\mathbf{v}}_{:=\mathbf{z}}$$

   Thus, $U\mathbf{v}$ is eigenvector of S_U with corresponding eigenvalue λ.