Final test (Tuesday 17 December 2024, 15:00-17:00 CET)

Elements of Mathematics – Master in Bioinformatics for Health Sciences

1. Consider the matrix $A = BB^t$ where $B^t = [1\ \ 3\ \ 0]$.
   (a) **(0.5 points)** Provide a basis of the column space of A.
   (b) **(0.5 points)** Provide a basis of the null space of A.

2. **(1 point)** Let $H \subset \mathbb{R}^3$ be the vector space of solutions of the equation $x + y + z = 0$. Compute an orthonormal basis of H.

3. Consider the following vectors $u, v, w \in \mathbb{R}^3$ and matrix $A \in \mathbb{R}^{3\times 3}$:
$$u = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \quad v = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} \quad w = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad A = \begin{bmatrix} \cos(\pi/6) & -\sin(\pi/6) & 0 \\ \sin(\pi/6) & \cos(\pi/6) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$
   The matrix of a linear transformation f in coordinates the basis $\mathcal{B} = \{u, v, w\}$ is A.
   (a) **(0.5 points)** Is $\mathcal{B}$ an orthonormal basis of $\mathbb{R}^3$?
   (b) **(0.5 points)** Compute the inverse of the matrix $C = [u|v|w]$ with columns u, v and w.
   (c) **(1 point)** Compute the matrix of f in coordinates of the canonical basis.

4. We are analyzing a gene expression dataset $\mathcal{D}$ with readout for 5 genes (columns) across 100 cells (rows). We carried out the eigendecomposition of Ω, the covariance matrix of $\mathcal{D}$, and got the following:
   - The eigenvalues of Ω are $\lambda_1 = 4$, $\lambda_2 = 3$, $\lambda_3 = 2$, $\lambda_4 = 0.5$, $\lambda_5 = 0.5$
   - The respective eigenvectors are:
$$u_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad u_2 = \frac{1}{\sqrt{6}}\begin{bmatrix} -1 \\ 2 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad u_3 = \frac{1}{2\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 3 \\ 0 \\ 1 \end{bmatrix} \quad u_4 = \frac{1}{2\sqrt{5}}\begin{bmatrix} -1 \\ -1 \\ -1 \\ 4 \\ 1 \end{bmatrix} \quad u_5 = \frac{1}{\sqrt{5}}\begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$
   (a) **(0.25 points)** What is the proportion of variance explained by the first 3 principal components?
   (b) **(0.25 points)** What are the loadings of the second principal component?
   (c) **(0.5 points)** After centering and scaling the gene expression dataset $\mathcal{D}$, one of the cells (rows) has $(1, 0, 0, 0, 0)$ as feature values. What are the scores of this cell in coordinates of the principal directions?

5. **(1 point)** Find the second order Taylor approximation of $f(x) = (1-c)\log(1-x) + c\log(x)$ at $a = 1/2$.

6. If b is a positive real number, the *logarithm of x to base b*, denoted $\log_b(x)$, is a function that satisfies the identity $\log_b(b^z) = z$. The *natural logarithm*, which we simply denote $\log(x)$, is the logarithm to base e (Euler's number).

   (1 point) Use the definition of logarithm to base b and the identity $b^x = e^{\log(b)x}$ to provide an expression of the derivative of $\log_b(x)$.

7. A *quadratic form* is a function $f : \mathbb{R}^n \to \mathbb{R}$ with the form $f(x) = \frac{1}{2}x^t A x + b^t x + c$, where $A \in \mathbb{R}^{n\times n}$ is a matrix, $b \in \mathbb{R}^n$ is a vector and $c \in \mathbb{R}$ is a scalar. Consider the quadratic form $f : \mathbb{R}^2 \to \mathbb{R}$ with
$$A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ -8 \end{bmatrix} \quad c = 0$$
   (a) **(0.5 points)** Compute the gradient of f.
   (b) **(1 point)** Find the critical points f.
   (c) **(0.5 points)** Compute the Hessian of f at the critical points.
   (d) **(1 point)** Classify the critical points of f.