Practice exercises

Euclidean Spaces – Survey

1. Find, using Gram-Schmidt orthogonalization, an orthonormal basis of the subspace of $\mathbb{R}^4$ spanned by $a_1 = (2, 0, 2, 0)$, $a_2 = (1, -1, 1, 0)$, $a_3 = (2, -2, 0, 0)$.

2. Consider the vector subspace $U = \mathcal{L}\{(1, 0, 1)\}$ of $\mathbb{R}^3$, with the usual dot product. Find the implicit equations of the orthogonal complement $U^\perp$. Also, find the orthogonal projection of $(1, 1, 2)$ onto U and $U^\perp$, respectively.

3. Consider the vector subspace U of $\mathbb{R}^3$, furnished with the usual dot product, defined by $U \equiv x + y = 0$. Find the parametric and implicit equations of the orthogonal complement $U^\perp$, and the orthogonal projection of $b = (1, 2, 3)$ onto U.

4. In $\mathbb{R}^4$, with the usual dot product, consider the vector subspace

$$U = \mathcal{L}\left((1, -1, 1, -1), (1, 1, 1, 1), (1, -1, -1, -1)\right).$$

   (a) Find a basis and the implicit equations of the orthogonal complement $U^\perp$ of U.
   (b) Construct an orthonormal basis for U.
   (c) Express $v = (1, 2, 3, 4)$ as a sum of two vectors, one of them belonging to U and the other to $U^\perp$.
   (d) Find the distance from v to U and $U^\perp$.

5. Find an orthogonal basis of the subspace $U \subseteq \mathbb{R}^4$ of equation $x+y-z+w = 0$. Determine the projection of $v = (1, 0, 0, 1)$ onto U and the distance from v to U.

6. As in problem 3, determine the orthogonal projection of $b = (1, 2, 3)$ onto the subspace spanned by the columns of the matrix

$$A = \begin{pmatrix} -1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

   (a) Solve the *normal equations*
   $$A^T A x = A^T b,$$
   and after finding the solution $\overline{x}$ of this system, determine the projection as $p = A\overline{x}$.
   (b) Find the *projection matrix*
   $$P = A(A^T A)^{-1} A^T,$$
   and check that $p = Pb$ is the projection computed in (a).

7. Find the regression line $b = C + Dt$ fort he data $t_1 = 1$, $t_2 = 2$, $t_3 = 3$, $t_4 = 4$, $b_1 = 3$, $b_2 = 4$, $b_3 = 5$, $b_4 = 7$ (i.e. the line of equation $b = C + Dt$ that best approximates, in the least squares sense, the points $(1, 3)$, $(2, 4)$, $(3, 5)$, $(4, 7)$). Denoting the solution $p = Ab$ (the projection of b onto the column space of A, which provides the values at the points t_i of the computed regression line), explain why the following relationships, well-known in statistics, hold:

$$e_1 + e_2 + e_3 + e_4 = 0, \qquad t_1e_1 + t_2e_2 + t_3e_3 + t_4e_4 = 0.$$

8. Consider the problem of calculating a regression parabola $b = Dt + Et^2$ (i.e. we impose that the constant term is $C = 0$) for the data $t_1 = 1$, $t_2 = 2$, $t_3 = 3$, $t_4 = 4$. Let $\overline{x}$ be the solution vector of the normal equations $A^TAx = A^Tb$ and let $e = b - A\overline{x}$ tbe he residues vector. Explain why (regardless of the vector b) it holds that

$$e_1 + 2e_2 + 3e_3 + 4e_4 = 0,$$

although in this case we cannot ensure that the sum of all the residues

$$e_1 + e_2 + e_3 + e_4 = 0.$$

is zero. Deduce other relationship that is also fulfilled by e_1, e_2, e_3, e_4.